International Standard

ISO/IEC 13818-1

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

Part 1: Systems

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

Partie 1: Systèmes

Tenth edition
2025-08

Reference number
ISO/IEC 13818-1:2025(en)

© ISO/IEC 2025



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2025

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Foreword

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

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted.

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

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

This tenth edition cancels and replaces the ninth edition (ISO/IEC 13818-1:2023), which has been technically revised.

The main changes are as follows:

— the usage of certain codec parameters and the definition of certain program element descriptors is clarified;

— in Annex T.4 the interpretation of the optional 'codecs' parameter for the audio layer is clarified;

— a field name is adjusted in the syntax and semantics of the MPEG-H 3D audio descriptor;

— in 2.6.110 the syntax of the MPEG-H 3D audio scene descriptor is corrected and field names are aligned with the names in the specification to which they refer;

— in 2.6.115 some clarification is added for the MPEG-H 3D audio multi-stream descriptor with respect to the semantics of the fields *thisStreamID* and auxiliaryStreamID.

A list of all parts in the ISO/IEC 13818 series can be found on the ISO and IEC websites.

© ISO/IEC 2025 – All rights reserved

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

© ISO/IEC 2025 – All rights reserved

INTERNATIONAL STANDARD ISO/IEC 13818-1
RECOMMENDATION ITU-T H.222.0 v10

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

Summary

This Recommendation | International Standard specifies the system layer of the coding. It was developed in 1994 to principally support the combination and synchronization of video and audio coding methods defined in ISO/IEC 13818 Part 2 (ITU-T H.262) and Part 3. Since 1994, this standard has been extended to support additional video coding specifications (e.g., ISO/IEC 14496-2, ITU-T H.264 | ISO/IEC 14496-10, ITU-T H.265 | ISO/IEC 23008-2 and ITU-T T.800 | ISO/IEC 15444-1 Annex M JPEG 2000 video), audio coding specifications (e.g., ISO/IEC 13818-7 and ISO/IEC 14496-3), system streams (e.g., ISO/IEC 14496-1 and ISO/IEC 15938-1), ISO/IEC 23009-1 dynamic adaptive streaming over HTTP (DASH), ISO/IEC 13818-11 intellectual property management and protection (IPMP) as well as generic metadata. The system layer supports six basic functions:

1) the synchronization of multiple compressed streams on decoding;
2) the interleaving of multiple compressed streams into a single stream;
3) the initialization of buffering for decoding start up;
4) continuous buffer management;
5) time identification; and
6) multiplexing and signalling of various components in a system stream.

Recommendation ITU-T H.222.0 | ISO/IEC 13818-1 multiplexed bit stream is either a transport stream or a program stream. Both streams are constructed from packetized elementary stream (PES) packets and packets containing other necessary information. Both stream types support multiplexing of video and audio compressed streams from one program with a common time base. The transport stream additionally supports the multiplexing of video and audio compressed streams from multiple programs with independent time bases. For almost error-free environments the program stream is generally more appropriate, supporting software processing of program information. The transport stream is more suitable for use in environments where errors are likely.

Either multiplexed bit stream is constructed in two layers: the outermost layer is the system layer, and the innermost is the compression layer. The system layer provides the functions necessary for using one or more compressed data streams in a system. The video and audio parts of this Specification define the compression coding layer for audio and video data. Coding of other types of data is not defined by this Recommendation | International Standard, but is supported by the system layer provided that the other types of data adhere to the constraints defined in this Recommendation | International Standard.

This tenth edition was initially developed as Amd. 1 to the ninth edition and was eventually approved as a new edition.

The main changes are as follows:

– The usage of certain codec parameters and the definition of certain program element descriptors is clarified;
– in clause T.4 the interpretation of the optional 'codecs' parameter for the audio layer is clarified;
– a field name is adjusted in the syntax and semantics of the MPEG-H 3D audio descriptor;
– in 2.6.110 the syntax of the MPEG-H 3D audio scene descriptor is corrected and field names are aligned with the names in the specification to which they refer;
– in 2.6.115 some clarification is added for the MPEG-H 3D audio multi-stream descriptor with respect to the semantics of the fields thisStreamID and auxiliaryStreamID;

© ISO/IEC 2025 – All rights reserved

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T H.222.0	1995-07-10	15	11.1002/1000/1071
1.1	ITU-T H.222.0 (1995) Amd. 1	1996-11-11	16	11.1002/1000/3834
1.2	ITU-T H.222.0 (1995) Amd. 2	1996-11-11	16	11.1002/1000/4096
1.3	ITU-T H.222.0 (1995) Technical Cor. 1	1998-02-06	16	11.1002/1000/4532
1.4	ITU-T H.222.0 (1995) Amd. 3	1998-02-06	16	11.1002/1000/4228
1.5	ITU-T H.222.0 (1995) Amd. 4	1998-02-06	16	11.1002/1000/4229
1.6	ITU-T H.222.0 (1995) Amd. 5	1999-05-27	16	11.1002/1000/4498
1.7	ITU-T H.222.0 (1995) Amd. 6	1999-05-27	16	11.1002/1000/4671
1.8	ITU-T H.222.0 (1995) Amd. 7	2000-02-17	16	11.1002/1000/4926
2.0	ITU-T H.222.0	2000-02-17	16	11.1002/1000/5142
2.1	ITU-T H.222.0 (2000) Technical Cor. 1	2001-03-01	16	11.1002/1000/5419
2.2	ITU-T H.222.0 (2000) Technical Cor. 2	2002-03-29	16	11.1002/1000/5675
2.3	ITU-T H.222.0 (2000) Amd. 1	2002-12-14	16	11.1002/1000/6190
2.4	ITU-T H.222.0 (2000) Amd. 1 Cor. 1	2003-06-29	16	11.1002/1000/6449
2.5	ITU-T H.222.0 (2000) Amd. 2	2003-06-29	16	11.1002/1000/6363
2.6	ITU-T H.222.0 (2000) Amd. 3	2004-03-15	16	11.1002/1000/7208
2.7	ITU-T H.222.0 (2000) Technical Cor. 3	2005-01-08	16	11.1002/1000/7435
2.8	ITU-T H.222.0 (2000) Amd. 4	2005-01-08	16	11.1002/1000/7436
2.9	ITU-T H.222.0 (2000) Amd. 5	2005-01-08	16	11.1002/1000/7437
2.10	ITU-T H.222.0 (2000) Technical Cor. 4	2005-09-13	16	11.1002/1000/8560
3.0	ITU-T H.222.0	2006-05-29	16	11.1002/1000/8802
3.1	ITU-T H.222.0 (2006) Amd. 1	2007-01-13	16	11.1002/1000/9024
3.2	ITU-T H.222.0 (2006) Amd. 2	2007-08-29	16	11.1002/1000/9214
3.3	ITU-T H.222.0 (2006) Cor. 1	2008-06-13	16	11.1002/1000/9471
3.4	ITU-T H.222.0 (2006) Cor. 2	2009-03-16	16	11.1002/1000/9692
3.5	ITU-T H.222.0 (2006) Amd. 3	2009-03-16	16	11.1002/1000/9691
3.6	ITU-T H.222.0 (2006) Cor. 3	2009-12-14	16	11.1002/1000/10621
3.7	ITU-T H.222.0 (2006) Cor. 4	2009-12-14	16	11.1002/1000/10622
3.8	ITU-T H.222.0 (2006) Amd. 4	2009-12-14	16	11.1002/1000/10623
3.9	ITU-T H.222.0 (2006) Amd. 5	2011-05-14	16	11.1002/1000/11287
3.10	ITU-T H.222.0 (2006) Amd. 6	2011-05-14	16	11.1002/1000/11288
4.0	ITU-T H.222.0	2012-06-29	16	11.1002/1000/11655
4.1	ITU-T H.222.0 (2012) Amd. 1	2014-01-13	16	11.1002/1000/12054
4.2	ITU-T H.222.0 (2012) Amd. 2	2014-01-13	16	11.1002/1000/12055
4.3	ITU-T H.222.0 (2012) Amd. 3	2014-01-13	16	11.1002/1000/12056
4.4	ITU-T H.222.0 (2012) Amd. 4	2014-01-13	16	11.1002/1000/12057
4.5	ITU-T H.222.0 (2012) Amd. 5	2014-10-14	16	11.1002/1000/12306
5.0	ITU-T H.222.0	2014-10-14	16	11.1002/1000/12359
5.1	ITU-T H.222.0 (2014) Amd. 1	2015-04-29	16	11.1002/1000/12452
5.2	ITU-T H.222.0 (2014) Amd. 1 Cor. 1	2015-11-29	16	11.1002/1000/12625
5.3	ITU-T H.222.0 (2014) Amd. 2	2015-12-14	16	11.1002/1000/12632
5.4	ITU-T H.222.0 (2014) Amd. 3	2015-12-14	16	11.1002/1000/12633
5.5	ITU-T H.222.0 (2014) Amd. 1 Cor. 2	2016-07-14	16	11.1002/1000/12899
5.5	ITU-T H.222.0 (2014) Cor. 1	2016-07-14	16	11.1002/1000/12903
5.7	ITU-T H.222.0 (2014) Amd. 4	2016-07-14	16	11.1002/1000/12900
5.8	ITU-T H.222.0 (2014) Amd. 5	2016-07-14	16	11.1002/1000/12901
5.9	ITU-T H.222.0 (2014) Amd. 6	2016-07-14	16	11.1002/1000/12902

* To access the Recommendation, type the URL https://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID.

