INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND ASSOCIATED AUDIO

ISO/IEC JTC1/SC29/WG11 N0742

21 July 1994

Compliance

INFORMATION TECHNOLOGY -

GENERIC CODING OF MOVING PICTURES AND ASSOCIATED AUDIO

ISO/IEC 13818-4

Working Draft

TTmm ddd d Jul ŒŒŒŒ

I. CONTENTS

FOREWORD

The ITU-T (the ITU Telecommunication Standardisation Sector) is a permanent organ of the International Telecommunications Union (ITU). The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to developing telecommunication standards on a world-wide basis.

The World Telecommunication Standardisation Conference, which meets every four years, establishes the program of work arising from the review of existing questions and new questions among other things. The approval of new or revised Recommendations by members of the ITU-T is covered by the procedure laid down in the ITU-T Resolution No. 1 (Helsinki 1993). The proposal for Recommendation is accepted if 70% or more of the replies from members indicate approval.

ISO (the International Organisation for Standardisation) and IEC (the International Electrotechnical Commission) form the specialised system for world-wide standardisation. National Bodies that are members of ISO and IEC participate in the development of International Standards through technical committees established by the respective organisation to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organisations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75% of the national bodies casting a vote.

This specification is a committee draft that is being submitted for approval to the ITU-T, ISO-IEC/JTC1 SC29. It was prepared jointly by SC29/WG11, also known as MPEG (Moving Pictures Expert Group), and the Experts Group for ATM Video Coding in the ITU-T SG15. MPEG was formed in 1988 to establish standards for coding of moving pictures and associated audio for various applications such as digital storage media, distribution and communication. The Experts Group for ATM Video Coding was formed in 1990 to develop video coding standard(s) appropriate for B-ISDN using ATM transport.
In this specification Annex A, Annex B and Annex C contain normative requirements and are an integral part of this specification. Annex D, Annex E, Annex F and Annex G are informative and contain no normative requirements.

<u>ISO/IEC</u>
This International Standard is published in four Parts.

13818-1 Systems	specifies the system coding of the specification. It defines a multiplexed structure for combining audio and video data and means of representing the timing information needed to replay synchronised sequences in real-time.
13818-2 Video	specifies the coded representation of video data and the decoding process required to reconstruct pictures.
13818-3 Audio	specifies the coded representation of audio data.
13818-4 Compiance	specifies the procedures for determining the characteristics of coded bitstreams and for testing compliance with the requirements stated in 13818-1, 13818-2 and 13818-3.

I. Introduction

Parts 1, 2 and 3 of ISO/IEC 13818 specify a multiplex structure and coded representations of audiovisual information. Parts 1, 2 and 3 of ISO/IEC 13818 allow for large flexibility, achieving suitability of this International Standard for many different applications. The flexibility is obtained by including parameters in the bitstream that define the characteristics of coded bitstreams. Examples are the audio sampling frequency, picture size, picture rate and bitrate parameters.

This part of ISO/IEC 11172 specifies how tests can be designed to verify whether bitstreams and decoders meet the requirements as specified in parts 1, 2 and 3 of ISO/IEC 13818. These tests can be used for various purposes such as:

- manufacturers of encoders, and their customers, can use the tests to verify whether the encoder produces valid bitstreams.

- manufacturers of decoders and their customers can use the tests to verify whether the decoder meets the requirements specified in parts 1,2 and 3 of ISO/IEC 13818 for the claimed decoder capabilities.

- applications can use the tests to verify whether the characteristics of a given bitstream meet the application requirements, for example whether the size of the coded picture does not exceed the maximum value allowed for the application.

INTERNATIONAL STANDARD 13818-4

ITU-T RECOMMENDATION H.262

INFORMATION TECHNOLOGY -

GENERIC CODING OF MOVING PICTURES AND ASSOCIATED AUDIO

I. Scope

This Recommendation | International Standard specifies the coded representation of picture information for digital storage media and digital video communication and specifies the decoding process. The representation supports constant bitrate transmission, variable bitrate transmission, random access, channel hopping, scalable decoding, bitstream editing, as well as special functions such as fast forward playback, fast reverse playback, slow motion, pause and still pictures. This Recommendation | International Standard is forward compatible with ISO/IEC13818-2 and upward or downward compatible with EDTV, HDTV, SDTV formats.

This Recommendation | International Standard is primarily applicable to digital storage media, video broadcast and communication. The storage media may be directly connected to the decoder, or via communications means such as busses, LANs, or telecommunications links.

I. Normative references

The following ITU-T 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 editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards. The TSB (Telecommunication Standardisation Bureau) maintains a list of currently valid ITU-T Recommendations.

´ Recommendations and reports of the CCIR, 1990

XVIIth Plenary Assembly, Dusseldorf, 1990 Volume XI - Part 1

Broadcasting Service (Television) Rec. 601-2 "Encoding parameters of digital television for studios"

´ CCIR Volume X and XI Part 3 Recommendation 648: Recording of audio signals.

´ CCIR Volume X and XI Part 3 Report 955-2: Sound broadcasting by satellite for portable and mobile receivers, including Annex IV Summary description of advanced digital system II.

´ ISO/IEC 11172 (1993) "Information technology „ Coding of moving picture and associated audio for digital storage media at up to about 1,5 Mbit/s"

´ IEEE Standard Specifications for the Implementations of 8 by 8 Inverse Discrete Cosine Transform, IEEE Std 1180-1990, December 6, 1990.

´ IEC Publication 908:198, "CD Digital Audio System"

´ IEC Standard Publication 461 Second edition, 1986 "Time and control code for video tape recorders"

´ ITU-T Recommendation H.261 (Formerly CCITT Recommendation H.261) ñCodec for audiovisual services at px64 kbit/s" Geneva, 1990

´ ISO/IEC 10918-1 | ITU-T Rec. T.81 (JPEG) "Digital compression and coding of continuous-tone still images"

I.

II. Definitions

For the purposes of this Recommendation | International Standard, the following definitions apply.

1. **ac coefficient [video]**: Any DCT coefficient for which the frequency in one or both dimensions is non-zero.
2. **AC coefficient**: Any DCT coefficient for which the frequency in one or both dimensions is non-zero.
3. **access unit [system]**: A coded representation of a presentation unit. In the case of compressed audio an access unit is an Audio Access Unit. In the case of compressed video an access unit is the coded representation of a picture.
4. **adaptive bit allocation [audio]**: The assignment of bits to subbands in a time and frequency varying fashion according to a psychoacoustic model.
5. **adaptive multichannel prediction [audio]**: A method of multichannel data reduction exploiting statistical inter-channel dependencies.
6. **adaptive noise allocation [audio]**: The assignment of coding noise to frequency bands in a time and frequency varying fashion according to a psychoacoustic model.
7. **adaptive segmentation [audio]**: A subdivision of the digital representation of an audio signal in variable segments of time.
8. **alias [audio]**: Mirrored signal component resulting from sub-Nyquist sampling.
9. **almost compliant:** the degree to which a given decoder fails to decode "evil" bitstreams.
10. **analysis filterbank [audio]**: Filterbank in the encoder that transforms a broadband PCM audio signal into a set of subsampled subband samples.
11. **audio access unit [audio]**: For Layers I and II, an audio access unit is defined as the smallest part of the encoded bit stream which can be decoded by itself, where decoded means "fully reconstructed sound". For Layer III, an audio access unit is part of the bit stream that is decodable with the use of previously acquired main information.
12. **audio buffer [audio]**: A buffer in the system target decoder for storage of compressed audio data.
13. **audio sequence [audio]**: A non-interrupted series of audio frames in which the following parameters are not changed:- ID- Layer- Sampling Frequency- For Layer I and II: Bitrate index
14. **B-field picture**: A field structure B-Picture.
15. **B-frame picture**: A frame structure B-Picture.
16. **B-picture**; **bidirectionally predictive-coded picture:** A picture that is coded using motion compensated prediction from past and/or future reference fields or frames.
17. **backward compatibility**: A newer coding standard is backward compatible with an older coding standard if decoders designed to operate with the older coding standard are able to continue to operate by decoding all or part of a bitstream produced according to the newer coding standard.
18. **backward motion vector [video]**: A motion vector that is used for motion compensation from a reference picture at a later time in display order.
19. **backward motion vector**: A motion vector that is used for motion compensation from a reference frame or reference field at a later time in display order.
20. **Bark [audio]**: Unit of critical band rate. The Bark scale is a non-linear mapping of the frequency scale over the audio range closely corresponding with the frequency selectivity of the human ear across the band.
21. **bidirectionally predictive-coded picture; B-picture [video]**: A picture that is coded using motion compensated prediction from a past and/or future reference picture.
22. **bitrate:** The rate at which the coded bitstream is delivered from the storage medium to the input of a decoder.
23. **block companding [audio]**: Normalising of the digital representation of an audio signal within a certain time period.
24. **block:** An 8-row by 8-column matrix of samples, or 64 DCT coefficients (source, quantised or dequantised).
25. **bottom field:** One of two fields that comprise a frame. Each line of a bottom field is spatially located immediately below the corresponding line of the top field.
26. **bound [audio]**: The lowest subband in which intensity stereo coding is used.
27. **byte aligned**: A bit in a coded bitstream is byte-aligned if its position is a multiple of 8-bits from the first bit in the stream.
28. **byte**: Sequence of 8-bits.
29. **center channel [audio]**: An audio presentation channel used to stabilise the central component of the frontal stereo image.

30. **channel [audio]**: A sequence of data representing an audio signal being transported.
31. **channel**: A digital medium that stores or transports a CD 13818 bit stream.
32. **chroma simulcast**: A type of scalability (which is a subset of SNR scalability) where the enhancement layer (s) contain only coded refinement data for the DC coefficients, and all the data for the AC coefficients, of the chrominance components.
33. **chrominance format:** Defines the number of chrominance blocks in a macroblock.
34. **chrominance (component) [video]**: A matrix, block or single pel representing one of the two colour difference signals related to the primary colours in the manner defined in CCIR Rec 601. The symbols used for the colour difference signals are Cr and Cb.
35. **citical band rate [audio]**: Psychoacoustic function of frequency. At a given audible frequency, it is proportional to the number of critical bands below that frequency. The units of the critical band rate scale are Barks.
36. **cocding parameters**: The set of user-definable parameters that characterise a coded video bitstream. Bitstreams are characterised by coding parameters. Decoders are characterised by the bitstreams that they are capable of decoding.
37. **coded audio bit stream [audio]**: A coded representation of an audio signal as specified in this part of the CD.
38. **coded B-frame**: A B-frame picture or a pair of B-field pictures.
39. **coded frame**: A coded frame is a coded I-frame, a coded P-frame or a coded B-frame.
40. **coded I-frame**: An I-frame picture or a pair of field pictures, where the first one is an I-picture and the second one is an I-picture or a P-picture.
41. **coded order [video]**: The order in which the pictures are stored and decoded. This order is not necessarily the same as the display order.
42. **coded order**: The order in which the pictures are transmitted and decoded. This order is not necessarily the same as the display order.
43. **coded P-frame**: A P-frame picture or a pair of P-field pictures.
44. **coded picture:** A coded picture is made of a picture header, the optionnal extensions immediately following it, and the following picture data. A coded picture may be a frame picture or a field picture.
45. **coded representation**: A data element as represented in its encoded form.
46. **coded video bit stream [video]**: A coded representation of a series of one or more pictures as specified in this CD.
47. **coding parameters [video]**: The set of user-definable parameters that characterise a coded video bit stream. Bit streams are characterised by coding parameters. Decoders are characterised by the bit streams that they are capable of decoding.
48. **component [video]**: A matrix, block or single pel from one of the three matrices (luminance and two chrominance) that make up a picture.
49. **component**: A matrix, block or single sample from one of the three matrices (luminance and two chrominance) that make up a picture.
50. **compression**: Reduction in the number of bits used to represent an item of data.
51. **constant bitrate coded video [video]**: A compressed video bit stream with a constant average bitrate.
52. **constant bitrate:** Operation where the bitrate is constant from start to finish of the coded bitstream.
53. **constrained parameters [video]**: The values of the set of coding parameters defined in 2.4.3.2 of ISO/IEC 11172-2.
54. **constrained system parameter stream (CSPS) [system]**: An ISO/IEC 11172 multiplexed stream for which the constraints defined in 2.4.6 of ISO/IEC 11172-1 apply.
55. **CRC**: The Cyclic Redundancy Check to verify the correctness of data.
56. **critical band [audio]**: Psychoacoustic measure in the spectral domain which corresponds to the frequency selectivity of the human ear. This selectivity is expressed in Bark.
57. **D-Picture**: A type of picture that shall not be used except in ISO/IEC 11172-2.
58. **data element**: An item of data as represented before encoding and after decoding.
59. **data partitioning**: A method for dividing a bitstream into two separate bitstreams for error resilience purposes. the two bitstreams have to be recombined before decoding.
60. **DC coefficient**: The DCT coefficient for which the frequency is zero in both dimensions.
61. **dc-coded picture; D-picture [video]**: A picture that is coded using only information from itself. Of the DCT coefficients in the coded representation, only the dc-coefficients are present.
62. **dc-coefficient [video]**: The DCT coefficient for which the frequency is zero in both dimensions.

63. **DCT coefficient**: The amplitude of a specific cosine basis function.
64. **de-emphasis [audio]**: Filtering applied to an audio signal after storage or transmission to undo a linear distortion due to emphasis.
65. **decoded stream**: The decoded reconstruction of a compressed bit stream.
66. **decoder input buffer [video]**: The first-in first-out (FIFO) buffer specified in the video buffering verifier.
67. **decoder input rate [video]**: The data rate specified in the video buffering verifier and encoded in the coded video bit stream.
68. **decoder**: An embodiment of a decoding process.
69. **decoding (process):** The process defined in this specification that reads an input coded bitstream and produces decoded pictures or audio samples.
70. **decoding time-stamp; DTS [system]**: A field that may be present in a PES packet header that indicates the time that an access unit is decoded in the system target decoder.
71. **dequantisation [video]**: The process of rescaling the quantised DCT coefficients after their representation in the bit stream has been decoded and before they are presented to the inverse DCT.
72. **dequantisation**: The process of rescaling the quantised DCT coefficients after their representation in the bitstream has been decoded and before they are presented to the inverse DCT.
73. **digital storage media; DSM**: A digital storage or transmission device or system.
74. **discrete cosine transform; DCT**: Either the forward discrete cosine transform or the inverse discrete cosine transform. The DCT is an invertible, discrete orthogonal transformation. The inverse DCT is defined in Annex A of this specification.
75. **display order [video]**: The order in which the decoded pictures should be displayed. Normally this is the same order in which they were presented at the input of the encoder.
76. **downmix [audio]**: A matrixing of n channels to obtain less than n channels.
77. **dual channel mode [audio]**: A mode, where two audio channels with independent programme contents (e.g. bilingual) are encoded within one bit stream. The coding process is the same as for the stereo mode.
78. **dynamic crosstalk [audio]**: A method of multichannel data reduction in which stereo-irrelevant signal components are copied to another channel.
79. **dynamic transmission channel switching [audio]**: A method of multichannel data reduction by allocating the most orthogonal signal components to the transmission channels.
80. **editing**: The process by which one or more compressed bit streams are manipulated to produce a new compressed bit stream. Conforming edited bit streams must meet the requirements defined in this CD.
81. **elementary stream; ES [system]**: A generic term for one of the coded video, coded audio or other coded bit streams.
82. **emphasis [audio]**: Filtering applied to an audio signal before storage or transmission to improve the signal-to-noise ratio at high frequencies.
83. **encoder**: An embodiment of an encoding process.
84. **encoding (process):** A process, not specified in this specification, that reads a stream of input pictures or audio samples and produces a valid coded bitstream as defined in this specification.
85. **entitlement control message; ECM:** Entitlement Control Messages are private conditional access information which specify control words and possibly other, typically stream-specific, scrambling and and/or control parameters.
86. **entitlement management message; EMM:** Entitlement Management Messages are private conditional access information which specify the authorization levels or the services of specific decoders. They may be addressed to single decoders or groups of decoders.
87. **entropy coding**: Variable length lossless coding of the digital representation of a signal to reduce redundancy.
88. **evil bitstreams:** bitstreams orthogonal to reality.
89. **fast reverse playback**: The process of displaying the picture sequence in the reverse of display order faster than real-time.
90. **fast forward playback [video]**: The process of displaying a sequence, or parts of a sequence, of pictures in display-order faster than real-time.
91. **FFT**: Fast Fourier Transformation. A fast algorithm for performing a discrete Fourier transform (an orthogonal transform).
92. **field period**: The reciprocal of twice the frame rate.

