ETSI TS 101 154 V1.7.1 (2005-06)

Technical Specification

Digital Video Broadcasting (DVB); Implementation guidelines for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream

Reference
RTS/JTC-DVB-170

Keywords
DVB, broadcasting, TV, digital, MPEG, video

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by Joint Technical Committee (JTC) Broadcast of the European Broadcasting Union (EBU), Comité Européen de Normalisation ELECtrotechnique (CENELEC) and the European Telecommunications Standards Institute (ETSI).

The original TR 101 154 was based on the DVB document A001 and it covered only the 25 Hz SDTV Baseline IRD. The first revision of TR 101 154 extended the scope to encompass both the 25 Hz SDTV Baseline IRD and the 25 Hz SDTV IRD with a digital interface intended for connection to a bitstream storage device such as a digital VCR. The second revision covered both the Baseline IRD and the IRD with digital interface for 25 Hz SDTV, 25 Hz HDTV, 30 Hz SDTV and 30 Hz HDTV. Subsequent revisions added optional support for the video Active Format Description (annex B), AC-3 audio and Enhanced AC-3 audio (annex C) and Ancillary Data for MPEG audio (annex D) and the Coding of Data Fields in the Private Data Bytes of the Adaptation Field (annex E) and optional support for DTS audio (annex F) and receiver-mixed audio (annex G). This revision adds optional support of H.264/AVC for video content and optional support of HE AAC and HE AACv2 (annex H) for audio content. The revisions to the TR have been developed in a largely backwards compatible manner, i.e. no changes to the mandatory functionality of a previously defined IRD have been made between one edition of the TR and the next.

The present document is complementary to TR 102 154, which provides Implementation Guidelines for the use of Video and Audio Coding in Contribution and Primary Distribution Applications based on the MPEG-2 Transport Stream.

The present document is complementary to TS 102 005, which provides Implementation Guidelines for the use of audio-visual content in DVB services delivered over IP.

NOTE: The EBU/ETSI JTC Broadcast was established in 1990 to co-ordinate the drafting of standards in the specific field of broadcasting and related fields. Since 1995 the JTC Broadcast became a tripartite body by including in the Memorandum of Understanding also CENELEC, which is responsible for the standardization of radio and television receivers. The EBU is a professional association of broadcasting organizations whose work includes the co-ordination of its members' activities in the technical, legal, programme-making and programme-exchange domains. The EBU has active members in about 60 countries in the European broadcasting area; its headquarters is in Geneva.

European Broadcasting Union
CH-1218 GRAND SACONNEX (Geneva)
Switzerland
Tel: +41 22 717 21 11
Fax: +41 22 717 24 81

Founded in September 1993, the DVB Project is a market-led consortium of public and private sector organizations in the television industry. Its aim is to establish the framework for the introduction of MPEG-2 based digital television services. Now comprising over 200 organizations from more than 25 countries around the world, DVB fosters market-led systems, which meet the real needs, and economic circumstances, of the consumer electronics and the broadcast industry.

Introduction

The present document presents guidelines covering coding and decoding using the MPEG-2 system layer, video coding and audio coding.

The guidelines presented in the present document for the Integrated Receiver-Decoder (IRD) are intended to represent a minimum functionality that all IRDs of a particular class are required to either meet or exceed. It is necessary to specify the minimum IRD functionality for basic parameters, if broadcasters are not to be prevented from ever using certain features. For example, if a significant population of IRDs were produced that supported only the Simple Profile, broadcasters would never be able to transmit Main Profile bit-streams.

IRDs are classified in five dimensions as:

- "25 Hz" or "30 Hz", depending on whether the nominal video frame rates based on 25 Hz or 30 000/1 001 Hz (approximately 29,97 Hz) are supported. It is expected that 25 Hz IRDs will be used in those countries where the existing analogue TV transmissions use 25 Hz frame rate and 30 Hz IRDs will be used in countries where the analogue TV transmissions use 30 000/1 001 Hz frame rate. There are also likely to be "dual-standard" IRDs which have the capabilities of both 25 Hz and 30 Hz IRDs.

- "SDTV" or "HDTV", depending on whether or not they are limited to decoding pictures of conventional TV resolution. The capabilities of an SDTV IRD are a sub-set of those of an HDTV IRD.

- "with digital interface" or "Baseline", depending on whether or not they are intended for use with a digital bitstream storage device such as a digital VCR. The capabilities of a Baseline IRD are a sub-set of those of an IRD with digital interface.

- MPEG-2 video or H.264/AVC video coding formats.

- Audio coding formats according to clause 6 or any of the annexes C, F or H.

To give a complete definition of an IRD, all five dimensions need to be specified, e.g.:

- 25 Hz SDTV Baseline IRD MPEG-2 video.25 Hz SDTV Baseline IRD MPEG-2 video, MPEG-1 Layer 2 audio, for an IRD able to decode 720 × 576 interlaced 25 Hz video pictures.

- 30 Hz HDTV Baseline IRD H264/AVC video, HE AAC Level 4 audio, for an IRD able to decode up to 1920 × 1080 interlaced 30 Hz video pictures or 1280 × 720 progressive 60 Hz video pictures.

All the formats supported by an IRD conforming to this specification are listed in annex A.

It should be noted that in DVB systems the source picture format, encoded picture format and display picture format do not need to be identical. For example, HDTV source material may be broadcast as an SDTV bitstream after down-conversion to SDTV resolution and encoding within the constraints of MPEG-2 video Main Profile at Main Level. The IRD receiving the bitstream may then up-convert the decoded picture for display at HDTV resolution.

Another notable feature of the DVB system is that a single Transport Stream may contain programme material intended for more than one type of IRD. A typical example of this is likely to be the simulcasting of SDTV and HDTV video material. In this case an SDTV IRD will decode and display SDTV pictures whilst an HDTV IRD will decode and display HDTV pictures from the same Transport Stream.

Where a feature described in the present document is mandatory, the word "shall" is used and the text is in italic; all other features are optional. The functionality is specified in the form of constraints on MPEG-2 systems, video and audio formats which the IRDs are required to decode correctly.

The specification of these baseline features in no way prohibits IRD manufacturers from including additional features, and should not be interpreted as stipulating any form of upper limit to the performance. The guidelines do not cover features, such as the IRDs up-sampling filter, which affect the quality of the displayed picture rather than whether the IRD is able to decode pictures at all. Such issues are left to the marketplace.

The guidelines presented for IRDs observe the following principles:

- wherever practical, IRDs should be designed to allow for future compatible extensions to the bit-stream syntax;
- all "reserved" and "private" bits in MPEG-2 systems, video and audio formats should be ignored by IRDs not designed to make use of them.

The rules of operation for the encoders are features and constraints which the encoding system should adhere to in order to ensure that the transmissions can be correctly decoded. These constraints may be mandatory or optional. Where a feature or constraint is mandatory, the word "shall" is used and the text is italic; all other features are optional.

Clauses 4 to 6 and the annexes, provide the guidelines for the Digital Video Broadcasting (DVB) systems layer, video, and audio respectively. For information, some of the key features are summarized below, but clauses 4 to 6 and the annexes should be consulted for all definitions:

Systems:

- MPEG-2 Transport Stream (TS) is used;
- Service Information (SI) is based on MPEG-2 program-specific information;
- Scrambling is as defined in ETR 289 [5];
- Conditional access uses the MPEG-2 Conditional Access CA_descriptor;
- Partial Transport Streams are used for digital VCR applications.

Video:

- MPEG-2 Main Profile at Main Level is used for MPEG-2 encoded SDTV;
- MPEG-2 Main Profile at High Level is used for MPEG-2 encoded HDTV;
- H.264/AVC Main Profile at Level 3 is used for H.264/AVC SDTV;
- H.264/AVC High Profile at Level 4 is used for H.264/AVC HDTV;
- The 25 Hz MPEG-2 SDTV IRD and 25 Hz H.264/AVC SDTV IRD support 25 Hz frame rate;
- The 25 Hz MPEG-2 HDTV IRD and 25 Hz H.264/AVC HDTV IRD support frame rates of 25 Hz or 50 Hz;
- The 30 Hz MPEG-2 SDTV IRD and 30 Hz H.264/AVC SDTV IRD support frame rates of 24 000/1 001, 24, 30 000/1 001 and 30 Hz;
- The 30 Hz MPEG-2 HDTV IRD and 30 Hz H.264/AVC HDTV IRD supports frame rates of 24 000/1 001, 24, 30 000/1 001, 30, 60 000/1 001 and 60 Hz;
- SDTV pictures may have either 4:3, 16:9 or 2.21:1 aspect ratio; IRDs support 4:3 and 16:9 and optionally 2.21:1 aspect ratio;
- MPEG-2 HDTV pictures have 16:9 or 2.21:1 aspect ratio; IRDs support 16:9 and optionally 2.21:1 aspect ratio;
- H.264/AVC HDTV pictures have 16:9 aspect ratio ; IRDs support 16:9 aspect ratio;
- MPEG-2 IRDs support the use of pan vectors to allow a 4:3 monitor to give a full-screen display of a 16:9 coded picture of SDTV resolution;
- IRDs may also optionally support the use of the Active Format Description (refer to annex B of the present document) as part of the logic to control the processing and positioning of the reconstructed image for display.