© ISO/IEC 2025 – All rights reserved

5.10	ITU-T H.222.0 (2014) Amd. 3 Cor. 1	2017-03-01	16	11.1002/1000/13184
5.10	ITU-T H.222.0 (2014) Cor. 2	2017-03-01	16	11.1002/1000/13188
5.12	ITU-T H.222.0 (2014) Amd. 7	2017-03-01	16	11.1002/1000/13186
5.13	ITU-T H.222.0 (2014) Amd. 8	2017-03-01	16	11.1002/1000/13187
6.0	ITU-T H.222.0	2017-03-01	16	11.1002/1000/13269
6.1	ITU-T H.222.0 (2017) Amd. 1	2017-12-14	16	11.1002/1000/13431
7.0	ITU-T H.222.0	2018-08-29	16	11.1002/1000/13664
7.1	ITU-T H.222.0 (2018) Cor. 1	2019-11-29	16	11.1002/1000/14106
7.2	ITU-T H.222.0 (2018) Amd. 1	2019-11-29	16	11.1002/1000/14105
8.0	ITU-T H.222.0 v8	2021-06-13	16	11.1002/1000/14658
8.1	ITU-T H.222.0 v8 (2021) Cor. 1	2022-12-14	16	11.1002/1000/15172
8.2	ITU-T H.222.0 v8 (2021) Amd. 1	2022-12-14	16	11.1002/1000/15171
9.0	ITU-T H.222.0 v9	2023-08-29	16	11.1002/1000/15655
10.0	ITU-T H.222.0 v10	2025-04-13	16	11.1002/1000/16266

Keywords

Audiovisual content, multimedia multiplexing, MPEG-2 system, multiplexed bit stream, program stream, transport stream.

© ISO/IEC 2025 – All rights reserved

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, and 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/software copyrights, 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 appropriate ITU-T databases available via the ITU-T website at https://www.itu.int/ITU-T/ipr/.

© ITU 2025

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

© ISO/IEC 2025 – All rights reserved

CONTENTS

Page

SECTION 1 – GENERAL 1
1.1 Scope 1
1.2 Normative references 1
SECTION 2 – TECHNICAL ELEMENTS 4
2.1 Definitions 4
2.2 Symbols and abbreviations 12
2.3 Method of describing bit stream syntax 14
2.4 Transport stream bitstream requirements 15
2.5 Program stream bitstream requirements 63
2.6 Program and program element descriptors 76
2.7 Restrictions on the multiplexed stream semantics 164
2.8 Compatibility with ISO/IEC 11172 169
2.9 Registration of copyright identifiers 169
2.10 Registration of private data format 169
2.11 Carriage of ISO/IEC 14496 data 170
2.12 Carriage of metadata 181
2.13 Carriage of ISO 15938 data 190
2.14 Carriage of Rec. ITU-T H.264 | ISO/IEC 14496-10 video 190
2.15 Carriage of ISO/IEC 14496-17 text streams 206
2.16 Carriage of auxiliary video streams 208
2.17 Carriage of HEVC 208
2.18 Carriage of green access units 222
2.19 Carriage of ISO/IEC 23008-3 MPEG-H 3D audio data 224
2.20 Carriage of Quality Access Units in MPEG-2 sections 226
2.21 Carriage of sample variants 227
2.22 Carriage of Media Orchestration Access Units 228
2.23 Carriage of VVC 228
2.24 Carriage of EVC 233
2.25 Carriage of LCEVC 236
Annex A CRC decoder model 238
A.1 CRC decoder model 238
Annex B Digital storage medium command and control (DSM-CC) 239
B.1 Introduction 239
B.2 General elements 240
B.3 Technical elements 242
Annex C Program-specific information 248
C.1 Explanation of program-specific information in transport streams 248
C.2 Introduction 248
C.3 Functional mechanism 248
C.4 The mapping of sections into transport stream packets 249
C.5 Repetition rates and random access 249
C.6 What is a program? 250
C.7 Allocation of program_number 250
C.8 Usage of PSI in a typical system 250
C.9 The relationships of PSI structures 251
C.10 Bandwidth utilization and signal acquisition time 253
Annex D Systems timing model and application implications of this Recommendation | International Standard 256
D.1 Introduction 256
Annex E Data transmission applications 265
E.1 General considerations 265
E.2 Suggestion 265
Annex F Graphics of syntax for this Recommendation | International Standard 266

© ISO/IEC 2025 – All rights reserved

Page
F.1 Introduction ... 266
Annex G General information ... 270
G.1 General information ... 270
Annex H Private data ... 271
H.1 Private data ... 271
Annex I Systems conformance and real-time interface ... 272
I.1 Systems conformance and real-time interface ... 272
Annex J Interfacing jitter-inducing networks to MPEG-2 decoders ... 273
J.1 Introduction ... 273
J.2 Network compliance models ... 273
J.3 Network specification for jitter smoothing ... 274
J.4 Example decoder implementations ... 275
Annex K Splicing transport streams ... 276
K.1 Introduction ... 276
K.2 The different types of splicing point ... 276
K.3 Decoder behaviour on splices ... 277
Annex L Registration procedure (see 2.9) ... 279
L.1 Procedure for the request of a Registered Identifier (RID) ... 279
L.2 Responsibilities of the Registration Authority ... 279
L.3 Responsibilities of parties requesting an RID ... 279
L.4 Appeal procedure for denied applications ... 279
Annex M Registration application form (see 2.9) ... 281
M.1 Contact information of organization requesting a Registered Identifier (RID) ... 281
M.2 Statement of an intention to apply the assigned RID ... 281
M.3 Date of intended implementation of the RID ... 281
M.4 Authorized representative ... 281
M.5 For official use only of the Registration Authority ... 281
Annex N Registration Authority Diagram of administration structure (see 2.9) ... 282
Annex O Registration procedure (see 2.10) ... 283
O.1 Procedure for the request of an RID ... 283
O.2 Responsibilities of the Registration Authority ... 283
O.3 Contact information for the Registration Authority ... 283
O.4 Responsibilities of parties requesting an RID ... 283
O.5 Appeal procedure for denied applications ... 283
Annex P Registration application form ... 285
P.1 Contact information of organization requesting an RID ... 285
P.2 Request for a specific RID ... 285
P.3 Short description of RID that is in use and date system that was implemented ... 285
P.4 Statement of an intention to apply the assigned RID ... 285
P.5 Date of intended implementation of the RID ... 285
P.6 Authorized representative ... 285
P.7 For official use of the Registration Authority ... 285
Annex Q T-STD and P-STD buffer models for ISO/IEC 13818-7 ADTS ... 286
Q.1 Introduction ... 286
Q.2 Leak rate from transport buffer ... 286
Q.3 Buffer size ... 286
Q.4 Conclusion ... 287
Annex R Carriage of ISO/IEC 14496 scenes in Rec. ITU-T H.222.0 | ISO/IEC 13818-1 ... 289
R.1 Content access procedure for ISO/IEC 14496 program components within a program stream ... 289
R.2 Content access procedure for ISO/IEC 14496 program components within a transport stream ... 290
Annex S Carriage of JPEG 2000 part 1 video over MPEG-2 transport streams ... 294
S.1 Introduction ... 294