93. **field picture; field structure picture** : A field structure picture is a coded picture with picture_structure is equal to "Top field" or "Bottom field".
94. **field:** For an interlaced video signal, a ñfieldî is the assembly of alternate lines of a frame. Therefore an interlaced frame is composed of two fields, a top field and a bottom field.
95. **filterbank [audio]**: A set of band-pass filters covering the entire audio frequency range.
96. **fixed segmentation [audio]**: A subdivision of the digital representation of an audio signal into fixed segments of time.
97. **flag**: A variable which can take one of only the two values defined in this specification.
98. **forbidden**: The term "forbidden" when used in the clauses defining the coded bit stream indicates that the value shall never be used. This is usually to avoid emulation of start codes.
99. **forced updating [video]**: The process by which macroblocks are intra-coded from time-to-time to ensure that mismatch errors between the inverse DCT processes in encoders and decoders cannot build up excessively.
100. **forward compatibility**: A newer coding standard is forward compatible with an older coding standard if decoders designed to operate with the newer coding standard are able to decode bitstreams of the older coding standard.
101. **forward motion vector [video]**: A motion vector that is used for motion compensation from a reference picture at an earlier time in display order.
102. **frame [audio]**: A part of the audio signal that corresponds to audio PCM samples from an Audio Access Unit.
103. **frame period**: The reciprocal of the frame rate.
104. **frame picture; frame structure picture** : A frame structure picture is a coded picture with picture_structure is equal to "Frame".
105. **frame rate**: The rate at which frames are be output from the decoding process.
106. **frame:** A frame contains lines of spatial information of a video signal. For progressive video, these lines contain samples starting from one time instant and continuing through successive lines to the bottom of the frame. For interlaced video a frame consists of two fields, a top field and a bottom field. One of these fields will commence one field period later than the other.
107. **free format [audio]**: Any bitrate other than the defined bitrates that is less than the maximum valid bitrate for each layer.
108. **future reference frame (field)**: A future reference frame(field) is a reference frame(field) that occurs at a later time than the current picture in display order.
109. **future reference picture [video]**: The future reference picture is the reference picture that occurs at a later time than the current picture in display order.
110. **granules [Layer III] [audio]**: 576 frequency lines that carry their own side information.
111. **group of pictures [video]**: A series of one or more coded pictures intended to assist random access. The group of pictures is one of the layers in the coding syntax defined in ISO/IEC 11172-2.
112. **Hann window [audio]**: A time function applied sample-by-sample to a block of audio samples before Fourier transformation.
113. **header**: A block of data in the coded bitstream containing the coded representation of a number of data elements pertaining to the coded data that follow the header in the bitstream.
114. **Huffman coding**: A specific method for entropy coding.
115. **hybrid filterbank [audio]**: A serial combination of subband filterbank and MDCT.
116. **hybrid scalability:** Hybrid scalability is the combination of two (or more) types of scalability.
117. **I-field picture**: A field structure I-Picture.
118. **I-frame picture**: A frame structure I-Picture.
119. **I-picture; intra-coded picture**: A picture coded using information only from itself.
120. **IMDCT [audio]**: Inverse Modified Discrete Cosine Transform.
121. **intensity stereo [audio]**: A method of exploiting stereo irrelevance or redundancy in stereophonic audio programmes based on retaining at high frequencies only the energy envelope of the right and left channels.
122. **interlace:** The property of conventional television frames where alternating lines of the frame represent different instances in time. In an interlaced frame, one of the field is meant to be displayed first. This field is called the first field. The first field can be the top field or the bottom field of the frame.
123. **intra coding [video]**: Coding of a macroblock or picture that uses information only from that macroblock or picture.
124. **intra-coded picture; I-picture [video]**: A picture coded using information only from itself.

125. **ISO/IEC 11172 (multiplexed) stream [system]**: A bit stream composed of zero or more elementary streams combined in the manner defined in ISO/IEC 11172-1.
126. **ISO/IEC 13818 (multiplexed) stream [system]**: A bit stream composed of 0 or more elementary streams combined in the manner defined in Part 1 of this Recommendation | International Standard.
127. **joint stereo coding [audio]**: Any method that exploits stereophonic irrelevance or stereophonic redundancy.
128. **joint stereo mode [audio]**: A mode of the audio coding algorithm using joint stereo coding.
129. **layer [audio]**: One of the levels in the coding hierarchy of the audio system defined in this part of the CD.
130. **layer [video and systems]**: One of the levels in the data hierarchy of the video and system specifications defined in ISO/IEC 11172-1 and ISO/IEC 11172-2.
131. **level** : A defined set of constraints on the values which may be taken by the parameters of this specification within a particular profile. A profile may contain one or more levels.
132. **low frequency enhancement channel [audio]**: A limited bandwidth channel for low frequency audio effects in a multichannel system.
133. **luminance (component) [video]**: A matrix, block or single pel representing a monochrome representation of the signal and related to the primary colours in the manner defined in CCIR Rec 601. The symbol used for luminance is Y.
134. **macroblock**: The four 8 by 8 blocks of luminance data and the two (for 4:2:0 chrominance format), four (for 4:2:2 chrominance format) or eight (for 4:4:4 chrominance format) corresponding 8 by 8 blocks of chrominance data coming from a 16 by 16 section of the luminance component of the picture. Macroblock is sometimes used to refer to the sample data and sometimes to the coded representation of the sample values and other data elements defined in the macroblock header of the syntax defined in this part of this specification. The usage is clear from the context.
135. **mapping [audio]**: Conversion of an audio signal from time to frequency domain by subband filtering and/or by MDCT.
136. **masking [audio]**: A property of the human auditory system by which an audio signal cannot be perceived in the presence of another audio signal .
137. **masking threshold [audio]**: A function in frequency and time below which an audio signal cannot be perceived by the human auditory system.
138. **MDCT [audio]**: Modified Discrete Cosine Transform which corredpond to the Time Domain Aliasing Cancellation Filter Bank.
139. **motion compensation**: The use of motion vectors to improve the efficiency of the rediction of sample values. The prediction uses motion vectors to provide offsets into the past and/or future reference frames or reference fields containing previously decoded sample values that are used to form the prediction error signal.
140. **motion estimation**: The process of estimating motion vectors during the encoding process.
141. **motion vector [video]**: A two-dimensional vector used for motion compensation that provides an offset from the coordinate position in the current picture to the coordinates in a reference picture.
142. **MS stereo [audio]**: A method of exploiting stereo irrelevance or redundancy in stereophonic audio programmes based on coding the sum and difference signal instead of the left and right channels.
143. **multichannel [audio]**: A combination of audio channels used to create a spatial sound field.
144. **multilingual [audio]**: A presentation of dialogue in more than one language.
145. **naughty bitstreams:** is less than "evil" but has some non-zero projection into reality.
146. **nice bitstream:** a bitstream which can be decoded by an almost-compliant decoder.
147. **non-intra coding**: Coding of a macroblock or picture that uses information both from itself and from macroblocks and pictures occurring at other times.
148. **non-tonal component [audio]**: A noise-like component of an audio signal.
149. **Nyquist sampling**: Sampling at or above twice the maximum bandwidth of a signal.
150. **P-field picture**: A field structure P-Picture.
151. **P-frame picture**: A frame structure P-Picture.
152. **P-picture; predictive-coded picture** : A picture that is coded using motion compensated prediction from past reference fields or frame.
153. **pack [system]**: A pack consists of a pack header followed by zero or more packets. It is a layer in the system coding syntax described in clause **Feil! Fant ikke referansekilden.**this Recommendation | International Standard.

154. **packet [system]**: A packet consists of a header followed by a number of contiguous bytes from an elementary data stream. It is a layer in the system coding syntax described in clause [XX] of this Recommendation | International Standard.
155. **packet data [system]**: Contiguous bytes of data from an elementary stream present in a packet.
156. **packet header [system]**: The data structure used to convey information about the elementary stream data contained in the packet data.
157. **packet identifier; PID [system]**: A unique integer value used to associate elementary streams of a program in a single or multi-program Transport Stream as described in [XXXX]
158. **padding [audio]**: A method to adjust the average length of an audio frame in time to the duration of the corresponding PCM samples, by conditionally adding a slot to the audio frame.
159. **parameter**: A variable within the syntax of this specification which may take one of a large range of values. A variable which can take one of only two values is a flag and not a parameter.
160. **past reference frame (field)**: A past reference frame(field) is a reference frame(field) that occurs at an earlier time than the current picture in display order.
161. **past reference picture [video]**: The past reference picture is the reference picture that occurs at an earlier time than the current picture in display order.
162. **pel [video]**: Picture element.
163. **pel aspect ratio [video]**: The ratio of the nominal vertical height of pel on the display to its nominal horizontal width.
164. **PES [system]:** An abbreviation for Packetized Elementary Stream.
165. **PES packet [system]**: The data structure used to carry elementary stream data. It consists of a PES packet header followed by PES packet payload and is described in clause [XXX]
166. **PES Stream [system]:** A PES Stream consists of PES packets, all of whose payloads consist of data from a single elementary stream, and all of which have the same stream_id. Specific semantic constraints apply.
167. **picture period [video]**: The reciprocal of the picture rate.
168. **picture rate [video]**: The nominal rate at which pictures should be output from the decoding process.
169. **picture**: Source, coded or reconstructed image data. A source or reconstructed picture consists of three rectangular matrices of 8-bit numbers representing the luminance and two chrominance signals. For progressive video, a picture is identical to a frame, while for interlaced video, a picture can refer to a frame, or the top field or the bottom field of the frame depending on the context.
170. **polyphase filterbank [audio]**: A set of equal bandwidth filters with special phase interrelationships, allowing for an efficient implementation of the filterbank.
171. **prediction [audio]**: The use of a predictor to provide an estimate of the subband sample in one channel from the subband samples in other channels.
172. **prediction [video]**: The use of a predictor to provide an estimate of the pel value or data element currently being decoded.
173. **prediction error [video]**: The difference between the actual value of a pel or data element and its predictor.
174. **prediction**: The use of a predictor to provide an estimate of the sample value or data element currently being decoded.
175. **predictive-coded picture; P-picture**: A picture that is coded using motion compensated prediction from past reference frames or reference fields.
176. **predictor**: A linear combination of previously decoded sample values or data elements.
177. **presentation channel [audio]**: audio channels at the output of the decoder corresponding to the loudspeaker positions left, center, right, left surround and right surround.
178. **presentation time-stamp; PTS [system]**: A field that may be present in a packet header that indicates the time that a presentation unit is presented in the system target decoder.
179. **presentation unit; PU [system]**: A decoded audio access unit or a decoded picture.
180. **profile**: A defined subset of the syntax of this specification.
181. **program [system]:** A program is a collection of elementary streams with a common time base.
182. **Program Specific Information; PSI [system]:** PSI consists of normative data which is necessary for the demultiplexing of Transport Streams and the successful regeneration of programs and is described in clause [XX]. One case of PSI, the non-mandatory network information table, is privately defined.
183. **progressive:** The property of film frames where all the samples of the frame represent the same instances in time.

184. **psychoacoustic model [audio]**: A mathematical model of the masking behaviour of the human auditory system.
185. **quantisation matrix [video]**: A set of sixty-four 8-bit values used by the dequantiser.
186. **quantised DCT coefficients**: DCT coefficients before dequantisation. A variable length coded representation of quantised DCT coefficients is transmitted as part of the compressed video bitstream.
187. **quantiser scalefactor [video]**: A data element represented in the bit stream and used by the decoding process to scale the dequantisation.
188. **random access**: The process of beginning to read and decode the coded bit stream at an arbitrary point.
189. **reconstructed frame:** A reconstructed frame consists of three rectangular matrices of 8-bit numbers representing the luminance and two chrominance signals. A reconstructed frame is obtained by decoding a coded frame.
190. **reconstructed picture:** A reconstructed picture is obtained by decoding a coded picture. A reconstructed picture is either a reconstructed frame (when decoding a frame picture), or one field of a reconstructed frame (when decoding a field picture). If the coded picture is a field picture, then the reconstructed picture is the top field or the bottom field of the reconstructed frame.
191. **reference field:** A reference field is one field of a reconstructed frame. Reference fields are used for forward and backward prediction when P-pictures and B-pictures are decoded. Note that when field P-pictures are decoded, prediction of the second field P-picture of a coded frame uses the first reconstructed field of the same coded frame as a reference field.
192. **reference frame:** A reference frame is a reconstructed frame that was coded in the form of a coded I-frame or a coded P-frame. Reference frames are used for forward and backward prediction when P-pictures and B-pictures are decoded.
193. **reference picture [video]**: Reference pictures are the nearest adjacent I- or P-pictures to the current picture in display order.
194. **reorder buffer [video]**: A buffer in the system target decoder for storage of a reconstructed I-picture or a reconstructed P-picture.
195. **requantisation [audio]**: Decoding of coded subband samples in order to recover the original quantised values.
196. **reserved**: The term "reserved" when used in the clauses defining the coded bitstream indicates that the value may be used in the future for ISO/IEC defined extensions.
197. **reverse playback [video]**: The process of displaying the picture sequence in the reverse of display order.
198. **sample aspect ratio:** (abbreviated to **SAR**). This specifies the distance between samples. It is defined (for the purposes of this specification) as the vertical displacement of the lines of luminance samples in a frame divided by the horizontal displacement of the luminance samples. Thus its units are (metres per line) … (metres per sample)
199. **scalability:** Scalability is the ability of a decoder to decode an ordered set of bitstreams to produce a reconstructed sequence. Moreover, useful video is output when subsets are decoded. The minimum subset that can thus be decoded is the first bitstream in the set which is called the base layer. Each of the other bitstreams in the set is called an enhancement layer. When addressing a specific enhancement layer, "lower layer" refer to the bitstream which precedes the enhancement layer.
200. **scalefactor [audio]**: Factor by which a set of values is scaled before quantisation.
201. **scalefactor band [audio]**: A set of frequency lines in Layer III which are scaled by one scalefactor.
202. **scalefactor index [audio]**: A numerical code for a scalefactor.
203. **sequence header [video]**: A block of data in the coded bit stream containing the coded representation of a number of data elements.
204. **side information:** Information in the bitstream necessary for controlling the decoder.
205. **skipped macroblock [video]**: A macroblock for which no data are stored.
206. **slice**: A series of macroblocks.
207. **slot [audio]**: A slot is an elementary part in the bit stream. In Layer I a slot equals four bytes, in Layers II and III one byte.
208. **SNR scalability**: A type of scalability where the enhancement layer (s) contain only coded refinement data for the DCT coefficients of the lower layer.
209. **source stream**: A single non-multiplexed stream of samples before compression coding.

210. **spatial scalability**: A type of scalability where an enhancement layer also uses predictions from sample data derived from a lower layer without using motion vectors. The layers can have different frame sizes, frame rates or chrominance formats
211. **spreading function [audio]**: A function that describes the frequency spread of masking effects.
212. **start codes [system and video]**: 32-bit codes embedded in that coded bitstream that are unique. They are used for several purposes including identifying some of the structures in the coding syntax.
213. **STD input buffer [system]**: A first-in first-out buffer at the input of the system target decoder for storage of compressed data from elementary streams before decoding.
214. **stereo mode [audio]**: Mode, where two audio channels which form a stereo pair (left and right) are encoded within one bit stream. The coding process is the same as for the dual channel mode.
215. **stereo-irrelevant [audio]**: a portion of a stereophonic audio signal which does not contribute to spatial perception.
216. **stuffing (bits); stuffing (bytes)** : Code-words that may be inserted into the compressed bit stream that are discarded in the decoding process. Their purpose is to increase the bitrate of the stream.
217. **subband [audio]**: Subdivision of the audio frequency band.
218. **subband filterbank [audio]**: A set of band filters covering the entire audio frequency range. In this part of the CD, the subband filterbank is a polyphase filterbank.
219. **subband samples [audio]**: The subband filterbank within the audio encoder creates a filtered and subsampled representation of the input audio stream. The filtered samples are called subband samples. From 384 time-consecutive input audio samples, 12 time-consecutive subband samples are generated within each of the 32 subbands.
220. **surround channel [audio]**: An audio presentation channel added to the front channels (L and R or L, R, and C) to enhance the spatial perception.
221. **syncword [audio]**: A 12-bit code embedded in the audio bit stream that identifies the start of a frame.
222. **synthesis filterbank [audio]**: Filterbank in the decoder that reconstructs a PCM audio signal from subband samples.
223. **system header [system]**: The system header is a data structure defined in clause of this Recommendation | International Standard that carries information summarizing the system characteristics of the ISO/IEC 13818 multiplexed stream.
224. **system target decoder; STD [system]**: A hypothetical reference model of a decoding process used to describe the semantics of an ISO/IEC 13818 multiplexed bit stream.
225. **temporal scalability**: A type of scalability where an enhancement layer also uses predictions from sample data derived from a lower layer using motion vectors. The layers have identical frame size, and chrominance formats, but can have different frame rates.
226. **time-stamp [system]**: A term that indicates the time of an event.
227. **tonal component [audio]**: A sinusoid-like component of an audio signal.
228. **top field:** One of two fields that comprise a frame. Each line of a top field is spatially located immediately above the corresponding line of the bottom field.
229. **Transport Stream packettransport packet header [system]:** A data structure used to convey information about the Transport Stream payload.
230. **triplet [audio]**: A set of 3 consecutive subband samples from one subband. A triplet from each of the 32 subbands forms a granule.
231. **variable bitrate**: Operation where the bitrate varies with time during the decoding of a compressed bit stream.
232. **variable length coding; VLC**: A reversible procedure for coding that assigns shorter code-words to frequent events and longer code-words to less frequent events.
233. **video buffering verifier; VBV**: A hypothetical decoder that is conceptually connected to the output of the encoder. Its purpose is to provide a constraint on the variability of the data rate that an encoder or editing process may produce.
234. **video sequence**: The highest syntactic structure of coded video bitstreams. It contains a series of one or more coded frames.
235. **zig-zag scanning order**: A specific sequential ordering of the DCT coefficients from (approximately) the lowest spatial frequency to the highest.