Audio:

- Audio content complies with MPEG-1 Layer I, MPEG-1 Layer II or MPEG-2 Layer II backward compatible audio or annexes C, F or H;

- Sampling rates of 32 kHz, 44,1 kHz and 48 kHz are supported by IRDs;
- The encoded bit-stream does not use emphasis;
- IRDs may also optionally support full multi-channel decoding of MPEG-2 Layer II backwards compatible multi-channel audio;
- The use of Layer II encoding is recommended for MPEG-1 audio bit-streams;
- IRDs may also optionally support the decoding of MPEG audio streams which include ancillary data (see annex D);
- IRDs may also optionally support receiver-mixed audio (see annex G).

1 Scope

The present document provides implementation guidelines for the use of audio-visual coding in satellite, cable and terrestrial broadcasting distribution systems that utilize MPEG-2 Systems. Both Standard Definition Television (SDTV) and High Definition Television (HDTV) are covered. Both MPEG-2 video and H.264/AVC video coding systems are covered. MPEG-1/MPEG-2 Layer II, Dolby AC-3, Enhanced AC-3, DTS, MPEG-4 HE AAC and MPEG-4 HE AAC v2 audio coding systems are covered. Guidelines for devices equipped with a digital interface intended for digital VCR applications are also given in the present document. It does not cover applications such as contribution services which are likely to be the subject of subsequent "Guidelines" documents.

The rules of operation for the encoders are features and constraints which the encoding system should adhere to in order to ensure that the transmissions can be correctly decoded. These constraints may be mandatory, recommended or optional.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication and/or edition number or version number) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

[1] ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 (2000): "Information Technology - Generic Coding of moving pictures and associated audio information: Systems"- and ITU-T Recommendation H.222.0 | Amendment 3 (2004)/ISO/IEC 13818-1: 2000/Amendment 3:(2004): "Transport of AVC video data over ITU-T Recommendation H.222.0 |ISO/IEC 13818-1 streams".

[2] ITU-T Recommendation H.262 (2000)/ISO/IEC 13818-2 (2000): "Information Technology - Generic Coding of moving pictures and associated audio information: Video".

[3] ISO/IEC 13818-3 (1998): "Information technology - Generic coding of moving picture and associated audio information; Part 3: Audio".

[4] ISO/IEC 13818-9 (1996): "Information technology - Generic coding of moving pictures and associated audio information - Part 9: Extension for real time interface for systems decoders".

[5] ETSI ETR 289: "Digital Video Broadcasting (DVB); Support for use of scrambling and Conditional Access (CA) within digital broadcasting systems".

[6] ETSI EN 300 468: "Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems".

[7] ETSI ETR 211: "Digital Video Broadcasting (DVB); Guidelines on implementation and usage of Service Information (SI)".

[8] ISO/IEC 11172-1: "Information Technology - Coding of moving pictures and associated audio for digital storage media up to about 1,5 Mbit/s - Part 1: Systems".

[9] ISO/IEC 11172-3 (1993): "Information Technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 3: Audio".

[10] ITU-T Recommendation J.17 (1988): "Pre-emphasis used on sound-programme circuits".

[11] EBU Recommendation R.68: "Alignment level in digital audio production equipment and in digital audio recorders".

[12] ETSI TS 102 366: "Digital Audio Compression (AC-3, Enhanced AC-3) Standard".

[13] ITU-R Recommendation BT.709: "Parameter values for the HDTV standards for production and international programme exchange".

[14] ETSI EN 300 294: "Television systems; 625-line television Wide Screen Signalling (WSS)".

[15] ETSI TS 102 114: "DTS Coherent Acoustics; Core and Extensions".

[16] ITU-T Recommendation H.264 (2003): "Advanced Video Coding for Generic Audiovisual Services" and ISO/IEC 14496-10 (2004): "Information technology - Coding of audio-visual objects- Part 10: Advanced Video Coding".

[17] ISO/IEC 14496-3: "Information technology -- Coding of audio-visual objects - Part 3: Audio, including amendment 1: "Bandwidth Extension" and amendment 2: "Parametric Coding for High Quality Audio".

[18] ISO/IEC 13818-1 (2000) Amd 5 (FPDAM): "New audio profile and level signalling and change to audio-type table entry".

[19] ETSI EN 300 401 (V.1.3.3): "Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers".

[20] ITU-T Recommendation T.35 (2000): "Procedure for the allocation of ITU-T defined codes for non-standard facilities".

[21] ISO/IEC 14496-10 (2004) amendment 1: "AVC professional extensions".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

25 Hz MPEG-2 SDTV IRD: IRD which is capable of decoding and displaying pictures based on a nominal video frame rate of 25 Hz from MPEG-2 Main Profile, Main Level bitstreams as specified in TS 101 154

25 Hz MPEG-2 SDTV Bitstream: bitstream which contains only MPEG-2 Main Profile, Main Level video at 25 Hz frame rate as specified in TS 101 154

25 Hz MPEG-2 HDTV IRD: IRD that is capable of decoding and displaying pictures based on a nominal video frame rate of 25 Hz or 50 Hz from MPEG-2 Main Profile, High Level bitstreams as specified in TS 101 154, in addition to providing the functionality of a 25 Hz SDTV IRD

25 Hz MPEG-2 HDTV Bitstream: bitstream which contains only MPEG-2 Main Profile, High Level (or simpler) video at 25 Hz or 50 Hz frame rates as specified in TS 101 154

30 Hz MPEG-2 SDTV IRD: IRD which is capable of decoding and displaying pictures based on a nominal video frame rate of 24 000/1001(approximately 23.98), 24, 30000/1001 (approximately 29,97) or 30 Hz from MPEG-2 Main Profile at Main Level bitstreams as specified in TS 101 154

30 Hz MPEG-2 SDTV Bitstream: bitstream which contains only MPEG-2 Main Profile, Main Level video at 24 000/1001, 24, 30000/1001 or 30 Hz frame rate as specified in TS 101 154

30 Hz MPEG-2 HDTV IRD: IRD that is capable of decoding and displaying pictures based on nominal video frame rates of 24 000/1001, 24, 30000/1001, 30, 60/1001 or 60 Hz from MPEG-2 Main Profile, High Level bitstreams as specified in TS 101 154, in addition to providing the functionality of a 30 Hz SDTV IRD

30 Hz MPEG-2 HDTV Bitstream: bitstream which contains only MPEG-2 Main Profile, High Level (or simpler) video at 24 000/1001, 24, 30000/1001, 30, 60/1001 or 60 Hz frame rates as specified in TS 101 154

MPEG-2 IRD: a collective term referring to the 25 Hz MPEG-2 SDTV IRD, 30 Hz MPEG-2 SDTV IRD, 25 Hz MPEG-2 HDTV IRD, 30 Hz MPEG-2 HDTV IRD

MPEG-2 Bitstream: a collective term referring to the 25 Hz MPEG-2 SDTV Bitstream, 30 Hz MPEG-2 SDTV Bitstream, 25 Hz MPEG-2 HDTV Bitstream, 30 Hz MPEG-2 HDTV Bitstream

25 Hz H.264/AVC SDTV IRD: IRD which is capable of decoding and displaying pictures based on a nominal video frame rate of 25 Hz from H.264/AVC Main Profile at Level 3 bitstreams as specified in TS 101 154

25 Hz H.264/AVC SDTV Bitstream: bitstream which contains only H.264/AVC Main Profile at Level 3 video at 25 Hz frame rate as specified in TS 101 154

25 Hz H.264/AVC HDTV IRD: IRD that is capable of decoding and displaying pictures based on a nominal video frame rate of 25 Hz or 50 Hz from H.264/AVC High Profile at Level 4 bitstreams as specified in TS 101 154, in addition to providing the functionality of a 25 Hz H.264/AVC SDTV IRD

25 Hz H.264/AVC HDTV Bitstream: bitstream which contains only H.264/AVC High Profile at Level 4 (or simpler) video at 25 Hz or 50 Hz frame rates as specified in TS 101 154

30 Hz H.264/AVC SDTV IRD: IRD which is capable of decoding and displaying pictures based on a nominal video frame rate of 24 000/1001(approximately 23.98), 24, 30000/1001 (approximately 29,97) or 30 Hz from H.264/AVC Main Profile at Level 3 bitstreams as specified in TS 101 154

30 Hz H.264/AVC SDTV Bitstream: bitstream which contains only H.264/AVC Main Profile at Level 3 video at 24 000/1001, 24, 30000/1001 or 30 Hz frame rate as specified in TS 101 154

30 Hz H.264/AVC HDTV IRD: IRD that is capable of decoding and displaying pictures based on nominal video frame rates of 24 000/1001, 24, 30000/1001, 30, 60/1001 or 60 Hz from H.264/AVC High Profile at Level 4 bitstreams as specified in TS 101 154, in addition to providing the functionality of a 30 Hz SDTV IRD