© ISO/IEC 2025 – All rights reserved

Page
S.2 J2K video access unit, J2K video elementary stream, J2K video sequence and J2K still picture 294
S.3 Optional J2K block mode for high resolution support 294
S.4 Optional J2K stripe mode for Ultra-Low Latency 295
S.5 Elementary stream header (elsm) and mapping to PES packets 295
S.6 J2K transport constraints 298
S.7 Interpretation of flags in adaptation and PES headers for J2K video elementary streams 299
S.8 T-STD extension for J2K video elementary streams 299
Annex T MIME type for MPEG-2 transport streams 302
T.1 Introduction 302
T.2 MIME type and subtype 302
T.3 Security considerations 303
T.4 Parameters 303
Annex U Carriage of timeline and external media information over MPEG-2 transport streams 305
U.1 Introduction 305
U.2 TEMI access unit and TEMI elementary stream 306
U.3 AF descriptors 307
Annex V Transport of HEVC tiles 316
V.1 Introduction 316
V.2 HEVC tile substream identification example 317
V.3 Subregion layout example 317
Annex W Carriage of JPEG XS part 1 video over MPEG-2 Transport Streams 319
W.1 Introduction 319
W.2 JPEG XS video access unit, JPEG XS video elementary stream, JPEG XS video sequence and JPEG XS still picture 319
W.3 Elementary stream header (jxes) and mapping to PES packets 319
W.4 JPEG XS transport constraints 320
W.5 Interpretation of flags in adaptation field and PES packet for JPEG XS video elementary streams 321
W.6 T-STD extension for JPEG XS video elementary streams 321
Bibliography 324

© ISO/IEC 2025 – All rights reserved

List of Tables

Page

Table 2-1 – Transport stream ... 26
Table 2-2 – Transport packet of this Recommendation | International Standard ... 26
Table 2-3 – PID table ... 27
Table 2-4 – Scrambling control values ... 27
Table 2-5 – Adaptation field control values ... 28
Table 2-6 – Transport stream adaptation field ... 28
Table 2-7 – Splice parameters Table 1 Simple Profile Main Level, Main Profile Main Level, SNR Profile Main Level (both layers), Spatial Profile High-1440 Level (base layer), High Profile Main Level (middle + base layers), Multi-view Profile Main Level (base layer) Video ... 35
Table 2-8 – Splice parameters Table 2 Main Profile Low Level, SNR Profile Low Level (both layers), High Profile Main Level (base layer), Multi-view Profile Low Level (base layer) Video ... 35
Table 2-9 – Splice parameters Table 3 Main Profile High-1440 Level, Spatial Profile High-1440 Level (all layers), High Profile High-1440 Level (middle + base layers), Multi-view Profile High-1440 Level (base layer) Video ... 36
Table 2-10 – Splice parameters Table 4 Main Profile High Level, High Profile High-1440 Level (all layers), High Profile High Level (middle + base layers), Multi-view Profile High Level (base layer) Video ... 36
Table 2-11 – Splice parameters Table 5 SNR Profile Low Level (base layer) Video ... 36
Table 2-12 – Splice parameters Table 6 SNR Profile Main Level (base layer) Video ... 36
Table 2-13 – Splice parameters Table 7 Spatial Profile High-1440 Level (middle + base layers) Video ... 37
Table 2-14 – Splice parameters Table 8 High Profile Main Level (all layers), High Profile High-1440 Level (base layer) Video ... 37
Table 2-15 – Splice parameters Table 9 High Profile High Level (base layer), Multi-view Profile Main Level (both layers) Video ... 37
Table 2-16 – Splice parameters Table 10 High Profile High Level (all layers), Multi-view Profile High-1440 Level (both layers) Video ... 37
Table 2-17 – Splice parameters Table 11 4:2:2 Profile Main Level Video ... 38
Table 2-18 – Splice parameters Table 12 Multi-view Profile Low Level (both layers) Video ... 38
Table 2-19 – Splice parameters Table 13 Multi-view Profile High Level (both layers) Video ... 38
Table 2-20 – Splice parameters Table 14 4:2:2 Profile High Level Video ... 38
Table 2-21 – PES packet ... 39
Table 2-22 – Stream_id assignments ... 42
Table 2-23 – PES scrambling control values ... 43
Table 2-24 – Trick mode control values ... 48
Table 2-25 – Field_id field control values ... 49
Table 2-26 – Coefficient selection values ... 49
Table 2-27 – Stream_id_extension assignments ... 51
Table 2-28 – Program-specific information ... 52
Table 2-29 – Program-specific information pointer ... 54
Table 2-30 – Program association section ... 54
Table 2-31 – table_id assignment values ... 55
Table 2-32 – Conditional access section ... 56
Table 2-33 – Transport stream program map section ... 57
Table 2-34 – Stream type assignments ... 58
Table 2-35 – Private section ... 61
Table 2-36 – The transport stream description table ... 62
Table 2-37 – Program stream ... 68
Table 2-38 – Program stream pack ... 68
Table 2-39 – Program stream pack header ... 68
Table 2-40 – Program stream system header ... 69
Table 2-41 – Program stream map ... 72
Table 2-42 – Program stream directory packet ... 74
Table 2-43 – Intra_coded indicator ... 76
Table 2-44 – Coding_parameters indicator ... 76
Table 2-45 – Program and program element descriptors ... 77
Table 2-46 – Video stream descriptor ... 78

© ISO/IEC 2025 – All rights reserved

Page

Table 2-47 – Frame rate code 78
Table 2-48 – Audio stream descriptor 79
Table 2-49 – Hierarchy descriptor 80
Table 2-50 – Hierarchy_type field values 81
Table 2-51 – Registration descriptor 81
Table 2-52 – Data stream alignment descriptor 82
Table 2-53 – Video stream alignment values 82
Table 2-54 – AVC video stream alignment values 83
Table 2-55 – HEVC video stream alignment values 83
Table 2-56 – Audio stream alignment values 83
Table 2-57 – VVC video stream alignment values 84
Table 2-58 – EVC video stream alignment values 84
Table 2-59 – Target background grid descriptor 85
Table 2-60 – Video window descriptor 85
Table 2-61 – Conditional access descriptor 86
Table 2-62 – ISO 639 language descriptor 86
Table 2-63 – Audio type values 87
Table 2-64 – System clock descriptor 88
Table 2-65 – Multiplex buffer utilization descriptor 88
Table 2-66 – Copyright descriptor 89
Table 2-67 – Maximum bitrate descriptor 89
Table 2-68 – Private data indicator descriptor 90
Table 2-69 – Smoothing buffer descriptor 90
Table 2-70 – STD descriptor 91
Table 2-71 – IBP descriptor 91
Table 2-72 – MPEG-4 video descriptor 92
Table 2-73 – MPEG-4 audio descriptor 92
Table 2-75 – IOD descriptor 95
Table 2-76 – SL descriptor 95
Table 2-77 – FMC descriptor 96
Table 2-78 – External_ES_ID descriptor 96
Table 2-79 – Muxcode descriptor 97
Table 2-80 – M4MuxBufferSize descriptor 97
Table 2-81 – MultiplexBuffer descriptor 98
Table 2-82 – M4MuxTiming descriptor 98
Table 2-83 – Content labelling descriptor 99
Table 2-84 – Metadata_application_format 99
Table 2-85 – Content_time_base_indicator values 100
Table 2-86 – Metadata pointer descriptor 101
Table 2-87 – Metadata format values 101
Table 2-88 – MPEG_carriage_flags 102
Table 2-89 – Metadata descriptor 103
Table 2-90 – decoder_config_flags 104
Table 2-91 – Metadata STD descriptor 105
Table 2-92 – AVC video descriptor 105
Table 2-93 – AVC timing and HRD descriptor 107
Table 2-94 – MPEG-2 AAC_audio_descriptor 108
Table 2-95 – MPEG-2_AAC_additional_information field values 109
Table 2-96 – MPEG-4 text descriptor 109
Table 2-97 – MPEG-4 audio extension descriptor 109
Table 2-98 – Auxiliary video stream descriptor 110
Table 2-99 – SVC extension descriptor 111
Table 2-100 – MVC extension descriptor 112
Table 2-101 – J2K video descriptor 113