I.

II. Abbreviations and symbols

The mathematical operators used to describe this specification are similar to those used in the C programming language. However, integer divisions with truncation and rounding are specifically defined. Numbering and counting loops generally begin from zero.

A. Arithmetic operators

+ Addition.

- Subtraction (as a binary operator) or negation (as a unary operator).

++ Increment.

- - Decrement.

× * Multiplication.

^ Power.

/ Integer division with truncation of the result toward zero. For example, 7/4 and -7/-4 are truncated to 1 and -7/4 and 7/-4 are truncated to -1.

// Integer division with rounding to the nearest integer. Half-integer values are rounded away from zero unless otherwise specified. For example 3//2 is rounded to 2, and -3//2 is rounded to -2.

DIV Integer division with truncation of the result toward minus infinity. For example 3 DIV 2 is rounded to 1, and -3 DIV 2 is rounded to -2.

… Used to denote division in mathematical equations where no truncation or rounding is intended.

% Modulus operator. Defined only for positive numbers.

Sign( )

Abs( )

A. Logical operators

|| Logical OR.

&& Logical AND.

! Logical NOT.

A. Relational operators

> Greater than.

>= Greater than or equal to.

< Less than.

<= Less than or equal to.

== Equal to.

!= Not equal to.

max [,...,] the maximum value in the argument list.

min [,...,] the minimum value in the argument list.

A. Bitwise operators

& AND

| OR

>> Shift right with sign extension.

<< Shift left with zero fill.

A. Assignment

= Assignment operator.

A. Mnemonics

The following mnemonics are defined to describe the different data types used in the coded bit stream.

bslbf Bit string, left bit first, where "left" is the order in which bit strings are written in ISO/IEC 11172. Bit strings are written as a string of 1s and 0s within single quote marks, e.g. '1000 0001'. Blanks within a bit string are for ease of reading and have no significance.

center_chan	Index of center channel.
center_limited	Variable which indicates whether a subband of the center is not transmitted. It is used in the case of phantom coding of center channel.
ch	Channel. If ch has the value 0, the left channel of a stereo signal or the first of two independent signals is indicated. (Audio)
gr	Granule of 3 * 32 subband samples in audio Layer II, 18 * 32 subband samples in audio Layer III. (Audio)
L, C, R, LS, RS	Left, center, right, left surround and right surround audio signals
L^W, C^W, R^W, LS^W, RS^W	Weighted left, center, right, left surround and right surround audio signals. The weighting is necessary for two reasons: 1) All signals have to be attenuated prior to encoding to avoid overload when calculating the compatible stereo signal. 2)The weighted and processed signals are actually coded and transmitted, and denormalised in the decoder
left_sur_chan	Index of left surround channel.
main_data	The main_data portion of the bit stream contains the scalefactors, Huffman encoded data, and ancillary information. (Audio)
mono_sur_chan	Index of the mono surround channel. This index is identical to the index of the left surround channel. (Audio)
msblimit	Maximum used subband
nch	Number of channels; equal to 1 for single_channel mode, 2 in other modes. (Audio)
nmch	Number of channels in the multichannel extension part
dyn_cross	dyn_cross means that dynamic crosstalk is used for a certain transmission channel and a certain subband.
npredcoeff	Number of prediction coefficients used.
part2_length	The number of main_data bits used for scalefactors. (Audio)
right_sur_chan	Index of right surround channel.
rpchof	Remainder polynomial coefficients, highest order first. (Audio)
sb	Subband. (Audio)
sbgr	Groups of individual subband according to subbandgroup table in subclause 3.5.2.10
sblimit	The number of the lowest subband for which no bits are allocated. (Audio)
scfsi	Scalefactor selection information. (Audio)
switch_point_l	Number of scalefactor band (long block scalefactor band) from which point on window switching is used. (Audio)
switch_point_s	Number of scalefactor band (short block scalefactor band) from which point on window switching is used. (Audio)
tc	Transmitted channel. (Audio)
uimsbf	Unsigned integer, most significant bit first.
vlclbf	Variable length code, left bit first, where "left" refers to the order in which the VLC codes are written.
window	Number of the actual time slot in case of block_type==2, 0 <= window <= 2. (Audio)

The byte order of multi-byte words is most significant byte first.

A. Constants

π	3,14159265359...
e	2,71828182845...

I.

II. Bitstream characteristics

Bitstream characteristics specify the constraints that are applied by the encoder in generating the bitstream. These syntactic and semantic constraints may for example restrict the range or the values of parameters that are encoded directly or indirectly in the bitstream. The constraints applied to a given bitstream may or may not be known *a priori* .

A. System bitstreams

System encoders may apply restrictions to the following parameters of system bitstreams (see ITU-T Rec. H.222.0 | ISO/IEC 13818-1):

1. Transport Streams

Reminder of Normative systems Requirements

Minimum frequency of encoding:
PCR / SCR
PTS / DTS

Maximum STC frequency variation
27 000 000 +/- 20 ppm
+/- 75 Hz per second

Maximum PCR tolerance
+/- 500 nsec

Timing-related Performance Measures

Maximum variation in audio/video synchronization:
how much can display of an audio AU be advanced with respect to an "overlapping" video AU?
How much can display of an audio AU be delayed with respect to an "overlapping" video AU?

Stability of derived decoder clock, across entire range of allowed PCR tolerance, STC frequency, and STC frequency drift and at minimum allowed PCR encoding rate

Acquisition time needed after start-up to reach specified decoder clock stability, at minimum allowed PCR encoding rate

Is handling of STC discontinuities supported while continuing to meet all other performance measures? How often can STC discontinuities be allowed while continuing to meet all other performance measures?

Does the decoder properly display video fields during trick-mode playback? Which trick modes are supported?

At what rate are AU«s discarded? At what rate are AU«s displayed more than once, (other than video AU«s in low-delay mode)? (These rates may be 0.)

Processing Capabilities

How often can the decoder process the following syntax elements?
PS pack headers
TS adaptation fields

PES packet headers
TS PSI sections
PS program_stream_directory/maps

How quickly can the decoder respond to new PSI sections?

How many TS PID«s can be processed simultaneously?
How many PS stream_id«s can be processed simultaneously?
Does the decoder conceal non-seamless splices? How?

a) Transport Stream basic functions

Table 1 -- Transport Stream basic functions

function	layer	syntax element
packet synchronization	TS	sync_byte
	PES	packet_start_code_prefix + stream_id *TS packet payload synchronization*
demultiplexing	TS	PID PSI
packet sequence integrity	TS	continuity_counter PID (transport_error_indicator)
payload synchronization	TS	payload_unit_start_indicator adaptation_field_control adaptation_field_length
	PES	random_access_indicator data_alignment_indicator *elementary stream start code detection*
decoder timebase recovery	TS	PCR PID PSI discontinuity_indicator
presentation synchronization	PES	PTS_DTS_flags PTS DTS
data transfer	TS	data_byte adaptation_field_length stuffing_byte
	PES	PES_packet_data_byte PES_packet_length stuffing_byte PES_header_data_length

a) Transport stream PSI functions

A. Find the elements of a program (any of them listed in the PAT) and "identify which streams you want"
and how fast you can do this.

Stages of this are

1. Acquire PAT (payload_unit_start_ind & pointer_section start)
2. Note the PMT of the relevant program
3. Acquire TPs of the PMT_PID
4. Ignore packets with sections with table_id<>02
5. Identify the section of the required program through reading the program # with the current_next_indicator = 1.
6. Read the PCR_PID to identify location of relevant PCRs
7. Less interesting to read the program info descriptors - but use length to skip if not read
8. Use stream_type to identify the type of a stream in order to route the stream to correct decoder (i.e. video to video decoder etc
9. Use elementary_PID to identify next stream_type field location but the descriptors should be read in order to respond to at least the following: Descriptors can be used to identify the wanted streams

B. Check whether the data is correct by reading the CRC_32... and checking you get a zero output of the registers of Annex B model.

C. CA table - If the decoder has the ability to descramble encrypted data it is interesting to receive PID 1, right from the start.

D. It is always interrupting

1. Look for PID 0 packets for any changes

2. Look for TPs of your PMT_PID to see if the PM-section relevant to your program has any changes

3. The same is true for PID 1 if relevant

Descriptors - the ability to make available associated with their stream_PID and for program_number

E. Speed of processing delivery of sections(implications on mechanism) Rate of PSI / Private section processing relative to Transport Stream rate.

F. How quickly tables can be compiled - how "close" sections can be together

G. Reaction time on new data - but this is also considered by timing group - PID switching

H. Are private sections extracted

I. A count of the number of PID filters required

J. Are "next" data required to be stored and when

K. Field filtering - maximum number specified/selected

L. decoder capacity for storing PSI data.

a) Transport Stream additional functions

Table 2 -- Transport Stream additional functions

function	layer	semantic elements
payload scrambling	TS	transport_scrambling_control PSI
	PES	PES_scrambling_control PSI
private data transfer	TS	transport_private_data_flag transport_private_data_length private_data_byte
	PES	PES_extension_flag PES_private_data_flag PES_private_data
trick mode	PES	DSM_trick_mode_flag trick_mode_control field_id intra_slice_refresh frequency_truncation field_rep_cntrl
network maintenance	PES	PES_CRC_flag previous_PES_packet_CRC
error resilience	TS	continuity_counter transport_error_indicator discontinuity_counter

1. Program Streams (fix)

program mux_rate
rate_bound
P-STD_buffer_size
P-STD_buffer_size_bound
delay caused by system target decoder input buffering
difference between two SCRs in successive packs
length of a pack
length of a packet
number of packets in a pack
presence of time stamps in packet headers (DTS, PTS)
CSPS_flag
use of private streams
packet rate
fixed or variable bitrate operation (fixed_flag parameter)
number of multiplexed audio streams (audio_bound parameter)
number of multiplexed video streams (video_bound parameter)
locking of audio sampling frequency and frequency of system clock (system_audio_lock_flag parameter)
locking of video picture rate and frequency of system clock (system_video_lock_flag parameter)

1. Program Stream

a) Program Stream basic functions

Table 3 -- Program Stream basic functions

function	semantic elements
synchronization	pack_start_code packet_start_code_prefix + stream_id PES_packet_length pack_stuffing_length stuffing_byte system_header_start_code header_length
demultiplexing	stream_id
packet sequence integrity	program_packet_sequence_counter_flag PES_extension_flag program_packet_sequence_counter
payload synchronisation	data_alignment_indicator PSM
timebase recovery	SCR program_mux_rate
presentation synchronization	PTS_DTS_flags PTS DTS
data transfer	PES_packet_length PES_packet_data_byte stuffing_byte PES_header_data_length
buffer management	P-STD_buffer_scale P-STD_buffer_size PES_extension_flag P-STD_buffer_flag P-STD_buffer_bound_scale P-STD_buffer_size_bound

a) Program Stream additional functions

Table 4 -- Program Stream additional functions

function	semantic elements
payload descrambling	PES_scrambling_control PSM
private data transfer	PES_extension_flag PES_private_data_flag PES_private_data
trick mode	DSM_trick_mode_flag trick_mode_control field_id intra_slice_refresh frequency_truncation field_rep_cntrl
network maintenance	PES_CRC_flag previous_PES_packet_CRC
padding	stream_id PES_packet_length padding_byte

A. Video bitstreams

A requirement for MPEG video encoders is that the arithmetic precision in the decoder process used in the encoder to produce the coded bitstream shall have the full accuracy specified in ISO/IEC 13818-2.

1. Parameter restrictions

Video encoders may apply restrictions to the following parameters of video bitstreams (see ISO/IEC 13818-2):

a) horizontal_size
b) vertical_size
c) pel_aspect_ratio
d) picture_rate
e) bit_rate
f) VBV_buffer_size
g) profiles and levels
h) f_code[][]
i) progressive_sequence
j) total number of macroblocks.
k) The number of macroblocks per second.
l) picture_structure
m) use of a non-default quantizer matrix for intra coded blocks
n) use of a non-default quantizer matrix for non-intra coded blocks
o) slice structure, that is the definition of where slices start and end within the picture.
p) IBPD structure. That is the picture coding types and sequences of different picture coding types, such as for example the number of consecutive B frames, may be restricted.

q) fixed and/or variable bitrate operation (encoded in the bit_rate field and in the vbv_delay field)

r) the occurrence and specification of user_data

1. Video compliance bitstreams

The following bitstreams test various Profiles of MPEG-2 decoders. The first set of tables summarizes the bitstream list. A second set describes each stream in more detail.

Main Profile

No	Bitstream description	Organization
1	**B-Field MC in Frame Pictures**	GI
2	Sequence with first **B pictures** coded with maximum **vbv_buffer_size**.	[done]
3	**Dual Prime MC, Frame pictures**	Nokia
4	All **MB_type transitions**	[done]
5	**Frequent slices** (e.g. Slice = MB)	TCELA
6	Alternating top/bottom, ZZ/Alternative, Intra/Non-intra VLC tables in I pictures, variable M/N, Field and Frame pictures	IBM
7	**Max. bit burst** and simultaneous Dual Prime MC	IBM
8	**All possible VLCs** symbols and **IDCT mismatch**. **Mismatch** and **saturation**.	Teracom
9	[merged with Paris No. 8]	[done]
10	**MCP** decoder test with chroma (MPEG 94/297)	CCETT
11	**Flat** distribution of VLC events on B and P pictures	LEP
12	**Burst** distribution with simultaneous Bi-directional MC	HHI
13	**Dual Prime MC, field pictures**	TCELA
14	**16x8 Bi-MC, Field pictures**, M > 3	[done]
15	Stream with R/P bits worth of **extra_bit_slice in picture**	Hitachi
16	**MPEG-1** bitstream	Matsushita
17	**Concatenated sequences** of different picture sizes with VBV continuity.	Hitachi.
18	**Low delay** sequence with **skipped pictures**	Norwegian Telecom
19	**IEEE 1180** bitstream test with mismatch	IBM

SNR@ML

Bitrates : 5 Mbit/s base layer and 10 Mbit/s enhancement layer

20	**Maximum VLD bandwidth on both layers** (base and enhancement) with burst of escape codes, bursts of short VLCs and maximum buffer size on both layers	TCEH
21	**Skipped MBs** on base layer, on enhancement layer, and on both layers together. Test of the DCT type in the enhancement layer while Mbs are skipped in the base layer.	TI Japan
22	**Different weighting matrices, different scanning and different q_scale_types** on the two layers	HHI

SSP@H-14L

23	All MB transitions in enhancement layer, all posible VLC symbols in enhancement layer, and all cases of motion vector updating	HHI, TCEH

Conformance bitstream characteristics:

Where applicable, the motion vectors in the conformance bitstreams reconstruct to values such that half-pel interpolation is performed in the Motion Compensated Predictor stage of the decoder for all components (Y,Cb,Cr).

For the Main Level conformance bitstreams, the bitstream parameters are at the legal maximum:

sampling rate: 720x576x25 or 720x480x30 fps
bitrate: 15 Mbit/sec
vbv_buffer_size: 1.75 Mbit
f_code[][1] = 5

All bistreams observe the rule that only 2 macroblocks can exceed the legal 3:2 expansion factor specified in ISO/IEC 13818-4.