30 Hz H.264/AVC HDTV Bitstream: bitstream which contains only H.264/AVC High Profile at Level 4 (or simpler) video at 24 000/1001, 24, 30000/1001, 30, 60/1001 or 60 Hz frame rates as specified in TS 101 154

H.264/AVC SDTV IRD: collective term referring to the 25 Hz H.264/AVC SDTV IRD and the 30 Hz H.264/AVC SDTV IRD

H.264/AVC SDTV Bitstream: collective term referring to the 25 Hz H.264/AVC SDTV Bitstream and the 30 Hz H.264/AVC SDTV Bitstream

H.264/AVC HDTV IRD: collective term referring to the 25 Hz H.264/AVC HDTV IRD and the 30 Hz H.264/AVC HDTV IRD

H.264/AVC HDTV Bitstream: collective term referring to the 25 Hz H.264/AVC HDTV Bitstream and the 30 Hz H.264/AVC HDTV Bitstream

H.264/AVC IRD: collective term referring to the H.264/AVC SDTV IRD and the H.264/AVC HDTV IRD

H.264/AVC Bitstream: collective term referring to the H.264/AVC SDTV Bitstream and the H.264/AVC HDTV Bitstream

I picture: picture (frame or field) containing only intra macroblocks

Baseline IRD: IRD which provides the minimum functionality to decode transmitted bitstreams as recommended in TS 101 154. It is not required to have the ability to decode Partial Transport Streams as may be received from a digital interface connected to digital bitstream storage device such as a digital VCR

IRD with Digital Interface: IRD which has the ability to decode Partial Transport Streams received from a digital interface connected to digital bitstream storage device such as a digital VCR as specified in TS 101 154, in addition to providing the functionality of a Baseline IRD

Pan Vector: horizontal offset in video frame centre position specified by non zero value in the frame_centre_horizontal _offset field in the MPEG video stream

Partial Transport Stream: bitstream derived from an MPEG-2 Transport Stream by removing those Transport Stream Packets that are not relevant to one particular selected programme, or a number of selected programmes

H.264/AVC RAP: access unit with AU delimiter in an H.264/AVC Bitstream at which an IRD can begin decoding video successfully. This access unit must contain one Sequence Parameter Set NAL unit and one Picture Parameter Set NAL unit that are active or being activated when decoding the primary coded picture in this access unit. This access unit must contain an IDR picture or an I picture

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AAC	Advanced Audio Coding according to ISO/IEC 14496-3 [17]
AC-3	Dolby AC-3 audio coding system according to TS 102 366 [12]
AFD	Active Format Description
AOT	Audio Object Type
AVC	Advanced Video Coding
CA	Conditional Access
DAB	Digital Audio Broadcasting
DTS	DTS audio coding system according to TS 102 114 [15]
DVB	Digital Video Broadcasting
DVD	Digital Versatile Disc
ES	Elementary Stream
ESCR	Elementary Stream Clock Reference
H.264/AVC	Advanced Video Coding for Generic Audiovisual Services according to ITU-T Recommendation H.264 [16]
HDTV	High Definition Television
HE AAC	High-Efficiency Advanced Audio Coding according to ISO/IEC 14496-3 [17]
IDR	Instantaneous Decoding Refresh
I-frame	Intra-coded frame
IRD	Integrated Receiver-Decoder
LATM	Low overhead Audio Transport Multiplex
MPEG	Moving Pictures Experts Group
NIT	Network Information Table
PAT	Program Association Table
PCR	Program Clock Reference
PES	Packetized Elementary Stream
PID	Packet IDentifier
PMT	Program Map Table
PS	Parametric Stereo
PSI	Program Specific Information
RAP	Random Access Point
SBR	Spectral Band Replication
ScF-CRC	Scale Factor Cyclic Redundancy Check
SDTV	Standard Definition Television
SEI	Supplemental Enhancement Information
SI	Service Information
STD	System Target Decoder
TS	Transport Stream
TSDT	Transport Stream Description Table
T-STD	Transport stream System Target Decoder
VCR	Video Cassette Recorder
VUI	Video Usability Information

4 Systems layer

This clause describes the guidelines for encoding the systems layer of MPEG-2 in DVB broadcast bit-streams, and for decoding this layer in the IRD. The source bitstream may be transmitted via a satellite, cable or terrestrial channel, or via a digital interface. Clause 4.1 applies to the encoding of all source bitstreams and their decoding by a Baseline IRD. Clause 4.2 gives specific information relating to bitstreams transmitted via a digital interface intended for VCR applications and decoding by IRDs equipped with such an interface.

4.1 Broadcast bitstreams and Baseline IRDs

The multiplexing of baseband signals and associated data conforms to ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1]. Some of the parameters and fields are not used in the DVB System and these restrictions are described below.

To allow full compliance to ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1] and upward compatibility with future enhanced versions, a DVB IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the IRD. As an example of this capability, a descriptor tag not yet defined within the DVB System shall be interpreted as a no-action tag, its length field correctly decoded and subsequent data skipped.

For the same reason, IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1] may occur in the broadcast stream even if presently reserved or unused. Therefore the following is assumed:

- private data shall only be acted upon by decoders which are so enabled;
- filling out the bit-stream shall be carried out using the normal stuffing mechanism. Reserved fields shall not be used for this purpose. Data of reserved fields shall be set to 0xFF.

The headings in this clause are based on ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1]. The numbers in brackets after the headings are the relevant chapter and clause headings of ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].

4.1.1 Introduction (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 Introduction)

MPEG-2 systems specify two types of multiplexed data stream: the transport stream and the program stream.

Encoding: *The transmitted multiplex shall use the transport stream.*

Decoding: *All Baseline IRDs shall be able to demultiplex the MPEG-2 transport stream.* Demultiplexing of program streams (as described in clauses Intro .2 and Intro .3 of ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1]) is optional.

4.1.2 Packetized Elementary Stream (PES) (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause Intro .4)

Encoding: The creation of a physical Packetized Elementary Stream (PES) by an encoder is not required. ESCR fields and ES rate fields need not be coded.

Decoding: ESCR fields and ES rate fields need not be decoded.

4.1.3 Transport stream system target decoder (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause 2.4.2)

Encoding: *The system clock frequency shall conform to the tolerance specified in clause 2.4.2.1 of ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].* It is recommended that the tolerance is within 5 parts per million.

Decoding: *The IRD shall operate over the full tolerance range of the system clock frequency specified in clause 2.4.2.1 of ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

4.1.4 Transport packet layer (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause 2.4.3.2)

4.1.4.1 Null packets

Encoding: *The encoding of null packets (those with PID value 0x1FFF) shall be as specified in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

4.1.4.2 Transport packet header

4.1.4.2.1 transport_error_indicator

Encoding: It is recommended that any error detecting devices in a transmission path should set the **transport_error_indicator** bit when uncorrecTable errors are detected.

Decoding: Whenever the **transport_error_indicator** flag is set in the transmitted stream it is recommended that the IRD should then invoke a suitable concealment or error recovery mechanism.

4.1.4.2.2 transport_priority

Decoding: The **transport_priority** bit has no meaning to the IRD, and may be ignored.

4.1.4.2.3 transport_scrambling_control

Encoding: *The transport_scrambling_control bits shall be set according to table 1, in accordance with ETR 289 [5].*

Table 1: Coding of transport_scrambling_control bits

Value	Description
00	no scrambling of TS packet payload
01	reserved for future DVB use
10	TS packet scrambled with Even key
11	TS packet scrambled with Odd key

Decoding: *These bits shall be read by the IRD, and the IRD shall respond in accordance with table 1.*

4.1.4.2.4 Packet IDentifier (PID) values for Service Information (SI) Tables

Encoding: The assignment of PID values for SI data is given in EN 300 468 [6].

4.1.5 Adaptation field (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause 2.4.3.4)

4.1.5.1 Random_access_indicator

For MPEG-2 Bitstreams, the following applies.

Encoding: It is recommended that the **random_access_indicator** bit is set whenever a random access point occurs in video streams (i.e. video sequence header immediately followed by an I-frame).

For H.264/AVC Bitstreams, the following applies.

Encoding: *The **random_access_indicator** bit shall be set whenever an H.264/AVC RAP occurs in video streams* (see H.264/AVC RAP definition in clauses 3.1 and 5.5.5).

Decoding: The **random_access_indicator** bit may be ignored by the IRD. It can be beneficially utilized together with the **elementary_stream_priority** indicator to identify RAP.

4.1.5.2 Elementary_stream_priority_indicator

For MPEG-2 IRDs, the following applies:

Decoding: The **elementary_stream_priority_indicator** bit may be ignored by the IRD.

For H.264/AVC Bitstreams, the following applies:

Encoding: *The **elementary_stream_priority_indicator** bit shall be set whenever an access unit containing an I picture is present in H264/AVC video streams.*

NOTE: The **elementary_stream_priority_indicator** shall be set in the adaptation header of the transport packet that contains the first slice start code of this I picture (per ISO/IEC 13818-1 [1]). This adaptation header may be in the transport packet after the packet containing the **random_access_indicator**.