© ISO/IEC 2025 – All rights reserved

Page

Table 2-102 – Example frame rates based on DEN_frame_rate and NUM_frame_rate values ... 115
Table 2-103 – MVC operation point descriptor ... 117
Table 2-104 – MPEG2_stereoscopic_video_format_descriptor syntax ... 118
Table 2-105 – Stereoscopic_program_info_descriptor syntax ... 118
Table 2-106 – Stereoscopic_service_type values ... 119
Table 2-107 – Stereoscopic_video_info_descriptor syntax ... 119
Table 2-108 – Upsampling factor values ... 120
Table 2-109 – Extension descriptor ... 120
Table 2-110 – Extension descriptor tag values ... 123
Table 2-111 – Transport_profile_descriptor syntax ... 124
Table 2-112 – Transport_profile values ... 124
Table 2-113 – HEVC video descriptor ... 125
Table 2-114 – Semantics of HDR_WCG_idc ... 127
Table 2-115 – HEVC timing and HRD descriptor ... 127
Table 2-116 – Adaptation field extension descriptor ... 128
Table 2-117 – HEVC operation point descriptor ... 129
Table 2-118 – HEVC hierarchy extension descriptor ... 131
Table 2-119 – Semantics of extension dimension bits ... 131
Table 2-120 – Green extension descriptor ... 132
Table 2-121 – MPEG-H 3D audio descriptor ... 133
Table 2-122 – MPEG-H 3D audio config descriptor ... 133
Table 2-123 – MPEG-H 3D audio scene descriptor ... 134
Table 2-124 – MPEG-H 3D audio text label descriptor ... 137
Table 2-125 – MPEG-H 3D audio multi-stream descriptor ... 139
Table 2-126 – MPEG-H 3D audio DRC and Loudness descriptor() ... 140
Table 2-127 – MPEG-H 3D audio command descriptor ... 142
Table 2-128 – Quality extension descriptor ... 143
Table 2-129 – Virtual segmentation descriptor ... 144
Table 2-130 – HEVC tile substream descriptor ... 145
Table 2-131 — HEVC subregion descriptor ... 146
Table 2-132 – JPEG XS video descriptor ... 148
Table 2-133 – VVC video descriptor ... 150
Table 2-134 – Semantics of HDR_WCG_idc ... 151
Table 2-135 – SDR widely used video property combinations ... 152
Table 2-136 – WCG widely used video property combinations ... 152
Table 2-137 – HDR/WCG widely used video property combinations ... 152
Table 2-138 – No Indication ... 153
Table 2-139 – VVC timing and HRD descriptor ... 153
Table 2-140 – EVC video descriptor ... 154
Table 2-141 – EVC timing and HRD descriptor ... 156
Table 2-142 – LCEVC video descriptor ... 157
Table 2-143 – LCEVC linkage descriptor ... 157
Table 2-144 – Media service kind descriptor ... 158
Table 2-145 – media_description_flag ... 159
Table 2-146 – Media type indicator ... 159
Table 2-147 – ID_length_code ... 159
Table 2-148 – ID_type ... 159
Table 2-149 – configuration type values ... 160
Table 2-150 – lang_len_indicator ... 160
Table 2-151 – Media service type values ... 160
Table 2-152 – Carriage of individual ISO/IEC 14496 streams in Rec. ITU-T H.222.0 | ISO/IEC 13818-1 ... 170
Table 2-153 – Section syntax for transport of ISO/IEC 14496 stream ... 175
Table 2-154 – ISO/IEC defined options for carriage of an ISO/IEC 14496 scene and associated streams in Rec. ITU-T H.222.0 | ISO/IEC 13818-1 ... 179
Table 2-155 – Metadata Access Unit Wrapper ... 185

© ISO/IEC 2025 – All rights reserved

Page

Table 2-156 – Metadata AU cell ... 185
Table 2-157 – Cell fragment indication ... 185
Table 2-158 – Section syntax for transport of metadata ... 186
Table 2-159 – Section fragment indication ... 187
Table 2-160 – View and dependency representation delimiter NAL unit ... 196
Table 2-161 – Implied hierarchy_layer_index if no hierarchy descriptors are used ... 216
Table 2-162 – Green access unit section syntax ... 223
Table 2-163 – Green access unit ... 223
Table 2-164 – Quality Access Unit ... 226
Table B.1 – DSM-CC syntax ... 243
Table B.2 – Command_id assigned values ... 243
Table B.3 – DSM-CC control ... 244
Table B.4 – Select mode assigned values ... 245
Table B.5 – DSM-CC Acknowledgement ... 246
Table B.6 – Time code ... 247
Table C.1 – Composite_descriptor ... 253
Table C.2 – Sub-descriptor ... 253
Table C.3 – Program association table bandwidth usage (bit/s) Number of programs per transport stream ... 254
Table C.4 – Program map table bandwidth usage (bit/s) Number of programs per transport stream ... 254
Table D.1 – Re-multiplexing strategy ... 261
Table E.1 – PES packet header example ... 265
Table S.1 – J2K Access unit elementary stream header ... 296
Table S.2 – Operating levels and maximum buffer size for JPEG 2000 broadcast profiles (from Table A.49 in Rec. ITU-T T.800 (2015) | ISO/IEC 15444-1:2016) ... 301
Table T.1 – 'codecs' parameter values for some specific stream_type values ... 303
Table U.1 – Variable field length notation example ... 305
Table U.1*bis* – Table U.1 in equivalent full notation ... 306
Table U.2 – TEMI access unit ... 306
Table U.3 – AF descriptor tags ... 307
Table U.4 – TEMI location descriptor ... 308
Table U.5 – TEMI URL scheme types ... 308
Table U.6 – TEMI service types ... 309
Table U.7 – TEMI base URL descriptor ... 309
Table U.8 – TEMI timeline descriptor ... 310
Table U.9 – TEMI MPEG-H_3dAudio_extStreamID descriptor ... 312
Table U.10 – Boundary descriptor ... 313
Table U.11 – sequence_number_length_code interpretation ... 313
Table U.12 – Labelling Descriptor ... 314
Table U.13 – label_length_code interpretation ... 314
Table U.14 – label_type values ... 314
Table U.15 – HEVC tile substream af_descriptor ... 315
Table W.1 – JPEG XS Access unit elementary stream header (jxes header) ... 319

© ISO/IEC 2025 – All rights reserved

List of Figures

Page
Figure Intro. 1 – Simplified overview of the scope of this Recommendation | International Standard xviii
Figure Intro. 2 – Prototypical transport demultiplexing and decoding example xx
Figure Intro. 3 – Prototypical transport multiplexing example xx
Figure Intro. 4 – Prototypical transport stream to program stream conversion xx
Figure Intro. 5 – Prototypical decoder for program streams xxi
Figure 2-1 – Transport stream system target decoder notation 16
Figure 2-2 – Program stream system target decoder notation 63
Figure 2-3 – Target background grid descriptor display area 84
Figure 2-4 – Media Service Kind Descriptor semantics at program level 163
Figure 2-5 – Media Program Kind Descriptor semantics at elementary stream level 164
Figure 2-6 – T-STD model extensions for individual ISO/IEC 14496 elementary streams 171
Figure 2-7 – T-STD model for ISO/IEC 14496 content 176
Figure 2-8 – P-STD model for ISO/IEC 14496 Systems stream 179
Figure 2-9 – Timing model for delivery of content and metadata 182
Figure 2-10 – Delivery of metadata in PES packets 183
Figure 2-11 – Metadata signalling and referencing 189
Figure 2-12 – Metadata decoding in the STD 189
Figure 2-13 – T-STD model extensions for Rec. ITU-T H.264 | ISO/IEC 14496-10 video 193
Figure 2-14 – P-STD model extensions for Rec. ITU-T H.264 | ISO/IEC 14496-10 video 195
Figure 2-15 – T-STD model extensions for Rec. ITU-T H.264 | ISO/IEC 14496-10 video with scalable video sub-bitstreams 196
Figure 2-16 – P-STD model extensions for Rec. ITU-T H.264 | ISO/IEC 14496-10 Video with scalable video sub-bitstreams 199
Figure 2-17 – T-STD model extensions for Rec. ITU-T H.264 | ISO/IEC 14496-10 Video with MVC video sub-bitstreams 201
Figure 2-18 – P-STD model extensions for Rec. ITU-T H.264 | ISO/IEC 14496-10 Video with MVC video sub-bitstreams 205
Figure 2-19 – T-STD model extensions for ISO/IEC 14496-17 text streams 207
Figure 2-20 – T-STD model extensions for single layer HEVC 210
Figure 2-21 – T-STD model extensions for layered transport of HEVC temporal video subsets 212
Figure 2-22 – T-STD model extensions for bitstream-partition-specific CPB operation 214
Figure 2-23 – T-STD model extensions for transport of HEVC tiles through individual ESs 218
Figure 2-24 – T-STD model extensions for transport of HEVC tiles in a common ES using AF descriptors 220
Figure 2-25 – T-STD model extension for transport of HEVC tiles in a common ES ignoring AF descriptors 222
Figure 2-26 – T-STD model extension for green access units 224
Figure 2-27 – Transport stream system target decoder for multiple audio elementary streams 226
Figure 2-28 – Quality Access Unit decoder processing model 227
Figure 2-29 – T-STD model extensions for single layer VVC 229
Figure 2-30 – T-STD model extensions for layered transport of VVC temporal video subsets 231
Figure 2-31 – T-STD model extensions for single layer EVC 234
Figure A.1 – 32-bit CRC decoder model 238
Figure B.1 – Configuration of DSM-CC application 241
Figure B.2 – BSM-CC bitstream decoded as a stand-alone bitstream 241
Figure B.3 – DSM-CC bitstream decoded as part of the system bitstream 242
Figure C.1 – Program and network mapping relationships 251
Figure D.1 – Constant delay model 256
Figure D.2 – STC recovery using PLL 260
Figure F.1 – Transport stream syntax diagram 266
Figure F.2 – PES packet syntax diagram 267
Figure F.3 – Program association section diagram 267
Figure F.4 – Conditional access section diagram 267
Figure F.5 – TS program map section diagram 268
Figure F.6 – Private section diagram 268
Figure F.7 – Program stream diagram 269