Bitstreams

Bitstream	1
Title	B-field MC in frame pictures
Profile and Level	MP@ML
Description	Frame coded picture, B-directional macroblock predictions in field mode.Reconstructed motion vectors have half-pel coordinates for Y,Cb,Cr blocks
Functional stage	prediction bandwidth
Purpose	field MB predictions in frame pictures have the largest prediction bandwidth overhead. Picture buffers organized as frames (interleaved fields) and macroblocks stored in contiguous address page segments would have the greatest penalty. Effective filtered block size is 16x8.
Parameters	4 and 60 frames
Organization	GI

Bitstream	2 (already generated)
Title	Sequence containing B-picture coded with maximum vbv_buffer_size (1.75 Mbit)
Profile and Level	MP@ML
Description	Maximum number of coded bits for the first B picture. Encourage use of long VLCs (not via escapes)
Functional stage	VLD
Purpose	A large B-picture located after several smaller coded pictures can catch a decoder off guard.
Parameters	4 and 60 frames, 720x576x25 fps
Organization	TCE-Hannover

Bitstream	3
Title	Dual Prime MC, frame pictures
Profile and Level	MP@ML
Description	All dual-prime macroblocks in frame-coded picture. All half-pel motion vectors
Functional stage	prediction bandwidth
Purpose	prediction bandwidth is at a maximum in this mode due to the small block sizes and two prediction sources.
Parameters	4 and 60 frames
Organization	Nokia

Bitstream	4.
Title	All MB_type transitions
Profile and Level	MP@ML
Description	All MB_type transitions (frame coding only
Functional stage	parser
Purpose	Sloppy decoders may not handle all scenarios in parsing tree.
Parameters	
Organization	Tektronix

Bitstream	5
Title	Frequent slices (average of 11 MB/slice over picture period)
Profile and Level	MP@ML
Description	Frequent slices (e.g. Slice = MB)
Functional stage	VLD and parser
Purpose	CPU-oriented designs have large overhead for each MB header.
Parameters	60 frames
Organization	TCE-LA

Bitstream	6.
Title	Alternating top/bottom, ZZ/Alternative, Intra/Non-intra VLC tables in I pictures, Field and Frame pictures, variable M/N, downloaded quantization weighting matrices.
Profile and Level	MP
Description	Various syntax switches are toggled from frame-to-frame
Functional stage	parser and control
Purpose	Sloppy decoders may not handle all scenarios
Parameters	This is a purely static test so it may not be necessary to specifiy bitstream sampling parameters.
Organization	IBM

Bitstream	7
Title	Max. bit burst and simultaneous Dual Prime MC
Profile and Level	MP@ML
Description	Simultaneous burst of coded bits and maximum bandwidth Dual Prime MC, followed by remaining macroblocks outside the burst with Dual Prime MC. Half-pel
Functional stage	VLD and prediction bandwidth
Purpose	DRAM is shared by VLD, MCP, and Display functions. This combination presents the longest sustainable period (whole picture) for DRAM bandwidth
Parameters	4 and 60 frames
Organization	IBM

Bitstream	8
Title	All possible VLCs symbols and IDCT mismatch. Mismatch and saturation.
Profile and Level	MP@ML
Description	All possible VLCs symbols and IDCT mismatch. Mismatch and saturation.
Functional stage	parser ; IDCT accuracy
Purpose	Sloppy decoders may not have included the complete VLC tables and may not implement sufficient drift correction
Parameters	IDCT subwindow 32x32
Organization	Swedish Telecom, TRAB [merged the old No. 8 and 9]

Bitstream	10
Title	MCP decoder test (MPEG 94/297)
Profile and Level	MP@ML
Description	Modification of MCP decoder test (MPEG 93/953) which adds chrominance prdiction and motion vectors that point towards the edge of pictures
Functional stage	MCP
Purpose	Test motion compensation stages for accuracy and that they execute all stages.
Parameters	[this is mostly a static test]
Organization	CCETT

Bitstream	11
Title	Flat distribution of VLC events on B and P pictures
Profile and Level	MP@ML
Description	Flat distribution of VLC events (worst case for constant rate symbolic VLDs) on B and P pictures.
Functional stage	VLD
Purpose	Badly designed systems will rely on statistically low count of symbols over global areas to meet real-time constraints.
Parameters	4 and 60
Organization	LEP

Bitstream	12
Title	Burst distribution of coded bits with simultaneous Bi-dir. MC
Profile and Level	MP@ML
Description	Bursty case for number of bits per MB with different burst location within picture (top, bottom), followed Bi-directional macroblocks. All motion vectors with half pel components.Macroblocks outside the burst concentration have all bidrectional prediction
Functional stage	VLD and prediction bandwidth
Purpose	Poor constant-bitrate designs will rely upon statistically small number of coded bits over local areas.
Parameters	4 and 60 frames
Organization	HHI

Bitstream	13
Title	Dual Prime MC in Field Pictures
Profile and Level	MP@ML
Description	Field-coded pictures, all Dual Prime. Half pel
Functional stage	prediction bandwidth
Purpose	some designs may behave differently on fields than on frames
Parameters	4 and 60 frames
Organization	TCE-LA

Bitstream	14
Title	16x8 bidirectional MC in Field Pictures, M > 3
Profile and Level	MP@ML
Description	Field coded pictures with 16x8 bidirectional macroblock motion compensation. Sequence contains many consequitive B pictures.
Functional stage	prediction bandwidth
Purpose	represents worst case bandwidth
Parameters	4 and 60 frames
Organization	TCE-LA

Bitstream	15.
Title	R/P bits for extra bit slice
Profile and Level	MP@ML
Description	Stream with R/P bits worth of extra_bit_slice in picture
Functional stage	Parser
Purpose	Test whether a decoder is capable of handling a large number of bits concentrated in the extra bit slice loop.
Parameters	4 and 60 frames.
Organization	Hitachi

Bitstream	16
Title	MPEG-1 stream
Profile and Level	MP@ML
Description	MPEG-1 bistream
Functional stage	overall
Purpose	Test whether MPEG-2 decoders observe requirement that they should decode MPEG-1 bitstreams coded at the same level parameters.
Parameters	4 and 60 pictures.
Organization	Matsushita

Bitstream	17
Title	Concatenated sequences of different picture sizes with VBV continuity.
Profile and Level	MP@ML of differing sampling dimensions and bitrates for each of the N concatenated sequences.
Description	Concatenated sequences of different picture sizes with VBV continuity
Functional stage	controller
Purpose	Realistic decoders will encounter concatenated sequences where each sequenece will possess different bitstream parameters (frame dimensions and rates).
Parameters	2*4 and 2*60 frames
Organization	Hitachi

Bitstream	18
Title	Low delay sequence with skipped pictures.
Profile and Level	SP@ML
Description	Low delay sequence with skipped pictures.
Functional stage	controller
Purpose	MPEG-2 Main Profile video decoders should be capable of decoding low delay mode and know how to recognize and deal with skipped pictures and buffer underflows in the VBV model.
Parameters	60 frames (4 may be insufficient to have a skipped picture).
Organization	Norwegian Telecom Research

Bitstream	19
Title	IEEE 1180 bitstream test with mismatch
Profile and Level	MP@ML
Description	IEEE 1180 IDCT mismatch test mapped into a bitstream to test the decoder's IDCT statistical accuracy.
Functional stage	IDCT
Purpose	Test IDCT decoder accuracy. This is not a drift test since all macroblocks are of type Intra.
Parameters	whatever number of frames are required to satisfy statistic count
Organization	IBM

Bitstream	20
Title	Maximum VLD bandwidth on both layers (base and enhancement) with burst of escape codes, bursts of short VLCs and maximum buffer size on both layers
Profile and Level	SNRP@ML
Description	Test of the maximum VLD bandwidth on both layers (base and enhancement) with burst of escape codes, bursts of short VLCs and maximum buffer size on both layers
Functional stage	test of parser(s)
Purpose	some designs may not be able to handle both layers
Parameters	5 Mbit/s base layer and 10 Mbit/s enhancement layer
Organization	TCE-Hannover

Bitstream	21
Title	Skipped MBs on base layer, on enhancement layer, and on both layers together. Test of the DCT type in the enhancement layer while Mbs are skipped in the base layer.
Profile and Level	SNRP@ML
Description	Test of skipped Mbs on base layer, on enhancement layer, and on both layers together. Test of the DCT type in the enhancement layer while Mbs are skipped in the base layer.
Functional stage	test of parser
Purpose	sloppy decoders may not be able to handle skipped Mbs in one of the layers
Parameters	5 Mbit/s base layer and 10 Mbit/s enhancement layer
Organization	TI Japan

Bitstream	22
Title	Different weighting matrices, different scanning on the two layers
Profile and Level	SNRP@ML
Description	Test of different weighting matrices, different scanning on the two layers
Functional stage	test of decoder
Purpose	sloppy decoders may not be able to handel different weighting matrices or scannings
Parameters	5 Mbit/s base layer and 10 Mbit/s enhancement layer
Organization	HHI

Bitstream	23
Title	All MB transitions in enhancement layer, all posible VLC symbols in enhancement layer, and all cases of motion vector updating
Profile and Level	SSP@H-14L
Description	Test of all MB transitions in enhancement layer, all posible VLC symbols in enhancement layer, and all cases of motion vector updating
Functional stage	test of spatially scalable decoder
Purpose	sloppy decoders may not be able to handle all possible cases
Parameters	*to be defined*
Organization	HHI and TCE-Hannover

A. Audio bitstreams

Audio encoders may apply restrictions to the following parameters of audio bitstreams (see ISO/IEC 13818-3):

2.3.3.1 Extension of ISO/IEC 11172-3 Audio Coding to Lower Sampling Frequencies

a) layer
b) bitrate_index
c) sampling_frequency
d) mode
e) mode_extension
f) emphasis
g) generation of crc_check
h) value of fixed bitrate when coding in free format mode.
i) generation of ancillary data

2.3.3.2 Low bit rate coding of Multichannel Audio

a) layer
b) bitrate_index

c) sampling_frequency
d) mode in MPEG-1 header, and center, surround and LFE in MC-header
e) mode_extension in MPEG-1 header
f) emphasis
g) generation of crc_check
h) value of fixed bitrate when coding in free format mode.
i) generation of ancillary data
j) use of extension stream
k) audio_mix
l) dematrix_procedure
m) no_of_multi_lingual_ch
n) multi_lingual_fs
o) multi_lingual_layer
p) n_ad_bytes
q) tc_sbgr_select
r) dyn_cross_on
s) mc_prediction_on

The use of higher order prediction (more than zero-order), and/or delay compensation will limit the editability of the coded bit streams.

I. Decoder characteristics

The characteristics of a decoder specify the properties and capabilities of the decoding process applied in the decoder. An example of a property is the arithmetic accuracy that is applied. The capabilities of a decoder specify which coded bitstreams the decoder can reconstruct, by defining the subset of the standard that may be exploited in decodable bitstreams. A bitstream can be decoded by a decoder if the characteristics of the coded bitstream are within the subset of the standard defined by the decoder capabilities. Compliance to ISO/IEC 11172 by a decoder requires that the capabilities of the decoder are specified. That is, each constraint on the subset of the standard that may be exploited in bitstreams decodable by the decoder shall be specified.

A. System decoders

An ITU-T Rec. H.222.0 | ISO/IEC 13818-1 decoder may support specific values only, or a specific range of the following parameters in system bitstreams. These parameters are encoded directly or indirectly in the bitstream for program streams.

a) mux_rate
b) P-STD_buffer_size
c) packet rate

Furthermore, a decoder may constrain the support of fixed and/or variable bitrate operation (see definition of the fixed_flag field in xxxx of ITU-T Rec. H.222.0 | ISO/IEC 13818-1), and may require locking between the 27 MHz system clock and the audio sampling frequency and/or the video picture rate (see the system_audio_lock_flag and the system_video_lock_flag fields in xxxx of ITU-T Rec. H.222.0 | ISO/IEC 13818-1). Decoders shall specify how private streams are handled.

1. Transport Streams

1. Program Streams

A. Video decoders

An ISO/IEC 13818-2 video decoder may support specific values only, or a specific range, or a specific combination of values or ranges of the following parameters in video bitstreams. These parameters are encoded directly or indirectly in the bitstream.

a) horizontal_size
b) vertical_size
c) pel_aspect_ratio
d) picture_rate
e) bit_rate
f) VBV_buffer_size
g) The total number of macroblocks in a picture.
h) The total number of macroblocks per second.
i) range of motion vectors (encoded in the various fields related to motion vectors)

Furthermore, an ISO/IEC 13818-2 video decoder may constrain:

a) IBPD structure. That is the picture coding types and sequences of different picture coding types that are supported, such as for example the number of consecutive B frames, may be restricted.
b) the support of fixed and/or variable bitrate operation.

An ISO/IEC 13818-2 video decoder shall specify how sequence_error_codes and user data are handled.

For a video decoder to be compliant to ISO/IEC 13818-2 a statement of whether or not computation is carried out with the full accuracy specified in ISO/IEC 13818-2 is required. If the full arithmetic precision is not implemented, the accuracy shall be specified. In the case of the inverse DCT, compliance testing requires performing the tests described in IEEE standard P1180/D2, as indicated in annex A of ISO/IEC 13818-2, and a statement of the numerical results for peak error and mean square error.

The legal ranges of the above parameters are defined in Chapter 8 of ISO/IEC 13818-2.

A. Audio decoders

1. Extension of ISO/IEC 11172-3 Audio Coding to Lower Sampling Frequencies

An ISO/IEC 13818-3 Low Sampling Frequency audio decoder may support only specific values, or a specific range, or a specific combination of values or ranges of the following parameters in audio bitstreams. These parameters are encoded directly or indirectly in the bitstream.

a) layer
b) bitrate_index
c) sampling_frequency
d) mode
e) mode_extension
f) emphasis

Furthermore, an ISO/IEC 13818-3 Low Sampling Frequency audio decoder may constrain the support of free format mode. For an ISO/IEC 13818-3 Low Sampling Frequency audio decoder the handling of ancillary data and error protection (crc_check) shall be specified, as shall the single channel performance (single channel output at one or at both output channels).

Compliance of an audio decoder to ISO/IEC 13818-3 Low Sampling Frequency requires that the output signal of the decoder is reconstructed accurately. For actual tests see 2.6.3.

An ISO/IEC 13818-3 Low Sampling Frequency compliant audio decoder that is able to support at least one but not all combinations of the options defined in 2.4.3.1 such as bit rates, sampling rates and modes, will be designated as an ISO/IEC 13818-3 Low Sampling Frequency Layer N audio decoder. Decoders that support all combinations are designated as Full ISO/IEC 13818-3 Low Sampling Frequency Layer N audio decoders, where N indicates I, II or III.

1. Low bit rate coding of Multichannel Audio

An ISO/IEC 13818-3 Multichannel audio decoder may support only specific values, or a specific range, or a specific combination of values or ranges of the following parameters in audio bitstreams. These parameters are encoded directly or indirectly in the bitstream.

a) layer
b) bitrate_index
c) sampling_frequency
d) mode in MPEG-1 header, and center, surround and LFE in MC-header
e) mode_extension in MPEG-1 header
f) emphasis
g) generation of ancillary data
h) use of extension stream
i) dematrix_procedure
j) no_of_multi_lingual_ch
k) multi_lingual_fs
l) multi_lingual_layer
m) tc_sbgr_select
n) dyn_cross_on
o) mc_prediction_on

Furthermore, an ISO/IEC 13818-3 Multichannel audio decoder may constrain the support of free format mode. For an ISO/IEC 13818-3 Multichannel audio decoder the handling of ancillary data and error protection (crc_check) shall be specified, as shall the single channel performance (single channel output at one or at both output channels).

Compliance of an audio decoder to ISO/IEC 13818-3 Multichannel Audio requires that the output signal of the decoder is reconstructed accurately. For actual tests see 2.6.3.

An ISO/IEC 13818-3 Multichannel compliant audio decoder that is able to support at least one but not all combinations of the options defined in 2.4.3.2 such as bit rates, sampling rates and modes, will be designated as an ISO/IEC 13818-3 Multichannel Layer N audio decoder.

An ISO/IEC 13818-3 Multichannel compliant audio decoder that is able to support at least the features according to the following table will be designated as a Core ISO/IEC 13818-3 Multichannel Layer N decoder or as a Full ISO/IEC 13818-3 Multichannel Layer N decoder respectively.