Decoding: The **elementary_stream_priority_indicator** bit may be ignored by the IRD. It can be beneficially utilized to support trick modes.

4.1.5.3 Program Clock Reference (PCR)

Encoding: *The time interval between two consecutive PCR values of the same program shall not exceed 100 ms as specified in clause 2.7.2 of ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

Decoding: *The IRD shall operate correctly with PCRs for a program arriving at intervals not exceeding 100 ms.*

4.1.5.4 Other fields

This clause covers the following fields:

- original_program_clock_reference_base;
- original_program_clock_reference_extension;
- splice_countdown;
- private_data_byte;
- adaptation_field_extension (including fields within).

Encoding: These fields are optional in a DVB bit-stream. *The flags that indicate the presence or absence of each of these fields shall be set appropriately.*

Decoding: *IRDs shall be able to accept bit-streams which contain these fields.* IRDs may ignore the data within the fields.

4.1.6 Packetized Elementary Stream (PES) Packet (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause 2.4.3.6)

4.1.6.1 stream_id and stream_type

Encoding: *Elementary streams shall be identified by stream_id and stream_type in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], tables 2-18 and 2-29.*

4.1.6.2 PES_scrambling_control

Encoding: *The **PES_scrambling_control** bits shall be set according to table 2, in accordance with ETR 289 [5].*

Table 2: Coding of PES_scrambling_control bits

Value	Description
00	no scrambling of PES packet payload
01	reserved for future DVB use
10	PES packet scrambled with Even key
11	PES packet scrambled with Odd key

Decoding: *The **PES_scrambling_control** bits shall be read by the IRD, and the IRD shall respond in accordance with table 2.*

4.1.6.3 PES_priority

Decoding: The **PES_priority** bit may be ignored by the IRD.

4.1.6.4 Copyright and original_or_copy

Encoding: The copyright and **original_or_copy** bits may be set as appropriate.

Decoding: The IRD need not interpret these bits. *The setting of these bits shall not be altered in any digital output from the IRD.*

4.1.6.5 Trick mode fields

This clause covers the following fields:

- trick_mode_control;
- field_id;
- intra_slice_refresh;
- frequency_truncation;
- field_rep_cntrl.

Encoding: *These trick mode fields shall not be transmitted in a broadcast bit-stream.* Bit-streams for other applications (e.g. for non-broadcast interactive services, storage applications, etc.) may use these fields.

Decoding: The IRD may skip over any data which is flagged as being in a trick mode, if it does not support decoding of trick modes. If the IRD has a digital interface intended for digital VCR applications, it is recommended that it supports decoding of trick modes as indicated in clause 2.2.

4.1.6.6 additional_copy_info

Encoding: This field may be used as appropriate.

Decoding: The IRD need not interpret this field. *The coding of the field shall not be altered in any digital output from the IRD.*

4.1.6.7 Optional fields

This clause covers the following fields:

- ESCR;
- ESCR_extension;
- ES_rate;
- previous_PES_packet_CRC;
- PES_private_data;
- pack_header();
- program_packet_sequence_counter;
- MPEG1_MPEG2_identifier;
- original_stuff_length;
- P-STD_buffer_scale;
- P-STD_buffer_size.

Encoding: These fields are optional in a DVB bit-stream. *The flags that indicate the presence or absence of each of these fields shall be set appropriately.*

Decoding: *The IRD shall be able to accept bit-streams which contain these fields.* The IRD may ignore the data within the fields.

4.1.6.8 PES_extension_field

The **PES_extension_field** data field is currently "reserved".

Encoding: *This extension field shall not be coded unless specified in the future by MPEG.*

Decoding: *The IRD shall be able to accept bit-streams which contain this field.* The IRD may ignore the data within the field.

4.1.6.9 Multiple video pictures per PES packet

For MPEG-2 bitstreams, while there is no restriction against multiple video pictures in a single PES packet, there may be some MPEG-2 decoders that do not support this.

Encoding: The encoder should not put multiple video pictures in a single PES packet.

Decoding: The IRD may be able to accept and decode bit-streams which contain multiple video pictures in a single PES.

For H.264/AVC bitstreams, multiple video pictures are allowed in a single PES packet.

Encoding: *A PES packet per access unit start shall be sent* unless if multiple access units can be placed in a single transport packet. In this last case, the encoder may put multiple complete access units in a single PES packet. In applications where the IRD is capable of decoding and displaying bitstreams that contain fractions of access unit, the PES packet may contain fractions of Access Units and encoders are recommended to utilize this option for instance when bitrate savings can be achieved.

An access unit with H.264/AVC RAP shall be the first access unit in the PES packet (see clause 4.1.5.1) and shall always be preceded by a PES header. Changes to picture size or frame rate cannot occur between access units in the same PES packet. The maximum increment in PTS values between two successive PES packets shall be less than 700 ms with the exception case where video is coded using still pictures where the spacing shall be less than 5 seconds. **A single PES packet shall not contain multiple AVC Still pictures or multiple H.264/AVC RAPs.**

Decoding: *The IRD shall support decoding and displaying bitstreams, which contain multiple complete access units in a single PES packet.* It is strongly recommended that the IRD also supports decoding and displaying bitstreams that contain fractions of access units in PES packet.

4.1.6.10 Presentation Time Stamp and Decoding Time Stamp occurrence

For H.264/AVC

Encoding: *Every PES header shall contain the Presentation Time Stamp and the Decoding Time Stamp (only if it differs from the Presentation Time Stamp) of the first access unit in the PES packet. The start of the first access unit shall occur in the same transport packet as the PES header or the packet of same PID immediately following the packet with the PES header, if the data preceding the access unit start code forces the access unit start code into the next transport packet. When a PES packet contains multiple access units, for any access units following the first access unit in the same PES packet the H.264/AVC syntax elements num_units_in_tick, time_scale, pic_struct (if present), and the value of the H.264/AVC variables TopFieldOrderCnt and BottomFieldOrderCnt of the access unit shall allow the derivation of Presentation Time Stamp and the Decoding Time Stamp for the access unit.*

Decoding: *If Presentation Time Stamp is available and Decoding Time Stamp is not available for the first access unit in the ¨PES packet, the H.264/AVC IRD shall set the Decoding Time Stamp equal to the Presentation Time Stamp (per ISO/IEC 13818-1**).** The Presentation Time Stamp and the Decoding Time Stamp of any access units following the first access unit in the same PES packet shall be derived using the H.264/AVC syntax elements num_units_in_tick, time_scale, pic_struct (if present), and the value of the H.264/AVC variables TopFieldOrderCnt and BottomFieldOrderCnt of the access unit.*

4.1.7 Program Specific Information (PSI) (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause 2.4.4)

The data formats for the Transport Stream Description Table (TSDT) and Network Information Table (NIT) in DVB bit-streams are given in EN 300 468 [6]. The present document also defines additional tables for service information which use Program Specific Information (PSI) private_section structure defined in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].

It is recommended that the Program Association Table (PAT) and Program Map Table (PMT) are repeated with a maximum time interval of 100 ms between repetitions. It is recommended that the Transport Stream Description Table (TSDT) is repeated with a maximum time interval of 10 seconds between repetitions.

4.1.8 Program and elementary stream descriptors (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause 2.6)

4.1.8.1 video_stream_descriptor and audio_stream_descriptor

Encoding: *The **video_stream_descriptor** shall be used to indicate video streams containing still picture data, otherwise these descriptors may be used when appropriate. If **profile_and_level_indication** is not present, then the video bit-stream shall comply with the constraints of Main Profile at Main Level. The appropriate **profile_and_level_indication field** shall always be transmitted for Profiles and Levels other than Main Profile at Main Level.*

If the **audio_stream_descriptor** is not present, then the audio bit-stream shall not use sampling frequencies of 16 kHz, 22,05 kHz or 24 kHz, and all audio frames in the stream shall have the same bit rate.

Decoding: The IRD may use these descriptors when present to determine if it is able to decode the streams.

4.1.8.2 hierarchy_descriptor

Encoding: *The **hierarchy_descriptor** shall be used if, and only if, audio is coded as more than one hierarchical layer.*

4.1.8.3 registration_descriptor

Encoding: The **registration_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.4 data_stream_alignment_descriptor

Encoding: The **data_stream_alignment_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.5 target_background_grid_descriptor

Encoding: *The **target_background_grid_descriptor** shall be used when the horizontal or vertical resolution is other than 720 × 576 pixels for a 25 Hz bitstream or is other than 720 × 480 pixels for a 30 Hz bitstream, otherwise its use is optional.*

Decoding: *If this descriptor is absent, a default grid of 720 × 576 pixels shall be assumed by a 25 Hz IRD, a default grid of 720 ×480 pixels shall be assumed by a 30 Hz IRD.* The display of correctly windowed video on background grids other than 720 × 576 pixels is optional for a 25 Hz SDTV IRD, the display of correctly windowed video on background grids other than 720 × 480 pixels is optional for a 30 Hz SDTV IRD. *The HDTV IRD shall read this descriptor, when present, to override the default values.*

4.1.8.6 video_window_descriptor

Encoding: The **video_window_descriptor** may be used when appropriate, to indicate the required position of the video window on the screen.