© ISO/IEC 2025 – All rights reserved

Page

Figure F.8 – Program stream map diagram 269
Figure J.1 – Sending system streams over a jitter-inducing network 274
Figure J.2 – Jitter-smoothing using network-layer timestamps 274
Figure J.3 – Integrated dejittering and MPEG-2 decoding 275
Figure R.1 – Example of ISO/IEC 14496 content in a program stream 290
Figure R.2 – Example of ISO/IEC 14496 content in a transport stream 291
Figure R.3 – Usage of MPEG-4 in a transport stream with BIFS scene referring to native PES 292
Figure R.4 – Usage of MPEG-4 in a transport stream with an ODUpdate_descriptor carrying an image ObjectDescriptor in the PMT 293
Figure S.1 – Structure and order of JPEG 2000 access units 298
Figure S.2 – T-STD model extensions for J2K Video 299
Figure U.1 – Stream partitioning into 2 and 5 second segments 313
Figure V.1 – Illustration of HEVC tiled encoding of panoramic content beyond UHD 316
Figure V.2 – Example of HEVC tile substream identification 317
Figure V.3 – Example of subregion layout for a 3 x 3 RoI 317
Figure W.1 – Structure and order of JPEG XS access units 320
Figure W.2 – T-STD model extensions for JPEG XS Video 322

© ISO/IEC 2025 – All rights reserved

Introduction

The systems part of this Recommendation | International Standard addresses the combining of one or more elementary streams of video and audio, as well as other data, into single or multiple streams which are suitable for storage or transmission. Systems coding follows the syntactical and semantic rules imposed by this Specification and provides information to enable synchronized decoding of decoder buffers over a wide range of retrieval or receipt conditions.

System coding shall be specified in two forms: the transport stream and the program stream. Each is optimized for a different set of applications. Both the transport stream and program stream defined in this Recommendation | International Standard provide coding syntax which is necessary and sufficient to synchronize the decoding and presentation of the video and audio information, while ensuring that data buffers in the decoders do not overflow or underflow. Information is coded in the syntax using time stamps concerning the decoding and presentation of coded audio and visual data and time stamps concerning the delivery of the data stream itself. Both stream definitions are packet-oriented multiplexes.

The basic multiplexing approach for single video and audio elementary streams is illustrated in Figure Intro. 1. The video and audio data is encoded as described in Rec. ITU-T H.262 | ISO/IEC 13818-2 and ISO/IEC 13818-3. The resulting compressed elementary streams are packetized to produce PES packets. Information needed to use PES packets independently of either transport streams or program streams may be added when PES packets are formed. This information is not needed and need not be added when PES packets are further combined with system level information to form transport streams or program streams. This systems standard covers those processes to the right of the vertical dashed line.

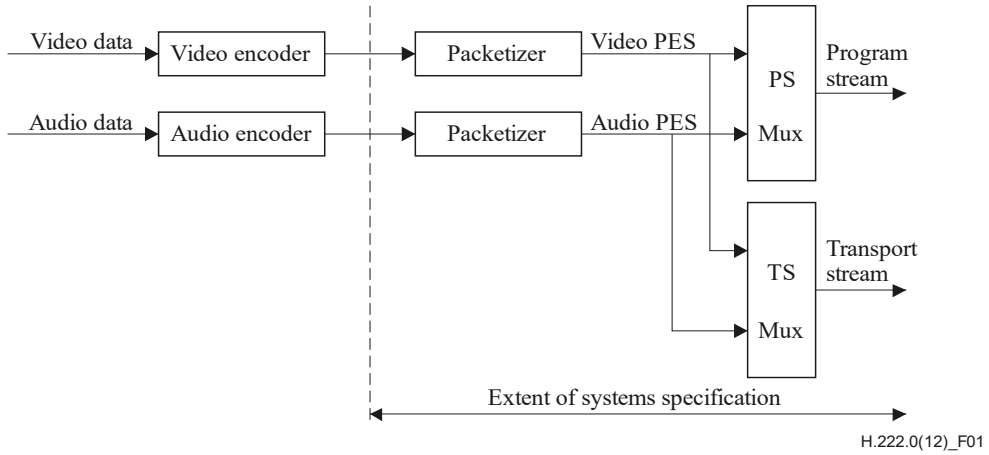


Figure Intro. 1 – Simplified overview of the scope of this Recommendation | International Standard

The program stream is analogous and similar to the ISO/IEC 11172 systems layer. It results from combining one or more streams of PES packets, which have a common time base, into a single stream.

For applications that require the elementary streams that comprise a single program to be in separate streams that are not multiplexed, the elementary streams can also be encoded as separate program streams, one per elementary stream, with a common time base. In this case the values encoded in the SCR fields of the various streams shall be consistent.

Like the single program stream, all elementary streams can be decoded with synchronization.

The program stream is designed for use in relatively error-free environments and is suitable for applications which may involve software processing of system information such as interactive multi-media applications. Program stream packets may be of variable and relatively great length.

The transport stream combines one or more programs with one or more independent time bases into a single stream. PES packets made up of elementary streams that form a program share a common time base. The transport stream is designed for use in environments where errors are likely, such as storage or transmission in lossy or noisy media. Transport stream packets are 188 bytes in length.

Program and transport streams are designed for different applications and their definitions do not strictly follow a layered model. It is possible and reasonable to convert from one to the other; however, one is not a subset or superset of the other. In particular, extracting the contents of a program from a transport stream and creating a valid program stream is possible and is accomplished through the common interchange format of PES packets, but not all of the fields needed in a program stream are contained within the transport stream; some must be derived. The transport stream may be used to span a range of layers in a layered model, and is designed for efficiency and ease of implementation in high bandwidth applications.

© ISO/IEC 2025 – All rights reserved

The scope of syntactical and semantic rules set forth in the systems specification differs: the syntactical rules apply to systems layer coding only, and do not extend to the compression layer coding of the video and audio specifications; by contrast, the semantic rules apply to the combined stream in its entirety.

The systems specification does not specify the architecture or implementation of encoders or decoders, nor those of multiplexors or demultiplexors. However, bit stream properties do impose functional and performance requirements on encoders, decoders, multiplexors and demultiplexors. For instance, encoders must meet minimum clock tolerance requirements. Notwithstanding this and other requirements, a considerable degree of freedom exists in the design and implementation of encoders, decoders, multiplexors, and demultiplexors.

Intro. 1 Transport stream

The transport stream is a stream definition which is tailored for communicating or storing one or more programs of coded data according to Rec. ITU-T H.262 | ISO/IEC 13818-2 and ISO/IEC 13818-3 and other data in environments in which significant errors may occur. Such errors may be manifested as bit value errors or loss of packets.

Transport streams may be either fixed or variable rate. In either case the constituent elementary streams may either be fixed or variable rate. The syntax and semantic constraints on the stream are identical in each of these cases. The transport stream rate is defined by the values and locations of program clock reference (PCR) fields, which in general are separate PCR fields for each program.

There are some difficulties with constructing and delivering a transport stream containing multiple programs with independent time bases such that the overall bit rate is variable. Refer to 2.4.2.3.

The transport stream may be constructed by any method that results in a valid stream. It is possible to construct transport streams containing one or more programs from elementary coded data streams, from program streams, or from other transport streams which may themselves contain one or more programs.

The transport stream is designed in such a way that several operations on a transport stream are possible with minimum effort. Among these are:

1) Retrieve the coded data from one program within the transport stream, decode it and present the decoded results as shown in Figure Intro. 2.
2) Extract the transport stream packets from one program within the transport stream and produce as output a different transport stream with only that one program as shown in Figure Intro. 3.
3) Extract the transport stream packets of one or more programs from one or more transport streams and produce as output a different transport stream (not illustrated).
4) Extract the contents of one program from the transport stream and produce as output a program stream containing that one program as shown in Figure Intro. 4.
5) Take a program stream, convert it into a transport stream to carry it over a lossy environment, and then recover a valid, and in certain cases, identical program stream.