	Layer I		Layer II		Layer III	
bitstream characteristic	Core	Full	Core	Full	Core	Full
MPEG-1 Layer I	Y	Y	N	Y	N	Y
MPEG-1 Layer II	N	N	Y	Y	N	Y
MPEG-1 Layer III	N	N	N	N	Y	Y
Configuration 3/2	Y	Y	Y	Y	Y	Y
Configurations 3/1, 3/0, 2/2, 2/1	N	Y	N	Y	N	Y
Configuration 2/0, 1/0	Y	Y	Y	Y	Y	Y
2nd stereo	N	Y	N	Y	N	Y

LFE	N	Y	N	Y	N	Y
free format	N	Y	N	Y	N	Y
extension bitstream	N	Y	N	Y	N	Y
channel switching	Y	Y	Y	Y	Y	Y
dynamic cross-talk	Y	Y	Y	Y		
prediction zero order	Y	Y	Y	Y	Y	Y
prediction higher order, delay	N	Y	N	Y		Y
dematrix procedure ’2Õ	N	Y	N	Y	N	Y
multilingual Layer II	N	Y	N	Y	N	Y
multilingual Layer II half Fs	N	Y	N	Y	N	Y
multilingual Layer III	N	N	N	N	N	Y
multilingual Layer III half Fs	N	N	N	N	N	Y

I. Procedures to test bitstream compliance

This clause describes procedures to verify compliance of system, video and audio decoders. All tests are performed using error free bit streams. For correct interpretation of syntax and semantics, test sequences covering a wide range of parameters shall be supplied to the decoder under test and its output sequence shall be compared with the output of a reference decoder. The comparison can be done, for example, by performing subjective tests, by evaluation of the difference signal and by comparing the timing performance. Each compliant decoder shall be able to decode all compliant ISO/IEC 11172 streams within the subset of the standard defined by the specified capabilities of the decoder. The procedures to test decoder compliance are given for parts 1, 2 and 3 of ISO/IEC 11172 in separate sub-clauses.

A. System bitstream tests

1. Introduction

The following tables identify functions of the Transport Stream and the Program Stream at the decoder. The tables identify syntax elements required to support the given function.
Each function specifies a decoder capability. A decoder that does not implement a specific function, performs no action when the associated fields are received. The presence of fields in the bitstream, relating to a decoder function that is not implemented, do not cause the decoder to fail.
For both the Transport Stream and the Program Stream a set of basic functions, and a set of additional functions, are given. A Transport Stream decoder must implement all of the Transport Stream basic functions. A Program Stream decoder must implement all of the Program Stream basic functions. In each case it is a decoders option as to whether the additional functions, respectively, are implemented.

1. System bitstream tests for Transport Streams

a) Tests at the transport_packet() level

sync_byte: each TS packet shall contain the value 0x47 as the first byte of the packet.
transport_error_indicator: if the transport_error_indicator is set to '1', all other test of fields within the packet may not yield valid results.
payload_unit_start_indicator: if the payload_unit_start_indicator is set to '1', and if the TS packet contains the start of a PES packet, the first five bytes of the TS packet payload shall contain the PES start code, which consists of packet_start_code_prefix followed by stream_id. The latter contains a value in the range permitted by Table 2-18 in ITU_T Rec. H.222.0 | ISO/IEC 13818-1. Otherwise if the payload_unit_start_indicator is set to '1', and if the TS packet contains PSI data, the first byte of data_byte shall be the first byte of a PSI section.

transport_priority: no tests are applied to transport_priority.
PID: the PID value shall be one of 0x0000, 0x0001, or 0x0010, through to 0x1FFF. Other values represent an error. If the PID value is 0x1FFF the payload_unit_start_indicator shall be set to '0', and the transport_scrambling_control field shall be '00'.
transport_scrambling_control: no tests are applied to transport_scrambling_control.
adaptation_field_control: if the adaptation_field_control is set to '00', then there is an error. If the adaptation_field_control is set to '01' there is no adaptation field in this TS packet. If the adaptation_field_control is set to '10', then the adaptation_field_length field shall have a value of 183. If the adaptation_field_control is set to '11', then the adaptation_field_length field shall have a value in the range of 0 to 183.
continuity_counter: if the continuity_counter field value is equal to the continuity_counter field value of the previous TS packet, then either the adaptation_field_control has a value of '00' or '10', the discontinuity_indicator field is set to '0' and the entire TS packet is a replica of the previous TS packet of the same PID, except for the PCR field, the discontinuity_indicator field is set to '1', or there is an error. If the continuity_counter field value is not equal to, or is not one greater (modulo 16) than the continuity_counter field value of the previous TS packet, then the either the discontinuity_indicator is set to '1', or there is an error.
data_byte: tests of packet payload are performed consistent with the stream_type.

a) Tests at the adaptation_field() level

adaptation_field_length: the adaptation_field_length shall be consistent with the adaptation_field_control field.
discontinuity_indicator: the discontinuity_indicator field shall be set to '1' if there is a PCR discontinuity. If the discontinuity_indicator is set to '1' the discontinuity state is true and remains true until the next packet of the same PID with the discontinuity_indicator equal to '0'. For each discontinuity state (consecutive packets of the same PID fo which the continuity state is true) the continuity_counter may be discontinous once. More than one continuity counter discontinuity constiutes an error. For each discontinuity state there may be onr PCR discontinuity. More than one PCR during a discontinuity state constitutes an error. If the PID is not a PCR_PID no more than two consecutive packets of the same PID shall have the discontinuity_indicator set to '1'. More than two such packets constiutes an error. In addition, the first byte after a continuity_counter should be the first byte of an elementary stream access point, *it* may be discontinous in version numbers of PSI sections, with section_length equal to 13, and *old versus new PTS and DTS values*.
random_access_indicator: if the TS packet contains PES data and the random_access_indicator is set to '1' and the payload_unit_start_indicator is set to '1', then the PES packet shall start in the transport packet. If the TS packet contains PES data and the random_access_indicator is set to '1', and the PID is PCR_PID then there shall be a PCR.
elementary_stream_priority_indicator: no tests are performed.
PCR_flag: if the PCR_flag is set to '1' then there shall be a PCR field present in this TS packet.
OPCR_flag: if the OPCR_flag is set to '1' then there shall be an OPCR field present in this TS packet.
splicing_point_flag: if the splicing_point_flag is set to '1' then there shall be a splice_countdown field present in this TS packet.
transport_private_data_flag: if the transport_private_data_flag is set to '1' then there shall be one or more private_data_byte fields present in this TS packet.
adaptation_field_extension_flag: if the adaptation_field_extension_flag is set to '1' then there shall be an adaptation_field_extension field present in this TS packet.
PCR:
OPCR: no test are perfomed on this field.
splice_countdown: this splice_countdown field is checked only in packets containing payload. Duplicate packets are not checked. If splice_countdown then for all consecutive packets with splicing_point_flag equal to '1', the splice_countdown field shall decrement with each packet. (2's complement numbers roll over from 1000 0000 (-128) to 0111 111 (+127)). If the splice_countdown is set to '1' then the packet payload starts with the first byte of a PES packet which starts with the first byte of an access point.
transport_private_data_length:
private_data_byte:
adaptation_field_extension_length:
ltw_flag:
piecewise_rate_flag:
seamless_splice_flag:

ltw_valid_flag:
ltw_offset:
piecewise_rate:
splice_type:
DTS_next_au:
stuffing_byte:

a) Tests at the PES header level

List of tests to be performed in bitstream verification,
concerning the syntax and semantics of the PES packet

Basic Assumptions

For the parsing of the PES packet, it is assumed that
a. The arrival time of each byte in the T- or P-STD is known,
b. It is known where in the bitstream a PES packet begins,
c. It is known whether the PES packet resides in a PS or a TS
d. If in a TS, it is known in what PID the PES packet comes from.
e. It is possible to pass information to the outside world (i.e. to the part of the software that creates log files)
Notation
N is the number of bytes that are read in following the PES_header_data_length field.
N should be properly incremented when data is read in.

Tests

packet_start_code_prefix
- Test that the value is 0x000001.

stream_id
- Test that the value of stream_id is not between 11110100-11111110
- If the PES packet resides in a TS, test that the value is not 10111100 or 11111111
- If the PES packet resides in a TS, test that the value is not different than other stream_ids of the same PID.
(Exception: padding?)
- Output the type of stream.

pes_packet_length
- In a TS, verify that the value is not zero in any PES packet that does not contain video.
- In a PS, verify that the value is not zero.

NOTE: If the stream_id is one of program_stream_map, padding_stream, private_stream_2, ECM, EMM,
program_stream_directory, the parsing stops here.

NOTE: There is a BUG in the PES header syntax!!! DSM-CC should be added to the list above. Perhaps, NOTE 5 is
wrong also/instead.

The next two bits should be 10

PES_scrambling_control
- Send to log file

PES_priority
- Send to log file

data_alignment _indicator
- if the stream_id indicates anything else than an MPEG video stream, check that this bit is not set to 1 (??).
- if the value is 1, and the stream_id indicates an MPEG video stream, check that the first bytes of the payload of
this PES packet correspond to a slice, picture, GOP or sequence start code. Note that this can only be performed if
the payload is NOT scrambled. (Comment: should there be interaction with PSI at this point to read the data
alignment descr.?)

copyright
- output to log file
(Comment: should there be interaction with PSI at this point to check the existence of the copyright descr.? If
yes, check that it is not present if this bit is set to zero)

original_or_copy
- output to log file

PTS_DTS_flags
- check that the value is not 01.

ESCR_flag
- no action

ES_rate_flag
- no action

DSM_trick_mode_flag
- no action

additional_copy_info_flag
- no action

PES_CRC_flag
- no action

PES_extension_flag
- no action

PES_header_data_length
- no action

```
if (PTS_DTS_flags == 10) {
        read PTS  field and report results to log file;
        verify that the previous PTS was received within the last 700 msec and report results to
        log file;
        verify that a picture_start_code or audio_start_code is present in this PES packet and
        report results to log file; (NOTE: Will depend on the resolution of the ongoing debate on
        the definition of DTS/PTS);
        Check for correctness of PTS, in particular, the following:
                * Should refer to a time in the future.
                * Should refer to a time later than the arrival time of the last byte of the access
                  unit (except in the case of skipped pictures when low_delay == 1).
                * If stream_type == video && low_delay == 0, verify that picture is B.
```

```
                    * For B pictures in trick mode, check that the B picture is not removed from the
                      buffer until after the last presentation time of this picture as defined by the
                      field-control_id.
}

if (PTS_DTS_flags == 11) {
          read PTS and DTS fields and report results to log file;
          verify that the previous PTS was received within the last 700 msec and report results to
          log file;
          verify that a picture_start_code or audio_start_code is present in this PES packet and
          report results to log file; (NOTE: Will depend on the resolution of the ongoing debate on
          the definition of DTS/PTS);
          Check for correctness of PTS and DTS, in particular, the following:
                    * ES type is not audio.
                    * If Video, low delay flag is not equal to 1.
                    * Both should refer to a time in the future.
                    * Both should refer to a time later than the arrival time of the last byte of the
                      access unit.
                    * PTS not equal to DTS.
}


if (ESCR_flag == 1) {
          read ESCR base and ESCR_extension and report results to log file;
          verify that the previous ESCR received was within the last 700 msec and report results to log
file;

}

if (ES_rate_flag == 1)
          read ES_rate and report results to log file;

If (DSM_trick_mode_flag == 1) {
          /* read the DSM_trick_mode_field */
          read trick_mode_control;
          output table 2-20 description to log file;
          based on trick_mode_control value read and output to log file the relevant fields
          (from the following list - field_id, intra_slice_refresh, frequency_truncation,
          field_rep_control)
          }

if (additional_copy_info_flag == 1) {
          read and output to log file additional_copy_info;
          }

if (PES_CRC_flag == 1) {
          read previous_pes_packet_crc and output to log file;
          calculate CRC for previous PES packet;
          compare to CRC read above and report results to log file;
          }

if (pes_extension_flag == 1) {
          read pes_private_data_flag and report results to log file;
          read pack_header_field_flag and report results to log file;
          read program_packet_sequence_counter_flag and and report results to log file;
          read P-STD_buffer_flag and report results to log file;
          read PES_extension_flag_2 and report results to log file;
          if (PES_private_data_flag == 1)
                    read pes_private_data and report results to log file;
```

```
        if (pack_header_field_flag == 1) {
                read pack_field_length and report results to log file;
                read pack_header and report results to log file;
                if possible,check for correctness of pack_field_length and report results to log file;
        }
        if (program_packet_sequence_counter_flag == 1) {
                read program_packet_sequence_counter and report results to log file;
                check for correct incrementing of this field and report results to log file;
                read MPEG1_MPEG2 identifiers and report results to log file;
                read original_stuff_length and report results to log file;
        }
        if (P-STD_buffer_flag == 1) {
                read P-STD_buffer_scale and report results to log file;
                read P-STD_buffer_size and report results to log file;
        }
        if (PES_extension_flag_2 == 1) {
                read PES_extension_field_length and report results to log file;
                read next PES_extension_field_length number of bytes and report results to log file;
                if possible, check for correctness of PES_extension_field_length;
        }
}

read PES_header_length - N bytes of stuffing and report results to log file;
read PES_packet_data_bytes till end of PES packet and report results to log file;
```

Uncertainties:

It is allowed to have more than one stream_id in one PID (perhaps for padding streams?)?
Is the padding stream allowed in the TS at all?
Is it an error to set the data_alignement_indicator to 1 in anything but video?

a) Tests on Program Specific Information (PSI)

Bitstream verification PSI

For the purpose of this text, the term section data refers to all bytes after the last_section_number field and before the CRC_32 field of any section.

PID0

1. Confirm that in the Transport Stream under test (TSUT), there is at least one Transport Stream packet with PID==0x0000 and length 188 bytes.

2. For all Transport Packets of PID==0x0000 check that the trasnport_scrambling_control bits are set to '00' - ie not scrambled.

3. Check that subsequent Transport Packets of PID==0x0000 have continuity_counter values which increase by +1 with each successive Transport Stream packet.

4. If the payload_unit_start_indicator is set to a value of '1' in a Transport Stream packet of PID==0x0000 then the first data byte of the Transport Packet shall be understood to be a pointer_field.

If the payload_unit_start_indicator is set to '0' then no pointer_field shall be present in that Trasnport Stream packet and it is not conformant for a section to start in that Transport Stream packet.

5. If the adaptation_field_control ==00 then the Transport Packet violates the semantics.

If the adaptation_field_control =='01' then check that
0 l.t.o.= pointer_field l.t.o.= 182

If the adaptation_field_control =='10' or '11' then check that
0 l.t.o.= pointer_field l.t.o.= (181 - adaptation_field_length)

(This checks the pointer does not point to beyond the end of the TP)

6. Check that the byte pointed to by the pointer_field takes the value '0x00' ie table_id=0x00

7. Check that the bits following the table_id field are '1011'

8. Check the section_length_field

9 l.t.o.= section_length l.t.o.= 1021

if section_length == n then byte n+1 == 0x00 or 0xFF

The nth byte of the section shall be checked to be the last byte of of the CRC_32.

9. The following 16 bits may take any value. For all table_id==0x00 sections in the TSUT this field should be coded the same value.

Q: What happens if you want to change the label?

10. Check that the following two bits are '11'.

11. The following four fields (version_number, current_next_indicator, section_number, last_section_number) must be analysed together. For this purpose, the following notation is used inthe draft.

A section is denoted with the following attributes:

S(v, c, s,l) where
v is the value of the version_number field
c is the value of the current_next_indicator flag
s is the value of the section_number field of the section
l is the value of the last_section_number field

12. program_number

Check that the program_number field does not take any single value more than once within all sections of any version of the program association table.

Check that if the program_number is 0x0000 then the following three bits shall be coded '111' and the bits after that do not take any value from 0x0000 to 0x000F and also do not take the value 0x1FFF. Check that they also do not take any values identified as Transport Packets carrying video or audio data.

Note that it is permitted for the network_PID to be the same as a program_map_PID

Check that if there are any Transport Stream packets with the PID denoted the network_PID then these packets contain private_sections.

Check that if the program_number field takes any value other than zero then the following three bits are '1'. The 16 bits after that define the program_map_PID for that program. Check that sections with table_id==0x02 and the program_number field equal to a program_number listed within the PAT only occur within Transport Stream packets which have the PID value of the program_map_PID listed for that program in the PAT.

13. PMT PID

Check that the program_map_PID field does not take any value from 0x0000 to 0x000F and also does not take the value 0x1FFF. Check that it also does not take any value identified as Transport Packets carrying video or audio data.

14. Check that the CRC_32 field is correctly caluclated such that the field value gives a zero output of the registers in the decoded defined in Annex B of part 1 of this standard after processing the entire conditional access section.