Decoding: *The IRD shall read this descriptor, when present, and position the video window accordingly.*

4.1.8.7 Conditional Access CA_descriptor

Encoding: *The **CA_descriptor** shall be encoded as defined in ETR 289 [5].*

Decoding: *The IRD shall interpret this descriptor as defined in ETR 289 [5].*

4.1.8.8 ISO_639_Language_descriptor

Encoding: *The **ISO_639_Language_descriptor** shall be present if more than one audio (or video) stream with different languages is present within a program.* It is optional otherwise. The use of the ISO_639_Language_descriptor is recommended for all audio, video and data streams.

Decoding: *The IRD shall use the data from this descriptor to assist the selection of appropriate audio (or video) stream of program, if more than one stream is available.*

4.1.8.9 system_clock_descriptor

Encoding: It is recommended that the **system_clock_descriptor** is included in the program_info part of the Program Map Table for each program.

Decoding: The IRD need not make use of this descriptor.

4.1.8.10 multiplex_buffer_utilization_descriptor

Encoding: The **multiplex_buffer_utilization_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.11 copyright_descriptor

Encoding: The **copyright_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.12 maximum_bitrate_descriptor

Encoding: The **maximum_bitrate_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.13 private_data_indicator_descriptor

Encoding: The **private_data_indicator_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.14 smoothing_buffer_descriptor

Encoding: It is recommended that the **smoothing_buffer_descriptor** is included in the program_info part of the Program Map Table for each program.

Decoding: The IRD need not make use of this descriptor, but the information may be of assistance to digital VCRs.

4.1.8.15 STD_descriptor

Encoding: *The **STD_descriptor** shall be used as specified in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

Decoding: The IRD need not make use of this descriptor.

4.1.8.16 IBP_descriptor

Encoding: The **IBP_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.17 MPEG-4_video_descriptor

Encoding: The **MPEG-4_video_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.18 MPEG-4_audio_descriptor

Encoding: The **MPEG-4_audio_descriptor** may be used when appropriate.

Decoding: The IRD need not make use of this descriptor.

4.1.8.19 AVC_video_descriptor

Encoding: The **AVC_video_descriptor** may be used when appropriate. *The **AVC_video_descriptor** shall be used to signal presence of AVC still pictures within the coded video sequence (see clause 5.5.4.3).*

Decoding: The IRD need not make use of this descriptor. However, the information may assist in support for AVC still pictures (see clause 5.5.4.3).

4.1.8.20 Descriptors related to ISO/IEC 14496-10

This clause covers the following descriptors:

- IOD_descriptor;
- SL_descriptor;
- FMC_descriptor;
- External_ES_ID_descriptor;
- MuxCode_descriptor;
- FmxBufferSize_descriptor;
- MultiplexBuffer_descriptor.

Encoding: These descriptors may be used when appropriate.

Decoding: The IRD need not make use of these descriptors.

Additional descriptors to those defined in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1] are defined in EN 300 468 [6], and guidelines for their use are provided in ETR 211 [7].

4.1.9 Compatibility with ISO/IEC 11172-1 (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause 2.8)

Decoding: Compatibility with ISO/IEC 11172-1 [8] (MPEG-1 Systems) is optional.

4.1.10 Storage Media Interoperability

It is recommended that the total bitrate of the set of components, associated PMT and PCR packets for an SDTV service anticipated to be recorded by a consumer, should not exceed 9 000 000 bit/s. It is recommended that the total bitrate of the set of components, associated PMT and PCR packets for an HDTV service anticipated to be recorded by a consumer, should not exceed 28 000 000 bit/s.

It is recommended that the parameters sb_size and sb_leak_rate in the smoothing_buffer_descriptor remain constant for the duration of an event. The value of the sb_leak_rate should be the peak attained during the event. The short_smoothing_buffer_descriptor is defined in EN 300 468 [6] and guidelines for its use are provided in ETR 211 [7].

4.2 Bitstreams from storage applications and IRDs with digital interfaces

This clause covers both the treatment of Partial Transport Streams which result from external program selection and Trick Play information received from a storage device. MPEG-2 PSI and DVB SI Tables for use specifically in storage applications are defined in EN 300 468 [6].

4.2.1 Partial Transport Streams

Partial transport streams for transfer on a digital interface, e.g. for digital VCR applications, have been defined in IEC CD - 100C/1883. A Partial Transport Stream may be created by selection of Transport Stream Packets from one or more program(s), including PSI Packets.

Encoding: *The Partial Transport Stream shall be fully MPEG compliant with reference to MPEG-2 "Extension for Real-Time-Interface for systems decoders" (ISO/IEC 13818-9 [4]).*

Decoding: *Devices equipped with a digital interface intended for digital VCR applications shall accept the bursty character of a Partial Transport Stream with gaps of variable length between the Transport Stream Packets.*

4.2.2 Decoding of Trick Play data (ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 clause 2.4.3.7)

Encoding: *Trick mode operation shall be signalled by use of the DSM_trick_mode flag in the header of the video Packetized Elementary Stream (PES) packets. During trick mode playback the storage device shall construct a bitstream which is syntactically and semantically correct, except as outlined in the note below.*

Decoding: It is recommended that devices decode the DSM_trick_mode_flag and the eight bit trick mode field. *Devices which decode the trick mode data shall follow the normative requirements detailed in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], 2 for all values of the trick_mode_control field.*

NOTE: Trick Mode Semantic Constraints.

The bitstream delivered to the decoder during trick mode shall comply with the syntax defined in the MPEG-2 standard. However, for the following video syntax elements, semantic exceptions apply in the presence of the DSM_trick_mode field:

- bit_rate;
- vbv_delay;
- repeat_first_field;
- v_axis_positive;
- field_sequence;
- subcarrier;
- burst_amplitude;
- subcarrier_phase.

A decoder cannot rely on the values encoded in these fields when in trick mode.

Similarly, for the systems layer, the following semantic exceptions apply in the presence of the DSM_trick_mode field:

- maximum spacing of PSI information may exceed 400 ms;
- maximum spacing of Presentation Time Stamp or Decoding Time Stamp occurrences may exceed 700 ms;
- PES packets may be void of video data to indicate a change in trick mode byte;
- a PES packet void of video data may contain a Presentation Time Stamp to indicate effective presentation time of new trick mode control;
- when trick_mode status is true, the elementary stream buffers in the T-STD may underflow.

5 Video

This clause describes the guidelines for encoding MPEG-2 video or H264/AVC video in DVB broadcast bit-streams, and for decoding this bit-stream in the IRD.

Clause 5.1 applies to 25 Hz MPEG-2 SDTV IRDs and broadcasts intended for reception by such IRDs.

Clause 5.2 applies to 25 Hz MPEG-2 HDTV IRDs and broadcasts intended for reception by such IRDs.

Clause 5.3 applies to 30 Hz MPEG-2 SDTV IRDs and broadcasts intended for reception by such IRDs.

Clause 5.4 applies to 30 Hz MPEG-2 HDTV IRDs and broadcasts intended for reception by such IRDs.

Clause 5.5 applies to 25 Hz H.264/AVC SDTV IRDs and broadcasts intended for reception by such IRDs.

Clause 5.6 applies to 25 Hz H.264/AVC HDTV IRDs and broadcasts intended for reception by such IRDs.

Clause 5.7 applies to 30 Hz H.264/AVC SDTV IRDs and broadcasts intended for reception by such IRDs.

Clause 5.8 applies to 30 Hz H.264/AVC HDTV IRDs and broadcasts intended for reception by such IRDs.

To allow full compliance to the MPEG-2 and H.264/AVC standard and upward compatibility with future enhanced versions, a DVB IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the IRD.

This clause is based on ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] and ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16].

5.1 25 Hz MPEG-2 SDTV IRDs and Bitstreams

The video encoding shall conform to ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]. Some of the parameters and fields are not used in the DVB System and these restrictions are described below. The IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] may occur in the broadcast stream even if presently reserved or unused.

5.1.1 Profile and level

Encoding: *Encoded bit-streams shall comply with the Main Profile Main Level restrictions, as described ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2], clause 8.2.* The **profile_and_level_indication** is "01001000" or, if appropriate, "0nnnnnnn", where "0nnnnnnn">"01001000", indicating a "simpler" profile or level than Main Profile, Main Level.

Decoding: *The 25 Hz MPEG-2 SDTV IRD shall support the decoding of Main Profile Main Level bitstreams.* Support for profiles and levels beyond Main Profile, Main Level is optional. *If the IRD encounters an extension which it cannot decode, such as one whose identification code is Reserved, Picture Sequence Scaleable, Picture Spatial Scaleable or Picture Temporal Scaleable, it shall discard the following data until the next start code (to allow backward compatible extensions to be added in the future).*

5.1.2 Frame rate

Encoding: *The frame rate shall be 25 Hz, i.e.* ***frame_rate_code*** *is "0011".*

Still pictures may be encoded by use of a video sequence consisting of a single intra-coded picture (see definition of still pictures in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.1.48).