Figure Intro. 2 and Figure Intro. 3 illustrate prototypical demultiplexing and decoding systems which take as input a transport stream. Figure Intro. 2 illustrates the first case, where a transport stream is directly demultiplexed and decoded. Transport streams are constructed in two layers:

– a system layer; and
– a compression layer.

The input stream to the transport stream decoder has a system layer wrapped about a compression layer. Input streams to the video and audio decoders have only the compression layer.

Operations performed by the prototypical decoder which accepts transport streams either apply to the entire transport stream ("multiplex-wide operations"), or to individual elementary streams ("stream-specific operations"). The transport stream system layer is divided into two sub-layers, one for multiplex-wide operations (the transport stream packet layer), and one for stream-specific operations (the PES packet layer).

A prototypical decoder for transport streams, including audio and video, is also depicted in Figure Intro. 2 to illustrate the function of a decoder. The architecture is not unique – some system decoder functions, such as decoder timing control, might equally well be distributed among elementary stream decoders and the channel-specific decoder – but this figure is useful for discussion. Likewise, indication of errors detected by the channel-specific decoder to the individual audio and video decoders may be performed in various ways and such communication paths are not shown in the diagram. The prototypical decoder design does not imply any normative requirement for the design of a transport stream decoder. Indeed non-audio/video data is also allowed, but not shown.

© ISO/IEC 2025 – All rights reserved

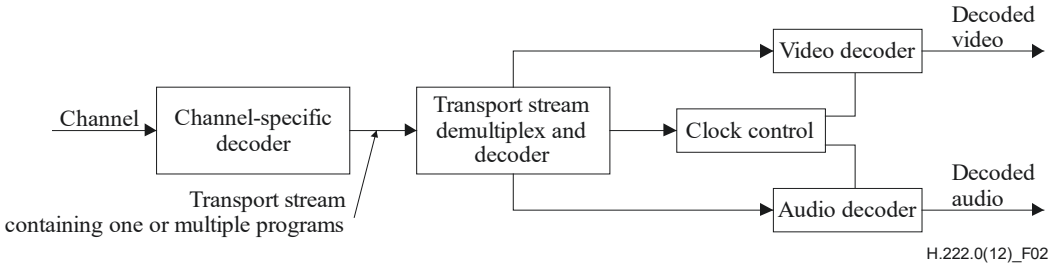


Figure Intro. 2 – Prototypical transport demultiplexing and decoding example

Figure Intro. 3 illustrates the second case, where a transport stream containing multiple programs is converted into a transport stream containing a single program. In this case the re-multiplexing operation may necessitate the correction of program clock reference (PCR) values to account for changes in the PCR locations in the bit stream.

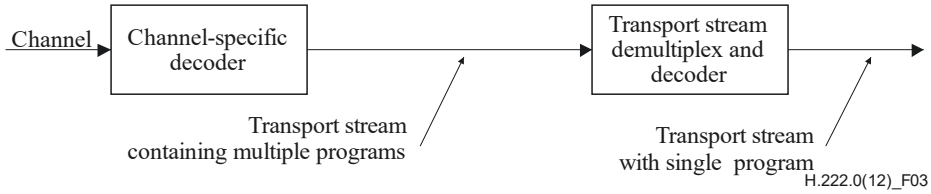


Figure Intro. 3 – Prototypical transport multiplexing example

Figure Intro. 4 illustrates a case in which a multi-program transport stream is first demultiplexed and then converted into a program stream.

Figures Intro. 3 and Intro. 4 indicate that it is possible and reasonable to convert between different types and configurations of transport streams. There are specific fields defined in the transport stream and program stream syntax which facilitate the conversions illustrated. There is no requirement that specific implementations of demultiplexors or decoders include all of these functions.

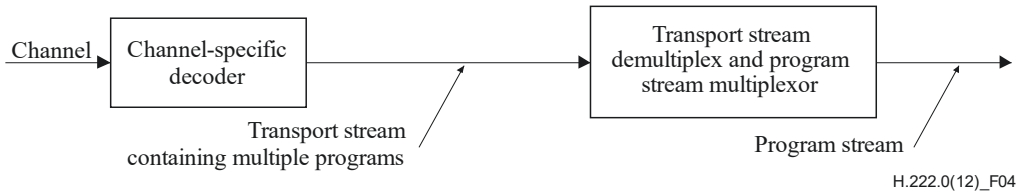


Figure Intro. 4 – Prototypical transport stream to program stream conversion

Intro. 2 Program stream

The program stream is a stream definition which is tailored for communicating or storing one program of coded data and other data in environments where errors are very unlikely, and where processing of system coding, e.g., by software, is a major consideration.

Program streams may be either fixed or variable rate. In either case, the constituent elementary streams may be either fixed or variable rate. The syntax and semantics constraints on the stream are identical in each case. The program stream rate is defined by the values and locations of the system clock reference (SCR) and mux_rate fields.

A prototypical audio/video program stream decoder system is depicted in Figure Intro. 5. The architecture is not unique – system decoder functions including decoder timing control might as equally well be distributed among elementary stream decoders and the channel-specific decoder – but this figure is useful for discussion. The prototypical decoder design does not imply any normative requirement for the design of a program stream decoder. Indeed, non-audio/video data is also allowed, but not shown.

© ISO/IEC 2025 – All rights reserved

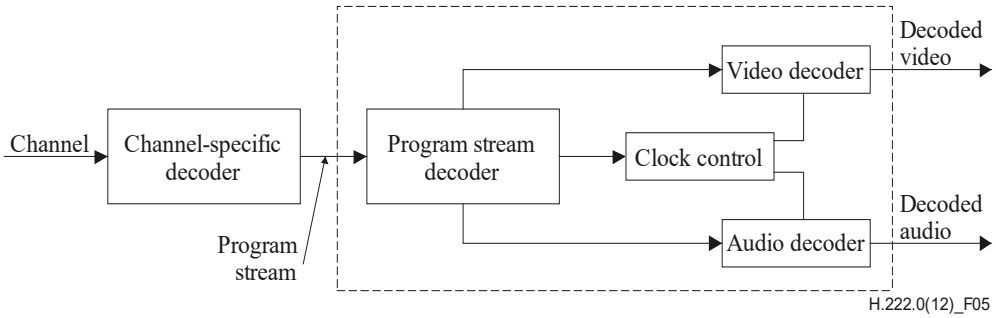


Figure Intro. 5 – Prototypical decoder for program streams

The prototypical decoder for program streams shown in Figure Intro. 5 is composed of system, video and audio decoders conforming to Parts 1, 2 and 3, respectively, of ISO/IEC 13818. In this decoder, the multiplexed coded representation of one or more audio and/or video streams is assumed to be stored or communicated on some channel in some channel-specific format. The channel-specific format is not governed by this Recommendation | International Standard, nor is the channel-specific decoding part of the prototypical decoder.

The prototypical decoder accepts as input a program stream and relies on a program stream decoder to extract timing information from the stream. The program stream decoder demultiplexes the stream, and the elementary streams so produced serve as inputs to video and audio decoders, whose outputs are decoded video and audio signals. Included in the design, but not shown in the figure, is the flow of timing information among the program stream decoder, the video and audio decoders, and the channel-specific decoder. The video and audio decoders are synchronized with each other and with the channel using this timing information.

Program streams are constructed in two layers: a system layer and a compression layer. The input stream to the program stream decoder has a system layer wrapped about a compression layer. Input streams to the video and audio decoders have only the compression layer.

Operations performed by the prototypical decoder either apply to the entire program stream ("multiplex-wide operations"), or to individual elementary streams ("stream-specific operations"). The program stream system layer is divided into two sub-layers, one for multiplex-wide operations (the pack layer), and one for stream-specific operations (the PES packet layer).

Intro. 3 Conversion between transport stream and program stream

It may be possible and reasonable to convert between transport streams and program streams by means of PES packets. This results from the specification of transport stream and program stream as embodied in 2.4.1 and 2.5.1 of the normative requirements of this Recommendation | International Standard. PES packets may, with some constraints, be mapped directly from the payload of one multiplexed bit stream into the payload of another multiplexed bit stream. It is possible to identify the correct order of PES packets in a program to assist with this if the program_packet_sequence_counter is present in all PES packets.

Certain other information necessary for conversion, e.g., the relationship between elementary streams, is available in tables and headers in both streams. Such data, if available, shall be correct in any stream before and after conversion.