15. Check that the first byte after the last byte of the CRC_32 field in any Transport Packet with PID==0x0000 takes either the value 0x00 (in which case this is the start of the following section) or the value 0xFF. If the value of 0xFF is found then check all following bytes to the end of the Transport Packet are also 0xFF.

If the value 0xFF was present then the it should be checked that following Transport Packet of PID==0x000o either i) consists only of adaptation_field
or ii) has the payload_unit_start_indicator set to '1', has the pointer_field set to zero and then has either a) table_id==0x00, starting a new section
or b) table_id==0xFF, with 0xFF stuffing to the end of the Transport Packet.

If no section started in the following Transport Packet, then the same test should be carried out untila Transport Packet is found with the table_id==0x00 condition is found.

16. Check that there are no sections with table_id == 00 in any Transport Packet with a PID value other than 0x0000.

17. If there is a change in the program_number fields of the programs present in the TSUT then it shouldbe checked that a new version of the PAT is present in the TSUT prior to the change, giving the new configuration. An example of such a change would be adding a new program, or deleting a program in order to be able to re-use the program_number.

18. Check that all the sections of the PAT for a given version_number contain a complete list of all the programs present within the TSUT. This is done by checking that every section in the TSUT with table_id==0x02 are listed through their program_numbers in the PAT. ie no sections with table_id==0x02 should be found in the TSUT unless their program number is listed in the PAT.

PID 1

1. If Transport Stream packets exist within the TSUT with transport_scrambling_control bits not set to '00' or PES packets with PES_scrambling_control bits not set to '00' then check that there are Transport Stream packets with PID==0x0001.

2. For all Transport Packets of PID==0x0001 check that the trasnport_scrambling_control bits are set to '00' - ie not scrambled.

3. Check that subsequent Transport Packets of PID==0x0001 have continuity_counter values which increase by +1 with each successive Trasnport Stream packet.

4. If the payload_unit_start_indicator is set to a value of '1' in a Transport Stream packet of PID==0x0001 then the first data byte of the Transport Packet shall be understood to be a pointer_field.

If the payload_unit_start_indicator is set to '0' then no pointer_field shall be present in that Trasnport Stream packet and it is not conformant for a section to start in that Transport Stream packet.

5. If the adaptation_field_control ==00 then the Transport Packet violates the semantics.

If the adaptation_field_control =='01' then check that
0 l.t.o.= pointer_field l.t.o.= 182

If the adaptation_field_control =='10' or '11' then check that
0 l.t.o.= pointer_field l.t.o.= (181 - adaptation_field_length)

(This checks the pointer does not point to beyond the end of the TP)

6. Check that the byte pointed to by the pointer_field takes the value '0x01' ie table_id=0x01

7. Check that the bits following the table_id field are '1011'

8. Check the section_length_field

9 l.t.o.= section_length l.t.o.= 1021

if section_length == n then byte n+1 == 0x01 or 0xFF

The nth byte of the section shall be checked to be the last byte of of the CRC_32.

9. Check that the following 18 bits all take the value '1'.

10. The following four fields (version_number, current_next_indicator, section_number, last_section_number) must be analysed together. The same rules apply for this as for program_association_sections above.

11. Following the last_section_number field descriptors may occur. There should be sufficient cases of the CA_descriptor to identify program related access control information for all scrambled streams. Therefore the first byte of any descriptor found should be checked to be 9. Since the content of the descriptor is primarily private, it is not possible to verify whether this private data is 'correct' or not. It is however possible to check whether the section is consistent with the standard, by assuming that the descriptor_length fields are correctly coded, and checking whether this then implies that the last four bytes of the section (determined by the section_length field) are a CRC_32 field correctly calculated over the complete section to ensure that the output of the registers of a decoder defined in Annex B of part 1 of this standard is zero.

Within the CA_descriptor the descriptor_length fields should together indicate the correct location of the CRC_32 within the section. Furthermore, the CA_PID should not be the same as any PID allocated for another purpose (0, 1, video PID, audio PID, PM_PID unless the CA data is coded in sections which do not have the table_id ==0x00, 0x01, 0x02).

12. Check that the CRC_32 field is correctly caluclated such that the field value gives a zero output of the registers in the decoded defined in Annex B of part 1 of this standard after processing the entire conditional access section.

13. Check that the first byte after the last byte of the CRC_32 field in any Transport Packet with PID==0x0001 takes either the value 0x01 (in which case this is the start of the following section) or the value 0xFF. If the value of 0xFF is found then check all following bytes to the end of the Transport Packet are also 0xFF.

If the value 0xFF was present then the it should be checked that following Transport Packet of PID==0x0001 either i) consists only of adaptation_field
or ii) has the payload_unit_start_indicator set to '1', has the pointer_field set to zero and then has either a) table_id==0x01, starting a new section
or b) table_id==0xFF, with 0xFF stuffing to the end of the Transport Packet.

If no section started in the following Transport Packet, then the same test should be carried out untila Transport Packet is found with the table_id==0x01 condition is found.

14. Check that there are no sections with table_id == 01 in any Transport Packet with a PID value other than 0x0001.

Program Map Section PIDs

1. The entries of the program_association_table should be used to obtain the values of PM PIDs. Each separate PM PID should then be checked with this procedure to ensure the PSI is compliant.

& Only sections with table_id == 0x02 are PMT sections.

& Other sections are allowed (short or long syntax doesn't matter) but they can only have table_id from 0x40 to 0xFE.

& descriptors should not be inconsistent with stream_types - eg don't use an audio descriptor on a video stream

& only one program per section

& section of prog p can only appear in PID indicated in PAT.

& program_number zero may not occur.

& PMTs PIDs can't be scrambled.

NIT

& if its there it follows the private section syntax

& if its there, its where it is referenced in the PAT under program_number zero.

Descriptors

Certain descriptors can only apply to certain stream types, and may be some only apply to either programs or elementary streams (Japanese DIS comment results will be relevant here).

Some descriptor tags are reserved.

Note specifically that the following are not required for a compliant bitstream.

a. There is no requirement to carry version of sections with the current_next_indicator set to a value of '0'.

b. There is no requirement for the sections of a table to be found in section_number order within the bitstream.

c. There is no requirement for all sections of a particular version number to be found in the Transport Stream an equal number of times.

d. There is no requirement for PSI data to be found at any minimum interval.

a) Timing tests
b) Buffer tests
1. System bitstream tests for Program Stream

a) Tests on the pack level

system_clock_reference: in successive packs, the system_clock_reference field contains encoded values which are samples of a nominal 27MHz system clock. The maximum interval between system_clock_reference fields is limited by the difference between encoded values in successive packs; this difference shall not exceed 0,7*90 000, as specified in xxxx of ITU-T Rec. H.222.0 | ISO/IEC 13818-1.

mux_rate (1): the value encoded in the mux_rate field shall be sufficently large that, if all bytes in the pack are transmitted at that rate, they are delivered to the system target decoder before the time the first byte of the subsequent pack is delivered. The time that the first byte of the subsequent pack is delivered may be derived from the system_clock_reference and mux_rate fields in that subsequent pack.

mux_rate (2): the mux_rate field shall not be encoded with the value zero.

packet rate: if the CSPS_flag in the system header is set to '1' and the rate specified in the mux_rate field is less than 4.5 Mbit/s then the rate at which packets arrive at the input of the system target decoder shall be less than 300 packets/s. If CSPS_flag is set to '1' and the rate specified in the mux_rate field is greater than 4.5 Mbit/s, the packet rate is bounded by a linear relation to the value encoded in the mux_rate field, as specified in 2.7.9 of ITU-T Rec. H.222.0 | ISO/IEC 13818-1.

length of a pack: the length of a pack may be determined by counting the bytes between successive pack start codes.

length of a packet: the length of a packet may be determined by counting the bytes between successive packet start codes.

a) Tests on the system header

general test (1): the first pack of an ISO/IEC 13818 stream shall contain a system header.

general test (2): a system header, if present in a pack, shall immediately follow the pack header.

general test (3): if an ISO/IEC 13818 stream contains more than one system header, then the values encoded in all the headers shall be identical.

header_length: the header_length field shall be encoded with a value equal to the number of bytes in the system header that follow the header_length field.

rate_bound: the rate_bound field shall denote a bitrate which is greater than or equal to the maximum bitrate value encoded in any mux_rate field in the same ISO/IEC 13818 stream.

audio_bound (1): the value encoded in the audio_bound field shall be greater than or equal to the maximum number of simultaneously active ISO/IEC 13818-3 audio streams in the ISO/IEC 13818 stream. For the purpose of this clause, an ISO/IEC 13818-3 audio stream is active if :

a) the input buffer of the system target decoder of that ISO/IEC 13818-3 audio stream is not empty, or if
b) one of the presentation units decoded from that ISO/IEC 13818-3 audio stream is being presented.

audio_bound (2): the value encoded in the audio_bound field shall be less than or equal to 32.

fixed_flag: if the fixed_flag is set to "1", then the values encoded in all system_clock_ reference fields shall satisfy 2.4.4.2 in ISO/IEC 13818-1.

CSPS_flag: if the CSPS_flag is set to "1", then the packet rate and the system target decoder buffer size shall satisfy 2.4.6 in ISO/IEC 13818-1.

system_audio_lock_flag: if the system_audio_lock_flag is set to "1", then the difference between the values encoded in any two presentation_time_stamp fields in audio packets of the same ISO/IEC 13818-3 audio stream shall correspond to the duration of the decoded audio access units in that ISO/IEC 13818-3 audio stream. For this purpose the duration (in terms of units of the system clock frequency) shall be derived from the number of samples and the sampling frequency of the decoded access units and the ratio SCASR as specified in 2.4.4.2 of ISO/IEC 13818-1. This assumes that no discontinuities occurred in the ISO/IEC 13818-3 audio stream in the presentation of the access units during the presentation period defined by both fields. See 2.4.5.4 of ISO/IEC 13818-1 for the definition of discontinuities.

system_video_lock_flag: if the system_video_lock_flag is set to "1", then the difference between the values encoded in any two presentation_time_stamp fields in video packets of the same ISO/IEC 13818-2 video stream shall correspond to the duration of the decoded pictures in that ISO/IEC 13818-2 video stream. For this purpose the duration (in terms of units of the system clock frequency) shall be derived from the picture rate of the decoded pictures and the ratio SCPR as specified in 2.4.4.2 of ISO/IEC 13818-1. This assumes that no discontinuities occurred in the ISO/IEC 13818-2 video stream in the presentation of the access units during the presentation period defined by both fields. See 2.4.5.4 of ISO/IEC 13818-1 for the definition of discontinuities.

video_bound (1): the value encoded in the video_bound field shall be greater than or equal to the maximum number of simultaneously active ISO/IEC 13818-2 video streams in the ISO/IEC 13818 stream. For the purpose of this clause, an ISO/IEC 13818-2 video stream is active if:

a) the input buffer of the system target decoder of that ISO/IEC 13818-2 video stream is not empty, or if
b) the reorder buffer of the system target decoder of that ISO/IEC 13818-2 video stream is not empty, or if
c) one of the presentation units decoded from that ISO/IEC 13818-2 video stream is being presented.

video_bound (2): the value encoded in the audio_bound field shall be less than or equal to 16.

stream_id (1): the value encoded in the stream_id field shall be one of the values permitted by table 2-18 (stream_id) in 2.4.3.7 in ISO/IEC 13818-1.

stream_id (2): the stream_id mechanism refers exactly once to each elementary stream in the multiplex.

STD_buffer_bound_scale: if the stream_id shall refer to an ISO/IEC 13818-3 audio stream, the STD_buffer_bound_scale shall be set to "0". If the stream_id refers to an ISO/IEC 13818-2 video stream, the STD_buffer_bound_scale shall be set to "1".

STD_buffer_size_bound: in the STD_buffer_size_bound field a value shall be encoded greater than or equal to the maximum value encoded in any of the STD_buffer_size fields in packets of the same elementary stream.

a) Tests on the PES level

stream_id: the value encoded in the stream_id field shall be one of the values permitted by table 1 (stream_id) in 2.4.4.2 in ISO/IEC 13818-1.

header_length: the header_length field shall be encoded with a value equal to the number of bytes in the packet that follow the header_length field.

stuffing bytes: a packet header shall contain no more than sixteen stuffing bytes.

STD_buffer_scale: if the stream_id refers to an ISO/IEC 13818-3 audio stream, the STD_buffer_scale shall be set to "0". If the stream_id refers to an ISO/IEC 13818-2 video stream, the STD_buffer_scale shall be set to "1",

STD_buffer_size (1): the STD_buffer_size for each elementary stream shall be encoded in the first packet of each elementary stream.

STD_buffer_size (2): if the CSPS_flag is set to "1" in the system header, and if the stream_id refers to an ISO/IEC 13818-3 audio stream, then the value encoded in the STD_buffer_size field shall define a buffer size value less than or equal to 4 096 Bytes, as specified in 2.4.6 in ISO/IEC 13818-1.

STD_buffer_size (3): if the CSPS_flag is set to "1" in the system header, and if the stream_id refers to an ISO/IEC 13818-2 video stream, then the value encoded in the STD_buffer_size field shall satisfy 2.4.6 in ISO/IEC 13818-1.

presentation_time_stamp (1): the presentation_time_stamp field may be encoded in a packet containing ISO/IEC 13818-3 audio data only if an audio access unit (an audio frame) commences in that packet.

presentation_time_stamp (2): the presentation_time_stamp field may be encoded in a packet containing ISO/IEC 13818-2 video data only if a video access unit (a video picture) commences in that packet.

presentation_time_stamp (3): the difference between the encoded values in any two successive presentation_time_stamps fields in packets of the same elementary stream, shall satisfy 2.4.5.3 in ISO/IEC 13818-1.

presentation_time_stamp (4): the presentation_time_stamp field shall satisfy 2.4.5.4 in ISO/IEC 13818-1.

decoding_time_stamp: the encoding of the decoding_time_stamp field shall satisfy 2.4.5.4 in ISO/IEC 13818-1.

buffer management (1): the access unit containing a data byte from an ISO/IEC 13818-3 audio or ISO/IEC 13818-2 video packet shall be decoded within one second after the data byte is received by the system target decoder in the STD model.

buffer management (2): the STD buffer shall neither overflow nor underflow; see 2.5.1.5.

a) Program Stream Map and Directory

b) Tests on timing accuracy

The entire ISO/IEC 13818 stream shall be constructed so that, in addition to meeting all other requirements, when the values and semantics of the presentation_time_stamp fields are compared with the audio and video samples to which they refer and their nominal sample rates, the calculated tolerance betweeen the actual sample frequency and the system_clock_frequency shall not exceed 100 parts per million. The timing relationships can be verified in practice solely by examining the bitstream. For example, to verify that the relationship between the audio sample rate and the system clock frequency meets this constraint, the following test can be performed:

a) Determine the nominal audio sample frequency as specified in the audio bitstream.

b) Find a PTS (PTS_1) associated with audio and the sample it refers to.

c) Count the number of audio samples until any other PTS (PTS_2) is found that refers to audio in the same stream, as long as no discontinuities occur between these two points.

d) Find the value of that second PTS (PTS_2), and the difference between PTS_2 and PTS_1.

e) Calculate the effective audio sample rate by dividing the number of samples passed by the difference in time as indicated by the PTSs (PTS_2 minus PTS_1) times 1/90 000).

f) If the calculated sample rate differs from the nominal by more than 100 ppm (allowing for integer arithmetic rounding effects) then this bitstream is not compliant.

a) Buffer overflow/underflow test

ISO/IEC 13818-1 specifies the requirement for the bitstream to be decodable by a System Target Decoder (STD). This means that, using the hypothetical STD model, when all audio and video streams are decoded and presented with exact synchronisation as specified by the respective Presentation Time Stamps, the STD buffers never underflow nor overflow. In other words, all bytes required for decoding audio and video are present in the STD buffers when they are needed, and the STD buffers are never filled beyond the capacity of the STD buffer sizes specified. The STD model is defined and described in 2.4.2 of ISO/IEC 13818-1.

In an ISO/IEC 13818 multiplexed bitstream, the buffer control mechanism specified for ISO/IEC 13818-2 video, using the vbv_delay field and the VBV model, is not directly applicable. In this case it is superseded by the STD model. The STD model allows for direct decoding of elementary streams from the multiplexed stream without the need to reconstruct the elementary streams first. However, it is possible to reconstruct elementary streams from the multiplex. In the case that an ISO/IEC 13818-2 video stream is reconstructed from the multiplexed stream, the VBV model again becomes applicable. For buffer overflow/underflow tests in cases where the VBV model is applicable, see 2.5.2.3.