Decoding: *All 25 Hz MPEG-2 SDTV IRDs shall support the decoding and display of video material with a frame rate of 25 Hz interlaced (i.e.* ***frame_rate_code*** *of "0011").* Support of other frame and field rates is optional.

25 Hz MPEG-2 SDTV IRDs shall be capable of decoding and displaying still pictures, i.e. video sequences consisting of a single intra-coded picture (see definition of still pictures in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.1.48).

5.1.3 Aspect ratio

Encoding: *The source aspect ratio in 25 Hz MPEG-2 SDTV bit-streams shall be either 4:3, 16:9 or 2.21:1.* Note that decoding of 2.21:1 aspect ratio is optional for the 25 Hz MPEG-2 SDTV IRD.

The **aspect_ratio_information** in the sequence header shall have one of the following three values:

- 4:3 aspect ratio source: "0010";
- 16:9 aspect ratio source: "0011";
- 2.21:1 aspect ratio source: "0100".

It is recommended that pan vectors for a 4:3 window are included in the transmitted bit-stream when the source aspect ratio is 16:9 or 2.21:1. *The vertical component of the transmitted pan vector shall be zero.*

If pan vectors are transmitted then the **sequence_display_extension** shall be present in the bit-stream and the **aspect_ratio_information** shall be set to ´0010´ (4:3 display). The display_vertical_size shall be equal to the **vertical_size**. *The* ***display_horizontal_size*** *shall contain the resolution of the target 4:3 display. The value of the* ***display_horizontal_size*** *field may be calculated by the following equation*:

$$\text{display\_horizontal\_size} = \frac{4}{3} \times \frac{\text{horizontal\_size}}{\text{source aspect ratio}}$$

Table 3 gives some typical examples:

Table 3: Values for display_horizontal_size

horizontal_size × vertical_size	Source aspect ratio	display_horizontal_size
720 × 576	16:9	540
544 × 576	16:9	408
480 × 576	16:9	360
352 × 576	16:9	264
352 × 288	16:9	264

Decoding: The 25 Hz MPEG-2 SDTV IRD *shall be able to decode bit-streams with values of* ***aspect_ratio_information*** *of "0010" and "0011", corresponding to 4:3 and 16:9 aspect ratio respectively.* If the IRD has a digital interface, this should be capable of outputting bit-streams with aspect ratios which are not directly supported by the IRD to allow their decoding and display via an external unit.

All 25 Hz MPEG-2 SDTV IRDs shall support the use of pan vectors and up sampling to allow a 4:3 monitor to give a full-screen display of a selected portion of a 16:9 coded picture with the correct aspect ratio. IRDs implementing the 2.21:1 aspect ratio should support the use of pan vectors and up sampling to allow a 4:3 monitor to give a full-screen display of a selected portion of the 2.21:1 picture with the correct aspect ratio. Support for pan vectors with non-zero vertical components is optional. *When no pan vectors are present in the transmitted bit-stream, the central portion of the wide-screen picture shall be displayed.* The support of vertical resampling to obtain the correct aspect ratio for a letterbox display of a 16:9 or 2.21:1 coded picture on a 4:3 monitor is optional.

5.1.4 Luminance resolution

Encoding: *The encoded picture shall have a full-screen luminance resolution (horizontal × vertical) of one of the following values:*

- 720 × 576;
- 544 × 576;
- 480 × 576;
- 352 × 576;
- 352 × 288.

In addition, non full-screen pictures may be encoded for display at less than full-size (when using one of the standard up-conversion ratios at the IRD).

Decoding: The 25 Hz MPEG-2 SDTV IRD *shall be capable of decoding pictures with luminance resolutions as shown in table 4 and applying up sampling to allow the decoded pictures to be displayed at full-screen size. In addition, IRDs shall be capable of decoding lower picture resolutions and displaying them at less than full-size after using one of the standard up-conversions, e.g. a horizontal resolution of 704 pixels within the 720 pixel full-screen display.*

Table 4: Resolutions for Full-screen Display from IRD

Coded Picture		**Displayed Picture Horizontal up sampling**	
Luminance resolution (horizontal × vertical)	**Aspect Ratio**	**4:3 Monitors**	**16:9 Monitors**
720 × 576	4:3 16:9 2.21:1	× 1 × 4/3 (see note 2) × 5/3 (see note 3)	× 3/4 (see note 1) × 1 × 5/4 (see note 4)
544 × 576	4:3 16:9 2.21:1	× 4/3 × 16/9 (see note 2) × 20/9 (see note 3)	× 1 (see note 1) × 4/3 × 5/3 (see note 4)
480 × 576	4:3 16:9 2.21:1	× 3/2 × 2 (see note 2) × 5/2 (see note 3)	× 9/8 (see note 1) × 3/2 × 15/8 (see note 4)
352 × 576	4:3 16:9 2.21:1	× 2 × 8/3 (see note 2) × 10/3 (see note 3)	× 3/2 (see note 1) × 2 × 5/2 (see note 4)
352 × 288	4:3 16:9 2.21:1	× 2 × 8/3 (see note 2) × 10/3 (see note 3) (and vertical up sampling × 2)	× 3/2 (see note 1) × 2 × 5/2 (see note 4) (and vertical up sampling × 2)

NOTE 1: Up sampling of 4:3 pictures for display on a 16:9 monitor is optional in the IRD, as 16:9 monitors can be switched to operate in 4:3 mode.
NOTE 2: The up sampling with this value is applied to the pixels of the 16:9 picture to be displayed on a 4:3 monitor.
NOTE 3: The up sampling with this value is applied to the pixels of the 2.21:1 picture to be displayed on a 4:3 monitor. Up sampling from 2.21:1 pictures for display on a 4:3 monitor is optional in the IRD.
NOTE 4: The up sampling with this value is applied to the pixels of the 2.21:1 picture to be displayed on a 16:9 monitor. Up sampling from 2.21:1 pictures for display on a 16:9 monitor is optional in the IRD.

5.1.5 Chromaticity Parameters

Encoding: It is recommended that the chromaticity co-ordinates of the ideal display, opto-electronic transfer characteristic of the ideal display and matrix coefficients used in deriving luminance and chrominance signals from the red, green and blue primaries be explicitly signalled in the encoded bitstream by setting the appropriate values for each of the following 3 parameters in the **sequence_display_extension()**: **colour_primaries**, **transfer_characteristics**, and **matrix_coefficients**.

*Within 25 Hz MPEG-2 SDTV bitstreams, if the **sequence_display_extension()** is not present in the bitstream or **colour_description** is zero, the chromaticity shall be implicitly defined to be that corresponding to **colour_primaries** having the value 5, the transfer characteristics shall be implicitly defined to be those corresponding to **transfer_characteristics** having the value 5 and the matrix coefficients shall be implicitly defined to be those corresponding **matrix_coefficients** having the value 5.* This set of parameter values corresponds signals compliance with ITU-R Recommendation BT.470-3 System B, G, I (see Bibliography).

5.1.6 Chrominance

Encoding: *The operation used to down sample the chrominance information from 4:2:2 to 4:2:0 shall be indicated by the parameter **chroma_420_type** in the picture coding extension.* A value of zero indicates that the fields have been down sampled independently. A value of one indicates that the two fields have been combined into a single frame before down sampling. It is desirable that the fields are down sampled independently (i.e. **chroma_420_type** = 0) to allow the IRD to use less memory for picture reconstruction.

Decoding: It is desirable that the operation used to up sample the chrominance information from 4:2:0 to 4:2:2 should be dependent on the parameter **chroma_420_type** in the picture coding extension.

5.1.7 Video sequence header

Encoding: It is recommended that a video sequence header, immediately followed by an I-frame, be encoded at least once every 500 ms. If quantizer matrices other than the default are used, the appropriate **intra_quantizer_matrix** and/or **non_intra_quantizer_matrix** are recommended to be included in every sequence header.

NOTE 1: Increasing the frequency of video sequence headers and I-frames will reduce channel hopping time but will reduce the efficiency of the video compression.

NOTE 2: Having a regular interval between I-frames may improve trick mode performance, but may reduce the efficiency of the video compression.

5.2 25 Hz MPEG-2 HDTV IRDs and Bitstreams

The video encoding shall conform to ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]. Some of the parameters and fields are not used in the DVB System and these restrictions are described below. The IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] may occur in the broadcast stream even if presently reserved or unused.

5.2.1 Profile and level

Encoding: *Encoded 25 Hz MPEG-2 HDTV bit-streams shall comply with the Main Profile High Level restrictions, as described ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2], clause 8.2.* The **profile_and_level_indication** is "01000100" or, if appropriate, "0nnnnnnn", where "0nnnnnnn">"01000100", indicating a "simpler" profile or level than Main Profile, High Level.