Intro. 4 Packetized elementary stream

Transport streams and program streams are each logically constructed from PES packets, as indicated in the syntax definitions in 2.4.3.6. PES packets shall be used to convert between transport streams and program streams; in some cases the PES packets need not be modified when performing such conversions. PES packets may be much larger than the size of a transport stream packet.

A continuous sequence of PES packets of one elementary stream with one stream ID may be used to construct a PES Stream. When PES packets are used to form a PES stream, they shall include elementary stream clock reference (ESCR) fields and elementary stream rate (ES_Rate) fields, with constraints as defined in 2.4.3.8. The PES stream data shall be contiguous bytes from the elementary stream in their original order. PES streams do not contain some necessary system information which is contained in program streams and transport streams. Examples include the information in the pack header, system header, program stream map, program stream directory, program map table, and elements of the transport stream packet syntax.

© ISO/IEC 2025 – All rights reserved

The PES stream is a logical construct that may be useful within implementations of this Recommendation | International Standard; however, it is not defined as a stream for interchange and interoperability. Applications requiring streams containing only one elementary stream can use program streams or transport streams which each contain only one elementary stream. These streams contain all of the necessary system information. Multiple program streams or transport streams, each containing a single elementary stream, can be constructed with a common time base and therefore carry a complete program, i.e., with audio and video.

Intro. 5 Timing model

Systems, video and audio all have a timing model in which the end-to-end delay from the signal input to an encoder to the signal output from a decoder is a constant. This delay is the sum of encoding, encoder buffering, multiplexing, communication or storage, demultiplexing, decoder buffering, decoding, and presentation delays. As part of this timing model all video pictures and audio samples are presented exactly once, unless specifically coded to the contrary, and the inter-picture interval and audio sample rate are the same at the decoder as at the encoder. The system stream coding contains timing information which can be used to implement systems which embody constant end-to-end delay. It is possible to implement decoders which do not follow this model exactly; however, in such cases it is the decoder's responsibility to perform in an acceptable manner. The timing is embodied in the normative specifications of this Recommendation | International Standard, which must be adhered to by all valid bit streams, regardless of the means of creating them.

All timing is defined in terms of a common system clock, referred to as a system time clock (STC). In the program stream this clock may have an exactly specified ratio to the video or audio sample clocks, or it may have an operating frequency which differs slightly from the exact ratio while still providing precise end-to-end timing and clock recovery.

In the transport stream the system clock frequency is constrained to have the exactly specified ratio to the audio and video sample clocks at all times; the effect of this constraint is to simplify sample rate recovery in decoders.

Intro. 6 Conditional access

Encryption and scrambling for conditional access to programs encoded in the program and transport streams is supported by the system data stream definitions. Conditional access mechanisms are not specified here. The stream definitions are designed so that implementation of practical conditional access systems is reasonable, and there are some syntactical elements specified which provide specific support for such systems.

Intro. 7 Multiplex-wide operations

Multiplex-wide operations include the coordination of data retrieval of the channel, the adjustment of clocks, and the management of buffers. The tasks are intimately related. If the rate of data delivery of the channel is controllable, then data delivery may be adjusted so that decoder buffers neither overflow nor underflow; but if the data rate is not controllable, then elementary stream decoders must slave their timing to the data received from the channel to avoid overflow or underflow.

Program streams are composed of packs whose headers facilitate the above tasks. Pack headers specify intended times at which each byte is to enter the program stream Decoder from the channel, and this target arrival schedule serves as a reference for clock correction and buffer management. The schedule need not be followed exactly by decoders, but they must compensate for deviations about it.

Similarly, transport streams are composed of transport stream packets with headers containing information which specifies the times at which each byte is intended to enter a transport stream decoder from the channel. This schedule provides exactly the same function as that which is specified in the program stream.

An additional multiplex-wide operation is a decoder's ability to establish what resources are required to decode a transport stream or program stream. The first pack of each program stream conveys parameters to assist decoders in this task. Included, for example, are the stream's maximum data rate and the highest number of simultaneous video channels. The transport stream likewise contains globally useful information.

The transport stream and program stream each contain information which identifies the pertinent characteristics of, and relationships between, the elementary streams which constitute each program. Such information may include the language spoken in audio channels, as well as the relationship between video streams when multi-layer video coding is implemented.

Intro. 8 Individual stream operations (PES packet layer)

The principal stream-specific operations are:

1) demultiplexing; and
2) synchronizing playback of multiple elementary streams.

© ISO/IEC 2025 – All rights reserved

Intro. 8.1 Demultiplexing

On encoding, program streams are formed by multiplexing elementary streams, and transport streams are formed by multiplexing elementary streams, program streams, or the contents of other transport streams. Elementary streams may include private, reserved, and padding streams in addition to audio and video streams. The streams are temporally subdivided into packets, and the packets are serialized. A PES packet contains coded bytes from one and only one elementary stream.

In the program stream both fixed and variable packet lengths are allowed subject to constraints as specified in 2.5.1 and 2.5.2. For transport streams the packet length is 188 bytes. Both fixed and variable PES packet lengths are allowed, and will be relatively long in most applications.

On decoding, demultiplexing is required to reconstitute elementary streams from the multiplexed program stream or transport stream. Stream_id codes in program stream packet headers, and packet ID codes in the transport stream make this possible.

Intro. 8.2 Synchronization

Synchronization among multiple elementary streams is accomplished with presentation time stamps (PTSs) in the program stream and transport streams. Time stamps are generally in units of 90 kHz, but the system clock reference (SCR), the program clock reference (PCR) and the optional elementary stream clock reference (ESCR) have extensions with a resolution of 27 MHz. Decoding of N-elementary streams is synchronized by adjusting the decoding of streams to a common master time base rather than by adjusting the decoding of one stream to match that of another. The master time base may be one of the N-decoders' clocks, the data source's clock, or it may be some external clock.

Each program in a transport stream, which may contain multiple programs, may have its own time base. The time bases of different programs within a transport stream may be different.

Because PTSs apply to the decoding of individual elementary streams, they reside in the PES packet layer of both the transport streams and program streams. End-to-end synchronization occurs when encoders save time stamps at capture time, when the time stamps propagate with associated coded data to decoders, and when decoders use those time stamps to schedule presentations.

Synchronization of a decoding system with a channel is achieved through the use of the SCR in the program stream and by its analogue, the PCR, in the transport stream. The SCR and PCR are time stamps encoding the timing of the bit stream itself, and are derived from the same time base used for the audio and video PTS values from the same program. Since each program may have its own time base, there are separate PCR fields for each program in a transport stream containing multiple programs. In some cases it may be possible for programs to share PCR fields. Refer to 2.4.4, program-specific information (PSI), for the method of identifying which PCR is associated with a program. A program shall have one and only one PCR time base associated with it.

Intro. 8.3 Relation to compression layer

The PES packet layer is independent of the compression layer in some senses, but not in all. It is independent in the sense that PES packet payloads need not start at compression layer start codes, as defined in Parts 2 and 3 of ISO/IEC 13818. For example, video start codes may occur anywhere within the payload of a PES packet, and start codes may be split by a PES packet header. However, time stamps encoded in PES packet headers apply to presentation times of compression layer constructs (namely, presentation units). In addition, when the elementary stream data conforms to Rec. ITU-T H.262 | ISO/IEC 13818-2 or ISO/IEC 13818-3, the PES_packet_data_bytes shall be byte aligned to the bytes of this Recommendation | International Standard.

Intro. 9 System reference decoder

Part 1 of ISO/IEC 13818 employs a "system target decoder" (STD), one for transport streams (refer to 2.4.2) referred to as "transport system target decoder" (T-STD) and one for program streams (refer to 2.5.2) referred to as "program system target decoder" (P-STD), to provide a formalism for timing and buffering relationships. Because the STD is parameterized in terms of Rec. ITU-T H.222.0 | ISO/IEC 13818-1 fields (for example, buffer sizes) each elementary stream leads to its own parameterization of the STD. Encoders shall produce bit streams that meet the appropriate STD's constraints. Physical decoders may assume that a stream plays properly on its STD. The physical decoder must compensate for ways in which its design differs from that of the STD.

Intro. 10 Applications

The streams defined in this Recommendation | International Standard are intended to be as useful as possible to a wide variety of applications. Application developers should select the most appropriate stream.

© ISO/IEC 2025 – All rights reserved

Modern data communications networks may be capable of supporting Rec. ITU-T H.222.0 | ISO/IEC 13818-1 video and ISO/IEC 13818 audio. A real-time transport protocol is required. The program stream may be suitable for transmission on such networks.

The program stream is also suitable for multimedia applications on CD-ROM. Software processing of the program stream may be appropriate.