Underflow and overflow of the STD buffer may be tested by constructing an STD model decoder. The test need not be performed in real time. For example, to test the behaviour of the STD video buffer, a test may be performed as follows:

a) The STD test model has a 90kHz time base (the system_clock_frequency) which is used to increment a counter.

b) Begin parsing the ISO/IEC 13818 stream starting at the first byte.

c) When a pack header is encountered, the system clock reference (SCR) is extracted, converted to its 33 bit value, and loaded onto a counter variable called System Time Clock which is incremented by the 90kHz time base.

d) Also from the pack header, extract the mux_rate field.

e) Calculate the number of bytes per unit of increments of the 90kHz time base as an accurate ratio of mux_rate to 90kHz (not an integer ratio). Note that mux_rate is coded in units of 50 bytes/s.

f) Read data from the source into the input buffer of the system target decoder at the rate indicated by the ratio of mux_rate to 90 kHz, calculated in step e). Note that accuracy must be maintained over the course of the bitstream, and the average number of bytes per system clock frequency increment is in general not an integer.

g) When packets of the video stream under test are encountered, the STD_buffer_scale and STD_buffer_size are extracted and converted to the STD buffer size in bytes. This is subsequently used as the upper limit on the STD buffer fullness. Presentation_time_ stamps and decoding_time_stamps are extracted and stored temporarily.

h) All packet_data bytes from the video stream under test are put into a fifo buffer in the order that they are received. This buffer is at least as large as the STD_buffer_size as specified for this stream in the packet header.

i) When a PTS or DTS is extracted, the PTS or DTS is saved temporarily until a video picture_start_code is encountered, and then the PTS or DTS is associated with the coded picture that immediately follows that picture_start_code. Note that if only the PTS is coded, the value of the DTS equals the value of the PTS.

j) When the System Time Clock value equals the DTS value of the picture that has been in the STD buffer the longest, the data of that coded picture is removed instantaneously from the STD buffer as specified in 2.4.2 in ISO/IEC 13818-1.

k) For a coded picture which has no DTS associated with it, the value of the DTS is derived from:
   - the DTS of the last coded picture to which a PTS or DTS is associated; this coded picture is referred to as coded picture (n);
   - the number (k) of coded pictures between coded picture (n) and the current coded picture;
   - the picture rate as encoded in the picture_rate field in the video sequence_header for the coded pictures between coded picture (n) and the current picture;
   - the derived DTS value is the sum of the previous coded DTS, which may have been implied by a coded PTS value, and the product of the number of succeeding pictures (k) and the picture period in units of (1/90 000) seconds.

l) Compliance requires that the STD buffer neither overflows nor underflows.

A. Video bitstream tests

1. Main Profile

a) Tests on the sequence header

horizontal_size (1): the value encoded in the horizontal_size field shall be at least 1.

horizontal_size (2): in 4:2:0 and 4:2:2 chroma structure sequences, horizontal_size must be a multiple of 2.

horizontal_size (3): within a video sequence, all horizontal_size fields shall be encoded with the same value.

horizontal_size (4): horizontal and vertical sizes are checked not to be a multiple of 4096.

vertical_size (1): the value encoded in the vertical_size field shall be at least 1.

vertical_size (2): In a frame picture, the derived value *mb_height* is checked against the actual number of coded macroblocks in the bitstream such that the line count is a multiple of 16. In field pictures, it is checked to be a multiple of 32.

vertical_size (3): within a video sequence, all vertical_size fields shall be encoded with the same value.

aspect_ratio_information (1): the binary value encoded in the aspect_ratio_information field shall be in the range from 0001 up to 0100

frame_rate_code (1): the binary value encoded in the picture_rate field shall be in the range from 0001 up to 1000.

frame_rate_code (2): The calculated value, *frame_rate*, generated by:

frame_rate = frame_rate_code * (frame_rate_extension_n+1)/(frame_rate_extension_d+1)

is checked such that it could not be represented by an entry in ISO/IEC 13818-2 Table 6.4

frame_rate_extension_n, frame_rate_extension_d (1): frame_rate_extension_n and frame_rate_extentsion_d shall not have the same value except if both have '0' value.

picture_rate (3): within a video sequence, all picture_rate fields shall be encoded with the same value.

bit_rate (1): the value encoded in the bit_rate field shall not be equal to zero.

bit_rate (2): within a video sequence, all bit_rate fields shall be encoded with the same value.

vbv_buffer_size (1): within a video sequence, all vbv_buffer_size fields shall be encoded with the same value.

vbv_buffer_size (2): if the constrained_parameters_flag is set to "1", then the value encoded in the vbv_buffer_size field shall be less than or equal to 40*1 024 bytes.

constrained_parameter_flag: within a 13818-2 video sequence, all constrained_parameter_flag fields shall have value '0'

intra_quantizer_matrix (1): the intra_quantizer_matrix field shall contain no values of intra_quant[i][j] equal to zero.

intra_quantizer matrix (2): the value of intra_quant[0][0] shall be equal to 8.

non_intra_quantizer_matrix (1): the non_intra_quantizer_matrix field shall contain no values of non_intra_quant[i][j] equal to zero.

user_data: user data shall not contain a string of 23 or more zero bits.

profile_and_level_indication (1): shall be a legal profile and level combination specified in ISO/IEC 13818-2 Chapter 8.

chroma_format(1): shall be in the range 01 to 11.
chroma_format(2): shall be consistent with Profile and Level indication.
low_delay(1): shall be consistent with Profile and Level indication.
video_format(1): shall be in the range 000 to 101.
chroma_primaries(1): shall be in the range 0000 0001 to 0000 0111.
transfer_characteristics(1): shall be in the range 0000 0001 and 0000 1000 or 0000 0100 and 0000 0111
matrix_coefficients(1): shall be in the range 0000 0001 and 0000 0010 or 0000 0100 and 0000 0111

a) Tests on the group of pictures layer

In this section, a number of parameters are used. The values of these parameters are obtained from the values encoded in the time_code fields at the Group of Pictures layer and in the picture_rate field in the Sequence Header as follows:

M is the value encoded in the time_code_minutes field;

S is the value encoded in the time_code_seconds field;

P is the value encoded in the picture_rate field in the sequence header, rounded to the nearest integer number.

drop_frame_flag: the drop_frame_flag may only be set to "1", if the picture_rate field in the Sequence Header indicates a picture rate of 29,97 Hz.

time_code_hours: the value encoded in the time_code_hours field shall be in the range from zero up to 23.

time_code_minutes: the value encoded in the time_code_minutes field shall be in the range from zero up to 59.

time_code_seconds: the value encoded in the time_code_seconds field shall be in the range from zero up to 59.

time_code_pictures: the value encoded in the time_code_pictures field shall be in the range from F0 up to (P-1), where

F0 = 2 , if the drop_frame_flag is set to "1" and S = 0 and M % 10 - 0
= 0 otherwise.

time_code: the time_code field shall be encoded in compliance with 2.4.3.3 in ISO/IEC 11172-2.

user_data: user data shall not contain a string of 23 or more zero bits.

a) Tests on the picture layer

temporal_reference (1): for the first picture, in display order, within a Group of Pictures the temporal_reference field shall be encoded with the value zero.

temporal_reference (2): the value encoded in the temporal_reference field shall be the display order number modulo 1 024 of the picture within the Group of Pictures.

temporal_reference (3): In the second field of a field-coded picture, temporal_reference should have the same value as the first coded field.

picture_coding_type (1): the binary value encoded in the picture_coding_type field shall be in the range from 001 up to 011.

picture_coding_type (2): the last picture, in display order, of a Group of Pictures shall be an I-, P picture

picture_coding_type (3): if a Group of Pictures contains at least two pictures of type I or P, the last two of which have coding order numbers m and n, then all B-pictures with coding order numbers m+1, ..., n-1 shall also be included in the Group of Pictures.

picture_coding_type (4): Let m and n be the natural display order of any I- or P-pictures and j and k be the decoding order. If m>n, then j>k and vice-versa.

picture_coding_type (4): A sequence shall contain at least one I or P picture.
picture_coding_type (5): Both fields in an interlaced picture shall have the same coding type with the exception that an Intra-field may have a P field-picture as the second field of the pair.
picture_coding_type (6): B pictures (011) shall not occur in Simple Profile or low delay sequences

vbv_delay (1): if the bit_rate field in the Sequence Header indicates variable bitrate operation, then the vbv_delay field shall be encoded with the hexadecimal value FFFF.

vbv_delay (2): if the bit_rate field in the Sequence Header indicates constant bitrate operation, then the vbv_delay field shall be encoded such that neither overflow nor underflow occurs in the VBV buffer. Overflow and underflow do not occur, if the following conditions apply:

a) all bytes of picture (i) shall arrive before picture (i) is to be decoded.

N(i) / R ^{2} V(i) / 90 000,

where

i indicates the number of the picture in coding order

N(i) is the number of bits after the final byte of the picture start code of picture (i) that are removed from the VBV buffer at decoding of picture (i) (see also ISO/IEC 11172-2, annex C)

R is the bitrate in bits/s, derived from the bit_rate field as encoded in the sequence header

V(i) is the value encoded in the vbv_delay field of picture (i) in units of the 90 kHz system clock

b) the bitrate at which all picture data are delivered shall be constant.

N(i)/T(i) = R(i)

where

i as specified in a)

N(i) as specified in a)

T(i) is the elapsed time in seconds between the arrival of the final bytes of the two successive picture start codes of pictures (i) and (i+1), as derived from :

T(i) = (V(i)/90 000) + (1/P) - (V(i+1)/90 000),

where

V(i) is the value of the vbv_delay of picture (i), as encoded in the vbv_delay field of picture i in units of the 90 kHz system clock

P is the picture rate

R(i) is the bitrate in bits/s during delivery of the coded data of picture (i) applied by the encoder. The encoder used constant bitrate operation. The value of R(i) found from this formula should therefore be constant for each picture (i) within the accuracy constraints of the parameters. If underflow occurs for picture (i), a higher value of R(i) is found. R(i) rounded upwards equals R, the value encoded in the Sequence Header in units of 400 bits/s.

c) the number of bits in the VBV buffer immediately before picture (i) is decoded shall be less than the size of the VBV buffer.

H(i) + V(i)*R(i)/90 000 ^{2} B(vbv),

where

i as specified in a)

H(i) is the number of bits of the picture start code and any preceding (header) data that is removed instantaneously at decoding of picture (i); see annex C of ISO/IEC 11172-2.

V(i) as specified in a)

R(i) as specified in b)

B(vbv) is the value of the vbv_buffer_size in bits

forward_f_code, backward_f_code (1): shall have the value 111 in MPEG-2 sequences.
full_pel_forward_vector, full_pel_backward_vector(1): shall have the value '0' in MPEG-2 sequences.
f_code(1): shall not have zero values..
f_code(2): shall have 1111 in Intra pictures with no error concealment vectors.
f_code(3): [r][1][t] backward shall have value 0xF in I and P pictures.
intra_dc_precision(1): shall agree with Profile and Level.
picture_structure(1): shall have value between 01 and 11.
picture_structure(2): 10 and 11 are illegal in progressive sequences.
picture_structure(3): the field parity shall be alternately top and bottom.
frame_pred_frame_dct(1): shall have value '1' in progressive sequences and pictures.
frame_pred_frame_dct(2): shall have value '0' in field pictures.
repeat_first_field(1): shall be zero in field pictures.
chroma_420_type(1): shall have the same parity as progressive_frame
progressive_frame(1): shall be '1' in progressive sequences.

user_data: user shall not contain a string of 23 or more zero bits.

a) Tests on the slice layer

slice_vertical_position (1): the value encoded in the slice_vertical_position field shall be in the range from 1 to 175.

slice_vertical_position (2): the value encoded in the slice_vertical_position shall not be greater than the vertical picture size in units of 16 lines, rounded upwards. The vertical picture size is indicated in the Sequence Header.

quantizer_scale: the value encoded in the quantizer_scale field shall not be equal to zero.

reserved_bits: shall have zero value.

extra_bit_slice is checked to have 0 value, however as a service to delinquent bitstreams, extra_information_slice may be parsed out.

a) Tests on the macroblock layer

macroblock_address_increment (1): the indicated macroblock shall be within the boundaries of the picture.

macroblock_address_increment (2): the first slice of the picture shall start with the first macroblock in the picture.

macroblock_address_increment (3): the first macroblock of a slice shall not be skipped.

macroblock_address_increment (4): the horizontal position of the first macroblock of a slice shall be less than or equal to the width of the picture in units of macroblocks.

macroblock_address_increment (5): the last macroblock of a slice shall not be skipped.

macroblock_address_increment (6): the last slice of the picture shall end with the last macroblock of the picture.

macroblock_address_increment (7): slices shall not overlap.

macroblock_address_increment (8): there shall be no gaps between slices in the same picture.

macroblock_address_increment (9): within Main Profile I pictures macroblocks shall not occur.

macroblock address increment (10): macroblock skipping is not permitted after an Intra macroblock

macroblock_address_increment(11): macroblock stuffing is illegal in MPEG-2 sequences.

macroblock_type (1): in a B-picture, a skipped macroblock shall not follow an intra-coded macroblock.

macroblock_type (2): each macroblock shall be intra-coded at least once per every 132 times it is coded in a P-picture without an intervening I-picture.

macroblock_type (3): if the macroblock_type field indicates macroblock_motion_forward and in case the picture is a B-picture that is preceded by exactly one I-picture in the same Group of Pictures, then the closed_gop flag shall be set to "0".

macroblock_type (4): if the macroblock_type field indicates macroblock_motion_forward and in case the picture is a B-picture that is preceded by exactly one I-picture in the same Group of Pictures, and if that Group of Pictures is the first Group of Pictures in the sequence, then the broken_link flag shall be set to "1".

macroblock_type(5): When intraa_slice is set, all macroblocks within that slice are checked to be of type Intra.

macroblock_type(6): Dual prime (11) shall not occur in Simple or Main Profile B pictures.

macroblock_type(7): No B-coded pictures shall exist between a picture containing at least one Dual Prime macroblock and the reference picture.

macroblock_type(8): for the special $^{I}_{P}$ case:

- (*macroblock_motion_forward==0 && macroblock_intra==0*)combination in macroblock_type is illegal.
- Dual Prime is checked not to occur.
- motion_vertical_field_select shall always have opposite parity.
- No skipped macroblocks

frame_motion_type and **field_motion_type** shall have values ranging from 01 to 11.

reconstructed motion vectors(1): All reconstructed vectors are checked such they fall within the legal range bounded by [*low, high*].

reconstructed motion vectors(1):

reconstructing motion vectors: Intermediate and final motion vectors for Dual Prime (same parity vectors, scaled opposite parity vectors, and the final scale+dmv vectors) shall lie within the [*low, high*] and picture boundaries.

reconstructed motion vectors:(1) Field MV in frame pictures are checked to have half the range (high:low) of the range specified by f_code.

reconstructed motion vectors(2): each reconstructed motion vector shall refer to a macroblock that is fully within the boundaries of the picture [horizontal_size - 16, vertical_size - 16].

a) Tests on the block layer

number of DCT coefficients: for each transformed block the block indices of the coefficients shall be in the range from zero to sixty-three. The block index defines the position of the coefficient in the array of quantized DCT coefficients in zig-zag or alternate scanning order.

1. SNR Scalability Profile

[this section is an inclusion of "Compliance Verifier Tool for MPEG-2 scalable video," Lameillieure, Herpel, Henot]

[To be modified to the style of the Main Profile section above]

The numbers listed here refer to the clause in ISO/IEC 13818-2 in which the Compliance test is first introduced.

7.8.1 Higher syntactic structures

The layer_id's of two bitstreams have to be consecutive numbers.

The sequence header of the enhancement layer has to be identical to the one in the lower layer except for the values of bit_rate, vbv_buffer_size, load_intra_quantiser_matrix, intra_quantiser_matrix, load_non_intra_quantiser_matrix and non_intra_quantiser_matrix. The value of load_intra_quantiser_matrix has to be zero in the enhancement layer.

The sequence extension has to be identical to the one in the lower layer except for the values of porfile_and_level_indication, chroma_format, bit_rate_extension and vbv_buffer_size_extension.

The chroma_format of the enhancement layer shall not be less than the one in the base layer.

The sequence_display_extension shall not be present in the enhancement layer.

The sequence_scalable_extension shall be present in the enhancement layer with scalable_mode = "SNR scalability"

If there is a GOP header in the enhancement layer, it will be identical to the one in the base layer.

The picture headers in the enhancement layer shall be identical to the ones in the base layer, except for the values of vbv_delay.