Decoding: *The 25 Hz MPEG-2 HDTV IRD shall support the decoding of Main Profile High Level bitstreams.* This requirement includes support for "simpler" profiles and levels, including Main Profile at Main Level, as defined in table 8-15 of ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]. Support for profiles and levels beyond Main Profile, High Level is optional. *If the IRD encounters an extension which it cannot decode, such as one whose identification code is Reserved, Picture Sequence Scaleable, Picture Spatial Scaleable or Picture Temporal Scaleable, it shall discard the following data until the next start code* (to allow backward compatible extensions to be added in the future*).*

5.2.2 Frame rate

Encoding: *The frame rate shall be 25 Hz or 50 Hz, i.e.* ***frame_rate_code*** *is "0011" or "0110".*

The source video format for 50 Hz frame rate material shall be progressive. The source video format for 25 Hz frame rate material may be interlaced or progressive.

Still pictures may be encoded by use of a video sequence consisting of a single intra-coded picture (see definition of still pictures in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.1.48).

Decoding: *All 25 Hz MPEG-2 HDTV IRDs shall support the decoding and display of video material with a frame rate of 25 Hz progressive, 25 Hz interlaced or 50 Hz progressive (i.e.* ***frame_rate_code*** *of "0011" or "0110") within the constraints of Main Profile at High Level.* Support of other frame and field rates is optional.

25 Hz MPEG-2 HDTV IRDs shall be capable of decoding and displaying still pictures, i.e. video sequences consisting of a single intra-coded picture (see definition of still pictures in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.1.48).

5.2.3 Aspect ratio

Encoding: *The source aspect ratio in 25 Hz MPEG-2 HDTV bit-streams shall be 16:9 or 2.21:1.* Note that decoding of 2.21:1 aspect ratio is optional for the 25 Hz MPEG-2 HDTV IRD.

The ***aspect_ratio_information*** *in the sequence header shall have the value "0011" or "0100".*

Decoding: *The 25 Hz MPEG-2 HDTV IRD shall be able to decode bit-streams with* ***aspect_ratio_information*** *of value "0011", corresponding to 16:9 aspect ratio.* The support of the aspect ratio 2.21:1 is optional. If the IRD has a digital interface, this should be capable of outputting bit-streams with aspect ratios which are not directly supported by the IRD to allow their decoding and display via an external unit.

5.2.4 Luminance resolution

Encoding: *The encoded picture shall have a full-screen luminance resolution within the constraints set by Main Profile at High Level, i.e. it shall not have more than:*

- 1 088 lines per frame;
- 1 920 luminance samples per line;
- 62 668 800 luminance samples per second.

It is recommended that the source video for 25 Hz MPEG-2 HDTV Bitstreams has a luminance resolution of:

- 1 080 lines per frame;
- 1 920 luminance samples per line;
- with an associated frame rate of 25 Hz, with two interlaced fields per frame.

The source video may or may not be down-sampled prior to encoding.

The use of other encoded video resolutions within the constraints of Main Profile at High Level is also permitted. annex A of the present document provides examples of supported full-screen luminance resolutions. In addition, non full-screen pictures may be encoded for display at less than full-size.

NOTE 1: The limit of 62 668 800 luminance samples per second of Main Profile at High Level excludes the use of the maximum allowed picture resolution at 50 Hz frame rate.

NOTE 2: If the recommended source video format is encoded without down-sampling it gives 51 840 000 luminance samples per second and therefore falls within the allowed range for Main Profile at High Level.

Decoding: *The 25 Hz MPEG-2 HDTV IRD shall be capable of decoding and displaying pictures with luminance resolutions within the constraints set by Main Profile at High Level.*

5.2.5 Chromaticity Parameters

Encoding: *The chromaticity co-ordinates of the ideal display, opto-electronic transfer characteristic of the source picture and matrix coefficients used in deriving luminance and chrominance signals from the red, green and blue primaries shall be explicitly signalled in the encoded HDTV bitstream by setting the appropriate values for each of the following 3 parameters in the* ***sequence_display_extension()****:* ***colour_primaries****,* ***transfer_characteristics****, and* ***matrix_coefficients****.*

It is recommended that ITU-R Recommendation BT.709 [13] colorimetry is used in the 25 Hz HDTV bitstream, which is signalled by setting **colour_primaries** to the value 1, **transfer_characteristics** to the value 1 and **matrix_coefficients** to the value 1.

Decoding: *The 25 Hz MPEG-2 HDTV IRD shall be capable of decoding bitstreams with any allowed values of* ***colour_primaries****,* ***transfer_characterstics*** *and* ***matrix_coefficients****.* It is recommended that appropriate processing be included for the accurate representation of pictures using ITU-R Recommendation BT.709 [13] colorimetry.

5.2.6 Chrominance

Encoding: The operation used to down sample the chrominance information from 4:2:2 to 4:2:0 shall be indicated by the parameter **chroma_420_type** in the picture coding extension. A value of zero indicates that the fields have been down sampled independently. A value of one indicates that the two fields have been combined into a single frame before down sampling. It is desirable that the fields are down sampled independently (i.e. **chroma_420_type** = 0) to allow the IRD to use less memory for picture reconstruction.

Decoding: It is desirable that the operation used to up sample the chrominance information from 4:2:0 to 4:2:2 should be dependent on the parameter **chroma_420_type** in the picture coding extension.

5.2.7 Video sequence header

Encoding: It is recommended that a video sequence header, immediately followed by an I-frame, be encoded at least once every 500 ms. If quantizer matrices other than the default are used, the appropriate **intra_quantizer_matrix** and/or **non_intra_quantizer_matrix** are recommended to be included in every sequence header.

NOTE 1: Increasing the frequency of video sequence headers and I-frames will reduce channel hopping time but will reduce the efficiency of the video compression.

NOTE 2: Having a regular interval between I-frames may improve trick mode performance, but may reduce the efficiency of the video compression.

5.2.8 Backwards Compatibility

Decoding: *In addition to the above, a 25 Hz MPEG-2 HDTV IRD shall be capable of decoding any bitstream that a 25 Hz MPEG-2 SDTV IRD is required to decode, as described in clause 5.1.*

5.3 30 Hz MPEG-2 SDTV IRDs and Bitstreams

The video encoding shall conform to ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]. Some of the parameters and fields are not used in the DVB System and these restrictions are described below. The IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] may occur in the broadcast stream even if presently reserved or unused.

5.3.1 Profile and level

Encoding: *Encoded bit-streams shall comply with the Main Profile Main Level restrictions, as described ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2], clause 8.2.* The **profile_and_level_indication** is "01001000" or, if appropriate, "0nnnnnnn", where "0nnnnnnn">"01001000", indicating a "simpler" profile or level than Main Profile, Main Level.

Decoding: *The IRD shall support the syntax of Main Profile.* Support for profiles and levels beyond Main Profile, Main Level is optional. *If the IRD encounters an extension which it cannot decode, such as one whose identification code is Reserved, Picture Sequence Scaleable, Picture Spatial Scaleable or Picture Temporal Scaleable, it shall discard the following data until the next start code (to allow backward compatible extensions to be added in the future).*

5.3.2 Frame rate

Encoding: *The frame rate shall be either 24 000/1 001, 24, 30 000/1 001 or 30 Hz, i.e. the **frame_rate_code** field shall be encoded with one of the following values: "0001", "0010", "0100" or "0101".*

Still pictures may be encoded by use of a video sequence consisting of a single intra-coded picture (see definition of still pictures in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.1.48).

Decoding: *All 30 Hz SDTV IRDs shall support the decoding and display of Main Profile @ Main Level video with a frame rate of 24 000/1001, 24, 30 000/1 001 or 30 Hz.* Support of other frame rates is optional.

IRDs shall be capable of decoding and displaying still pictures, i.e. video sequences consisting of a single intra-coded picture (see definition of still pictures in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.1.48).

5.3.3 Aspect ratio

Encoding: *The source aspect ratio in 30 Hz MPEG-2 SDTV bit-streams shall be either 4:3, 16:9 or 2.21:1.* Note that decoding of 2.21:1 aspect ratio is optional for the 30 Hz SDTV IRD.

*The **aspect_ratio_information** in the sequence header shall have one of the following three values:*

- 4:3 aspect ratio source: "0010";
- 16:9 aspect ratio source: "0011";
- 2.21:1 aspect ratio source: "0100".

It is recommended that pan vectors for a 4:3 window are included in the transmitted bit-stream when the source aspect ratio is 16:9 or 2.21:1. *The vertical component of the transmitted pan vector shall be zero.*

*If pan vectors are transmitted then the **sequence_display_extension** shall be present in the bit-stream and the **aspect_ratio_information** shall be set to ´0010´ (4:3 display). The display_vertical_size shall be equal to the **vertical_size**. The **display_horizontal_size** shall contain the resolution of the target 4:3 display*. The value of the **display_horizontal_size** field may be calculated by the following equation:

$$\text{display\_horizontal\_size} = \frac{4}{3} \times \frac{\text{horizontal\_size}}{\text{source aspect ratio}}$$

Table 5 gives some typical examples:

Table 5: Values for display_horizontal_size

horizontal_size × vertical_size	Source aspect ratio	display_horizontal_size
720 × 480	16:9	540
640 × 480	16:9	480
544 × 480	16:9	408
480 × 480	16:9	360
352 × 480	16:9	264
352 × 240	16:9	264

Decoding: *The 30 Hz MPEG-2 SDTV IRD shall be able to decode bit-streams with values of **aspect_ratio_information** of "0010" and "0011", corresponding to 4:3 and 16:9 aspect ratio respectively.* If the IRD has a digital interface, this should be capable of outputting bit-streams with aspect ratios which are not directly supported by the IRD to allow their decoding and display via an external unit.