The transport stream may be more suitable for error-prone environments, such as those used for distributing compressed bit-streams over long-distance networks and in broadcast systems.

Many applications require storage and retrieval of Rec. ITU-T H.222.0 | ISO/IEC 13818-1 bitstreams on various digital storage media (DSM). A digital storage media command and control (DSM-CC) protocol is specified in Annex B and Part 6 of ISO/IEC 13818 in order to facilitate the control of such media.

© ISO/IEC 2025 – All rights reserved

INTERNATIONAL STANDARD
ITU-T RECOMMENDATION

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

SECTION 1 – GENERAL

1.1 Scope

This Recommendation | International Standard specifies the system layer of the coding. It was developed principally to support the combination of the video and audio coding methods defined in Parts 2 and 3 of ISO/IEC 13818. The system layer supports six basic functions:

1) the synchronization of multiple compressed streams on decoding;
2) the interleaving of multiple compressed streams into a single stream;
3) the initialization of buffering for decoding start up;
4) continuous buffer management;
5) time identification;
6) multiplexing and signalling of various components in a system stream.

A Rec. ITU-T H.222.0 | ISO/IEC 13818-1 multiplexed bit stream is either a transport stream or a program stream. Both streams are constructed from PES packets and packets containing other necessary information. Both stream types support multiplexing of video and audio compressed streams from one program with a common time base. The transport stream additionally supports the multiplexing of video and audio compressed streams from multiple programs with independent time bases. For almost error-free environments the program stream is generally more appropriate, supporting software processing of program information. The transport stream is more suitable for use in environments where errors are likely.

A Rec. ITU-T H.222.0 | ISO/IEC 13818-1 multiplexed bit stream, whether a transport stream or a program stream, is constructed in two layers: the outermost layer is the system layer, and the innermost is the compression layer. The system layer provides the functions necessary for using one or more compressed data streams in a system. The video and audio parts of this Specification define the compression coding layer for audio and video data. Coding of other types of data is not defined by this Specification, but is supported by the system layer provided that the other types of data adhere to the constraints defined in 2.7.

1.2 Normative references

The following Recommendations and International Standards contain provisions which, through reference in this text, constitute provisions of this Recommendation | International Standard. At the time of publication, the editions indicated were valid. All Recommendations and Standards are subject to revision, and parties to agreements based on this Recommendation | International Standard are encouraged to investigate the possibility of applying the most recent edition of the Recommendations and Standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards. The Telecommunication Standardization Bureau of the ITU maintains a list of currently valid ITU-T Recommendations.

1.2.1 Identical Recommendations | International Standards

- Recommendation ITU-T H.262 (2012) | ISO/IEC 13818-2:2013, *Information technology – Generic coding of moving pictures and associated audio information: Video*.
- Recommendation ITU-T T.800 (2024) | ISO/IEC 15444-1:2024, *Information technology – JPEG 2000 image coding system: Core coding system.*

1.2.2 Paired Recommendations | International Standards equivalent in technical content

- Recommendation ITU-T H.264 (2024), *Advanced video coding for generic audiovisual services*.

  ISO/IEC 14496-10:2025, *Information technology – Coding of audio-visual objects – Part 10: Advanced video coding*.
- Recommendation ITU-T H.265 (2024), *High efficiency video coding.*

  ISO/IEC 23008-2:2025, *Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High efficiency video coding*.

© ISO/IEC 2025 – All rights reserved

– Recommendation. ITU-T H.273 (2024), *Coding-independent code points for video signal type identification*.

ISO/IEC 23091-2:2025, *Information technology — Coding-independent code points — Part 2: Video*.

– Recommendation ITU-T T.171 (1996), *Protocols for interactive audiovisual services: coded representation of multimedia and hypermedia objects*.

ISO/IEC 13522-1:1997, *Information technology – Coding of Multimedia and Hypermedia information – Part 1: MHEG object representation – Base notation (ASN.1)*.

– Recommendation ITU-T H.266 (2023), *Versatile video coding*.

ISO/IEC 23090-3:2024, *Information technology – Coded Representation of Immersive Media – Part 3: Versatile video coding*.

– Recommendation ITU-T H.274 (2023), *Versatile supplemental enhancement information messages for coded video bitstreams*.

ISO/IEC 23002-7:2024 – *Information Technology – MPEG Video technologies – Part 7: Versatile supplemental enhancement information messages for coded video bitstreams*.

1.2.3 Additional references

– Recommendation ITU-R BT.709-6 (2015), *Parameter values for the HDTV standards for production and international programme exchange*.

– Recommendation ITU-R BT.1886 (2011), *Reference electro-optical transfer function for flat panel displays used in HDTV studio production.*

– Recommendation ITU-R BT.2020 (2015), *Parameter values for ultra-high definition television systems for production and international programme exchange.*

– Recommendation ITU-R BT.2100-3 (2025), *Image parameter values for high dynamic range television for use in production and international programme exchange.*

– ISO 639:2023, *Code for individual languages and individual groups*.

– ISO 8859-1:1998, *Information technology – 8-bit single-byte coded graphic character sets – Part 1: Latin alphabet No. 1*.

– ISO 15706-1:2023, *Information and documentation – International Standard Audiovisual Number (ISAN) – Part 1: Audiovisual work identifier*.

– ISO 15706-2:2023, *Information and documentation – International Standard Audiovisual Number (ISAN) – Part 2: Version identifier*.

– ISO/IEC 11172-1:1993, *Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s – Part 1: Systems*.

– ISO/IEC 11172-2:1993, *Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s – Part 2: Video*.

– ISO/IEC 11172-3:1993, *Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s – Part 3: Audio*.

– ISO/IEC 13818-3:1998, *Information technology – Generic coding of moving pictures and associated audio information – Part 3: Audio*.

– ISO/IEC 13818-6:1998, *Information technology – Generic coding of moving pictures and associated audio information – Part 6: Extensions for DSM-CC*.

– ISO/IEC 13818-7:2006, *Information technology – Generic coding of moving pictures and associated audio information – Part 7: Advanced Audio Coding (AAC)*.

– ISO/IEC 13818-11:2004, *Information technology – Generic coding of moving pictures and associated audio information – Part 11: IPMP on MPEG-2 systems*.

– ISO/IEC 14496-1:2010, *Information technology – Coding of audio-visual objects – Part 1: Systems*.

– ISO/IEC 14496-2:2004, *Information technology – Coding of audio-visual objects – Part 2: Visual*.

– ISO/IEC 14496-3:2019, *Information technology – Coding of audio-visual objects – Part 3: Audio*.

– ISO/IEC 14496-17:2006, *Information technology, Coding of audio-visual objects – Part 17: Streaming text format*.

– ISO/IEC 21122-1:2024, *JPEG XS low-latency lightweight image coding system – Part 1: Core coding system*.

© ISO/IEC 2025 – All rights reserved

- ISO/IEC 21122-2:2024, *JPEG XS low-latency lightweight image coding system – Part 2: Profiles and buffer models*.
- ISO/IEC 21122-3:2024, *JPEG XS low-latency lightweight image coding system – Part 3: Transport and container formats*.
- ISO/IEC 23001-10:2020, *Information technology – MPEG systems technologies – Part 10: Carriage of timed metadata metrics of media in ISO base media file format.*
- ISO/IEC 23001-11:2019, *Information technology – MPEG systems technologies – Part 11: Energy-efficient media consumption (Green Metadata)*.
- ISO/IEC 23001-12:2018, *Information technology – MPEG systems technologies – Part 12: Sample Variants*.
- ISO/IEC 23001-13:2019, *Information technology – MPEG systems technologies – Part 13: Media Orchestration*.
- ISO/IEC 23003-4:2020, *Information technology – MPEG audio technologies – Part 4: Dynamic Range Control*.
- ISO/IEC 23008-3:2019, *Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 3: 3D audio.*
- ISO/IEC 23091-3:2018, *Information technology – Coding-independent code points – Part 3: Audio*.
- ISO/IEC 23094-1:2020*, Information technology – General video coding – Part 1: Essential video coding*.
- ISO/IEC 23094-2:2021, *Information technology – General video coding – Part 2: Low complexity enhancement video coding.*
- ANSI/SCTE 35 (2019), *Digital Program Insertion Cueing Message for Cable*.
- IETF BCP 47: IETF RFC 4647 (2006) *Matching of Language Tags* combined with IETF RFC 5646 (2009), *Tags for Identifying Languages*.
- IETF RFC 3986 (2005), *Uniform Resource Identifier (URI): Generic Syntax*.
- IETF RFC 5484 (2009), *Associating Time-Codes with RTP Streams*.

© ISO/IEC 2025 – All rights reserved