The picture_coding_extension in the enhancement layer shall be identical to the one in the base layer except for the vlaue of q_scale_type and alternate_scan

If there is a quant_matrix_extension in the enhancement layer, the values of load_intra_quantiser_matrix and load_chroma_intra_quantiser_matrix shall be zero.

There shall be no pan_scan_extension in the enhancement layer.

Slices of the enhancement layer have to coincide with the slices of the base layer.

7.8.2.1

IF dct_type is present in both layers, it shall have the same value.

8.5

If the level is Main level, the sum of the VBV buffer sizes of both layers shall not exceed 1835008 bit. For this level, the VBV buffer size of the base layer shall not exceed 1212416 bit.

The VBV buffer size and VBV delays shall not be zero.

1. Spatial scalability Profile

6.3.8

The values of lower_layer_prediction_horizontal_size and lower_layer_prediction_vertical_size shall have exactly the same value as the lower layer horizontal_size and vertical_size respectively.

7.7

There shall be no gaps between slices.

7.7.1

In the case of interlaced sequences in both layers, the picture_structure has to be the same in the coded picture of the enhancement layer and in the reference picture of the base layer.

7.7.3.2

No spatial prediction shall be made for macroblocks in the enhancement layer that not lie wholly within the upsampled lower layer reconstructed frame.

7.7.4

The value of spatial_temporal_weight_code-table_index shall only have values that are allowed in table 7-20, according to the scanning formats in both layers.

7.7.6

There shall be no skipped macroblock following a spatial-only macroblock in B-pictures.

A. Audio bitstream tests

2.5.3.1 Extension of ISO/IEC 11172-3 Audio Coding to Lower Sampling Frequencies

1. General tests

layer: the Layer field shall not be encoded with the binary value 00.

bitrate: the bitrate field shall not be encoded with the binary value 1111.

sampling_frequency: the sampling frequency field shall not be encoded with the binary value 11.

padding: padding shall be applied such that the accumulated length of the coded audio frames, after a certain number of audio frames, shall not deviate more than (+0,-1) slot from the value specified in 2.4.2.3 of ISO/IEC 11172-3. This shall apply only if the layer, the bitrate and the sampling frequency do not change in the course of the considered audio frames.

emphasis: the emphasis field shall not be encoded with the binary value 10.

protection: if the protection bit is set to "0", then the correct CRC16 value shall be in the crc_check field

ID: the ID flag shall be set to "0".

a) Tests on Layer I

allocation: the allocation[sb] or allocation[ch][sb] field shall not be encoded with the binary value 1111.

scalefactor: the scalefactor[sb] or scalefactor[ch][sb] field shall not refer to index 63.

samples: for the coded representation of subband samples the valid range is from zero up to (nlevels - 2), where nlevels equals the number of levels used for quantization of that sample, that is the coded representation of a sample shall not consist of a bitstring with only "1"s.

frame length (1): the bit allocation shall be such that the total number of bits for a frame does not exceed the frame length for Layer I.

frame length (2): for Layer 1 the frame length shall equal the number of slots times the slot size for Layer I.

a) Tests on Layer II

scalefactor: the scalefactor[sb][p] or scalefactor[ch][sb][p] field shall not refer to index 63

samples: for un-grouped samples the coded representation of subband samples the valid range is from zero up to (nlevels -2), where nlevels equals the number of levels used for quantization of that sample, that is the coded representation of a sample shall not consist of a bitstring with only "1"s. For grouped samples the range shall be from zero up to 26 if nlevels equals 3, from zero up to 124 if nlevels equals 5, and from zero up to 728 if nlevels equals 9.

frame length (1): the bit allocation and the scalefactor select information shall be such that the total number of bits for a frame does not exceed the frame length for Layer II.

frame length (2): for Layer II the frame length shall equal the number of slots times the slot size for Layer II.

a) Tests on Layer III

part2_3_length: the value encoded in the part2_3_length[gr] or part2_3_length[gr][ch] field shall correspond to the total length of scalefactors and Huffman encoded data.

table_select: the table_select[region][gr] or table_select[region][gr][ch] fields shall be encoded correctly.

frame_length (1): the Huffman code data shall be such that the total number of bits for a frame does not exceed the frame length for Layer III.

frame length (2): for Layer III the frame length shall equal the number of slots times the slot size for Layer III.

buffer control: the value of main_data_begin shall comply with the buffer considerations specified in 2.4.3.4 of ISO/IEC 11172-3.

2.5.3.2 Low bit rate coding of Multichannel Audio

Due to the compatibility of ISO/IEC 13818-3 multichannel audio coding with ISO/IEC 11172-3, ISO/IEC 11172-4 applies to the MPEG-1 part of the bit stream. Furthermore, the following tests apply to an ISO/IEC 13818-3 bit stream.

a) General tests

The following table gives an overview of the allowed combinations of mode in the MPEG-1 header and the multichannel options in the mc_header.

MPEG-1 mode	MPEG-2 multichannel option				
	Center	mono/stereo surround	2nd stereo programme	LFE	multilingual
mono/dual	no	no	yes	no	yes
stereo/joint stereo	yes	yes	yes	yes	yes

dematrix_procedure: The dematrix_procedure ’10Õ may only occur in 3/1 or in 3/2 configuration.

a) Tests on Layer I and Layer II

tc_allocation: The following combinations of configuration and tc_allocation are not allowed.

configuration	forbidden tc_allocation
3/1	5, 6, 7
3/0 or 3/0+2/0	3
2/1	3

If phantom coding is used, the combinations of configuration and tc_allocation are further restricted:

configuration	forbidden tc_allocation
3/2	1, 2, 6, 7
3/1	1, 2, 5, 6, 7
3/0 or 3/0+2/0	1, 2, 3
2/2 and 2/1	not applicable with phantom coding

dyn_cross_mode: The following combinations of configuration and dyn_cross_mode are not allowed.

configuration	forbidden dyn_cross_mode
3/2	15
3/1	5, 6, 7
2/2	5, 6, 7

lf_scalefactor, scalefactor[mch][sb][p],scalefactor[mlch][sb][p]: the lf_scalefactor and scalefactor[mch][sb][p] and scalefactor[mlch][sb][p] fields shall not refer to index 63.

lf_sample[gr], sample[mch,sb,s], sample[mlch,sb,s], samplecode[mch,sb,gr], samplecode[mlch,sb,gr]: for un-grouped samples the coded representation of subband samples the valid range is from zero up to (nlevels -2), where nlevels equals the number of levels used for quantization of that sample, that is the coded representation of a sample shall not consist of a bitstring with only "1"s. For grouped samples the range shall be from zero up to 26 if nlevels equals 3, from zero up to 124 if nlevels equals 5, and from zero up to 728 if nlevels equals 9.

frame length (1): the bit allocation and, in Layer II, the scalefactor select information of the MPEG-1 defined part, and the bit allocation, scalefactor select information, and composite status info of the multichannel extension, and the bit allocation and the scalefactor select information of the multilingual extension, shall be such that the total number of bits for a frame does not exceed the frame length plus the length of the extension frame.

frame length (2): the frame length shall equal the number of slots times the slot size.

a) Tests on Layer III

part2_3_length: the value encoded in the part2_3_length[gr] or part2_3_length[gr][ch] field shall correspond to the total length of scalefactors and Huffman encoded data.

table_select: the table_select[region][gr] or table_select[region][gr][ch] fields shall be encoded correctly.

frame_length (1): the Huffman code data shall be such that the total number of bits for a frame does not exceed the frame length for Layer III.

frame length (2): for Layer III the frame length shall equal the number of slots times the slot size for Layer III.

buffer control: the value of main_data_begin shall comply with the buffer considerations specified in 2.4.3.4 of ISO/IEC 13818-3.

I. Procedures to test decoder compliance

A. System decoder tests

1. Transport Stream

a) basic

b) additional

2. Program Stream

a) basic

b) additional

A. Video decoder tests

[This is an inclusion of Henot & Lameillieure, "Conformance testing of video decoders."]

1. statistical behaviour of decoders on conformance bitstreams

a) testing conditions

Confomance testing should be ideally possible on every type of decoder. However, the video MPEG2 specification only concerns the decoding process from a bitstream to a digital video output. This output, in the case of MP@ML, is only in the form of a 4:2:0 signal.

The conformance testing of systems that provide only an analog video output (NTSC, PAL, SECAM, RGB...), or systems of which the only output is the display, is not considered in this section. It will be considered in the next section on decoder breakdown

The proposed method can be applied to video systems with digital video output. In the case of MP@ML video decoders tests can be performed either on 4:2:0, or on 4:2:2 output. However, the results of tests on 4:2:2 signals should be carefully interpreted as the chroma conversion from 4:2:0 to 4:2:2 is outside the scope of the standard and will certainly bias the results on chrominance. As a consequence, it should be recommended to perform the test in 4:2:0 format, and , if not possible 4:2:2 should be considered as an alternative of limited value for chrominance signals.

The proposed set up for the test is shown on figure 1.

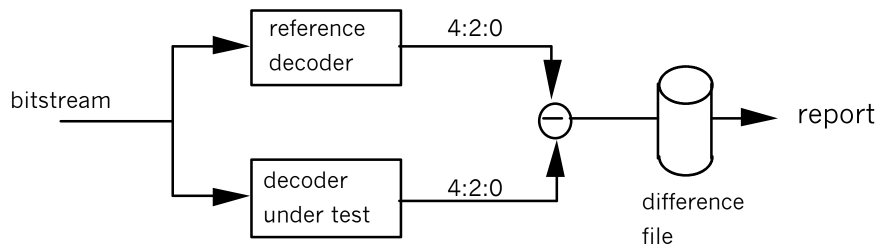

Figure 1: test configuration for a decoder with a digital output

a) report on the analysis of difference files

A report on the difference file is generated for each conformance bitstream.

The report consists of:

- the histogram of the pel difference per field between the output of the reference decoder and the decoder under test for each component (Y,U,V) of the signal

- the number and addresses and the Mean Absolute Error of erroneous [errors not explained by mismatch] MBs, and the erroneous component in each erroneous MB. A MB is declared erroneous when at least one pixel of the difference file has an absolute value that cannot be explained by mismatch.

A. Audio decoder tests

To test audio decoders, ISO/IEC JTC1 SC29/WG11 supplies a number of test sequences. Supplied sequences cover Full Layer N Decoders and Core Layer N decoders. For a supplied test sequence, testing can be done by comparing the output of a decoder under test with a reference output also supplied by ISO/IEC JTC1 SC29/WG11. Measurements are carried out relative to full scale where the output signals of the decoders are normalized to be in the range between -1 and +1.

To be called an ISO/IEC 13818-3 audio decoder, the decoder shall provide an output such that the rms level of the difference signal between the output of the decoder under test and the supplied reference output is less than 2^{-15}/sqrt(12) for the supplied sine sweep (20Hz-10kHz) with an amplitude of -20dB relative to full scale. In addition, the difference signal shall have a maximum absolute value of at most 2^{-14} relative to full-scale.

To be called a limited accuracy ISO/IEC 13818-3 audio decoder, the decoder shall provide an output for a provided test sequence such that the rms level of the difference signal between the output of the decoder under test and the supplied reference output is less than 2^{-11} /(sqrt(12) for the supplied sine sweep (20Hz-10kHz) with an amplitude of -20dB relative to full scale.

The above two tests only verify the computational accuracy of an implementation.

I. Annex A
II. Patent statements

(This annex does not form an integral part of this Recommendation | International Standard)
The user's attention is called to the possibility that, for some of the processes specified in this part of ISO/IEC 13818, conformance with this specification may require use of an invention covered by patent rights.
By publication of this part of ISO/IEC 13818, no position is taken with respect to the validity of this claim or of any patent rights in connection therewith. However, each company listed in this Annex has undertaken to file with the Information Technology Task Force (ITTF) a statement of willingness to grant a license under such rights that they hold on reasonable and non-discriminatory terms and conditions to applicants desiring to obtain such a license.
Information regarding such patents can be obtained from the following organisations.
The table summarises the formal patent statements received and indicates the parts of the standard to which the statement applies. The list includes all organisations that have submitted informal patent statements. However, if no "X" is present, no formal patent statement has yet been received from that organisation.

Company	**V**	**A**	**S**
AT&T	X	X	X
BBC Research Department			
Bellcore	X		
Belgian Science Policy Office	X		
BOSCH	X	X	X
CCETT			
CSELT	X		
David Sarnoff Research Center	X	X	X
Deutsche Thomson-Brandt GmbH	X	X	X
France Telecom CNET			
Fraunhofer Gesellschaft		X	X
GC Technology Corporation	X	X	X
General Instruments			
Goldstar			
Hitachi, Ltd.			
International Business Machines Corporation	X	X	X
IRT		X	
KDD	X		
Massachusetts Institute of Technology	X	X	X
Matsushita Electric Industrial Co., Ltd.	X	X	X
Mitsubishi Electric Corporation			
National Transcommunications Limited			
NEC Corporation		X	
Nippon Hoso Kyokai	X		
cont			

Company	**V**	**A**	**S**
Nippon Telegraph and Telephone	X		
Nokia Research Center	X		
Norwegian Telecom Research	X		
Philips Consumer Electronics	X	X	X
OKI			
Qualcomm Incorporated	X		
Royal PTT Nederland N.V., PTT Research (NL)	X	X	X
Samsung Electronics			
Scientific Atlanta	X	X	X
Siemens AG	X		
Sharp Corporation			
Sony Corporation			
Texas Instruments			
Thomson Consumer Electronics			
Toshiba Corporation	X		
TV/Com	X	X	X
Victor Company of Japan Limited			

I.

II. Annex B

III. Bibliography

(This annex does not form an integral part of this Recommendation | International Standard)

1 Arun N. Netravali & Barry G. Haskell "Digital Pictures, representation and compression" Plenum Press, 1988

2 Didier Le Gall "MPEG: A Video Compression Standard for Multimedia Applications" Trans. ACM, April 1991

3 C Loeffler, A Ligtenberg, G S Moschytz "Practical fast 1-D DCT algorithms with 11 multiplications" Proceedings IEEE ICASSP-89, Vol. 2, pp 988-991, Feb. 1989

4 See the Normative Reference for ITU-R Rec 601 (formerly CCIR Rec 601)

5 See the Normative Reference for IEC Standard Publication 461

6 See the Normative Reference for ITU-T Rec. H.261

7 See the Normative reference for IEEE Standard Specification P1180-1990

8 ISO/IEC 10918-1 | ITU-T T.81 (JPEG)

9 E Viscito and C Gonzales "A Video Compression Algorithm with Adaptive Bit Allocation and Quantization", Proc SPIE Visual Communications and Image Proc '91 Boston MA November 10-15 Vol. 1605 205, 1991

10 A Puri and R Aravind "Motion Compensated Video Coding with Adaptive Perceptual Quantization", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 1 pp 351 Dec. 1991.

11 C. Gonzales and E. Viscito, "Flexibly scalable digital video coding". Image Communications, Vol. 5, Nos. 1-2, February 1993

12 A.W.Johnson, T.Sikora and T.K. Tan, "Filters for Drift Reduction in Frequency Scalable Video Coding Schemes" *<Transmitted for publication to Electronic Letters.>*

13 R.Mokry and D.Anastassiou, "Minimal Error Drift in Frequency Scalability for Motion-Compensated DCT Coding". IEEE Transactions on Circuits and Systems for Video Technology, *<accepted for publication>*

14 K.N. Ngan, J. Arnold, T. Sikora, T.K. Tan and A.W. Johnson. "Frequency Scalability Experiments for MPEG-2 Standard". Asia-Pacific Conference on Communications, Korea, August 1993.

15 T. Sikora, T.K. Tan and K.N. Ngan, "A Performance Comparison of Frequency Domain Pyramid Scalable Coding Schemes Within the MPEG Framework". Proc. PCS, Picture Coding Symposium, Lausanne, pp. 16.1 - 16.2, Switzerland March 1993.

16 Masahiro Iwahashi, “Motion Compensation Technique for 2:1 Scaled-down Moving PicturesÓ. 8-14, Picture Coding Symposium '93.

17 Sikora, T. and Pang, K., "Experiments with Optimal Block-Overlapping Filters for Cell Loss Concealment in Packet Video", Proc. IEEE Visual Signal Processing and Communications Workshop, Melbourne, 21-22 Sept. 1993, pp. 247-250.

18 A. Puri "Video Coding Using the MPEG-2 Compression Standard", <to appear> Proc SPIE Visual Communications and Image Proc '93 Boston MA November,1993.

19 A. Puri and A. Wong "Spatial Domain Resolution Scalable Video Coding", <to appear> Proc SPIE Visual Communications and Image Proc '93 Boston MA November,1993.