All 30 Hz MPEG-2 SDTV IRDs shall support the use of pan vectors and up sampling to allow a 4:3 monitor to give a full-screen display of a selected portion of a 16:9 coded picture with the correct aspect ratio. IRDs implementing the 2.21:1 aspect ratio should support the use of pan vectors and up sampling to allow a 4:3 monitor to give a full-screen display of a selected portion of the 2.21:1 picture with the correct aspect ratio. Support for pan vectors with non-zero vertical components is optional. When no pan vectors are present in the transmitted bit-stream, the central portion of the wide-screen picture shall be displayed. The support of vertical resampling to obtain the correct aspect ratio for a letterbox display of a 16:9 or 2.21:1 coded picture on a 4:3 monitor is optional.

5.3.4 Luminance resolution

Encoding: The encoded picture shall have a full-screen luminance resolution (horizontal × vertical) of one of the following values:

- 720 × 480;
- 640 × 480;
- 544 × 480;
- 480 × 480;
- 352 × 480;
- 352 × 240.

In addition, non full-screen pictures may be encoded for display at less than full-size (when using one of the standard up-conversion ratios at the IRD).

Decoding: *The 30 Hz MPEG-2 SDTV IRD shall be capable of decoding pictures with luminance resolutions as shown in table 6 and applying up sampling to allow the decoded pictures to be displayed at full-screen size. In addition, IRDs shall be capable of decoding lower picture resolutions and displaying them at less than full-size after using one of the standard up-conversions, e.g. a horizontal resolution of 704 pixels within the 720 pixel full-screen display.*

Table 6: Resolutions for Full-screen Display from IRD

Coded Picture		Displayed Picture Horizontal up sampling	
Luminance resolution (horizontal × vertical)	**Aspect Ratio**	**4:3 Monitors**	**16:9 Monitors**
720 × 480	4:3 16:9 2:21:1	× 1 × 4/3 (see note 2) × 5/3 (see note 3)	× 3/4 (see note 1) × 1 × 5/4 (see note 4)
640 × 480	4:3	× 9/8	× 27/32 (see note 1)
544 × 480	4:3 16:9 2:21:1	× 4/3 × 16/9 (see note 2) ×20/9 (see note 3)	× 1 (see note 1) × 4/3 × 5/3 (see note 4)
480 × 480	4:3 16:9 2:21:1	× 3/2 × 2 (see note 2) × 5/2 (see note 3)	× 9/8 (see note 1) × 3/2 × 15/8 (see note 4)
352 × 480	4:3 16:9 2:21:1	× 2 × 8/3 (see note 2) × 10/3 (see note 3)	× 3/2 (see note 1) × 2 × 5/2 (see note 4)
352 × 240	4:3 16:9 2:21:1	× 2 × 8/3 (see note 2) × 10/3 (see note 3) (and vertical up sampling × 2)	× 3/2 (see note 1) × 2 × 5/2 (see note 4) (and vertical up sampling × 2)

NOTE 1: Up sampling of 4:3 pictures for display on a 16:9 monitor is optional in the IRD, as 16:9 monitors can be switched to operate in 4:3 mode.
NOTE 2: The up sampling with this value is applied to the pixels of the 16:9 picture to be displayed on a 4:3 monitor.
NOTE 3: The up sampling with this value is applied to the pixels of the 2.21:1 picture to be displayed on a 4:3 monitor. Up sampling from 2.21:1 pictures for display on a 4:3 monitor is optional in the IRD.
NOTE 4: The up sampling with this value is applied to the pixels of the 2.21:1 picture to be displayed on a 16:9 monitor. Up sampling from 2.21:1 pictures for display on a 16:9 monitor is optional in the IRD.

5.3.5 Chromaticity Parameters

Encoding: It is recommended that the chromaticity co-ordinates of the ideal display, opto-electronic transfer characteristic of the ideal display and matrix coefficients used in deriving luminance and chrominance signals from the red, green and blue primaries be explicitly signalled in the encoded bitstream by setting the appropriate values for each of the following 3 parameters in the **sequence_display_extension()**: **colour_primaries**, **transfer_characteristics**, and **matrix_coefficients**.

*Within 30 Hz SDTV bitstreams, if the **sequence_display_extension()** is not present in the bitstream or **colour_description** is zero, the chromaticity shall be implicitly defined to be that corresponding to **colour_primaries** having the value 6, the transfer characteristics shall be implicitly defined to be those corresponding to **transfer_characterstics** having the value 6 and the matrix coefficients shall be implicitly defined to be those corresponding **matrix_coefficients** having the value 6.* This set of parameter values signals compliance with SMPTE 170M.

5.3.6 Chrominance

Encoding: *The operation used to down sample the chrominance information from 4:2:2 to 4:2:0 shall be indicated by the parameter **chroma_420_type** in the picture coding extension.* A value of zero indicates that the fields have been down sampled independently. A value of one indicates that the two fields have been combined into a single frame before down sampling. It is desirable that the fields are down sampled independently (i.e. **chroma_420_type** = 0) to allow the IRD to use less memory for picture reconstruction.

Decoding: It is desirable that the operation used to up sample the chrominance information from 4:2:0 to 4:2:2 should be dependent on the parameter **chroma_420_type** in the picture coding extension.

5.3.7 Video sequence header

Encoding: It is recommended that a video sequence header, immediately followed by an I-frame, be encoded at least once every 500 ms. If quantizer matrices other than the default are used, the appropriate **intra_quantizer_matrix** and/or **non_intra_quantizer_matrix** are recommended to be included in every sequence header.

NOTE 1: Increasing the frequency of video sequence headers and I-frames will reduce channel hopping time but will reduce the efficiency of the video compression.

NOTE 2: Having a regular interval between I-frames may improve trick mode performance, but may reduce the efficiency of the video compression.

5.4 30 Hz MPEG-2 HDTV IRDs and Bitstreams

The video encoding shall conform to ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]. Some of the parameters and fields are not used in the DVB System and these restrictions are described below. The IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] may occur in the broadcast stream even if presently reserved or unused.

5.4.1 Profile and level

Encoding: *Encoded 30 Hz MPEG-2 HDTV bit-streams shall comply with the Main Profile High Level restrictions, as described ITU-T Recommendation H.262 | ISO/IEC 13818-2* [2]*, clause 8.2.*

The **profile_and_level_indication** is "01000100" or, if appropriate, "0nnnnnnn", where "0nnnnnnn">"01000100", indicating a "simpler" profile or level than Main Profile, High Level.

Decoding: *The 30 Hz MPEG-2 HDTV IRD shall support the decoding of Main Profile High Level bitstreams.* This requirement includes support for "simpler" profiles and levels, including Main Profile at Main Level, as defined in table 8-15 of ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]. Support for profiles and levels beyond Main Profile, High Level is optional. *If the IRD encounters an extension which it cannot decode, such as one whose identification code is Reserved, Picture Sequence Scaleable, Picture Spatial Scaleable or Picture Temporal Scaleable, it shall discard the following data until the next start code* (to allow backward compatible extensions to be added in the future).

5.4.2 Frame rate

Encoding: *The frame rate shall be 24 000/1 001, 24, 30 000/1 001, 30, 60 000/1 001 or 60 Hz, i.e. **frame_rate_code** is "0001", "0010", "0100", "0101", "0111" or "1000".*

The source video format for 24 000/1 001, 24, 60 000/1 001 and 60 Hz frame rate material shall be progressive. The source video format for 30 000/1 001 and 30 Hz frame rate material may be interlaced or progressive.

Still pictures may be encoded by use of a video sequence consisting of a single intra-coded picture (see definition of still pictures in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.1.48).

Decoding: *All 30 Hz MPEG-2 HDTV IRDs shall support the decoding of video material with a frame rate of 24 000/1 001, 24, 30 000/1 001, 30, 60 000/1 001 or 60 Hz (i.e.* ***frame_rate_code*** *of "0001", "0010", "0100", "0101", "0111" or "1000") within the constraints of Main Profile at High Level.* Support of other frame rates is optional.

30 Hz MPEG-2 HDTV IRDs shall support the display of video whose source frame rate is 24 000/1 001, 24, 30 000/1 001, 30, 60 000/1001 or 60 Hz progressive. 30 Hz MPEG-2 HDTV IRDs shall support the display of video whose source frame rate is 30000/1001 or 30 Hz interlaced.

30 Hz MPEG-2 HDTV IRDs shall be capable of decoding and displaying still pictures, i.e. video sequences consisting of a single intra-coded picture (see definition of still pictures in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.1.48).

5.4.3 Aspect ratio

Encoding: *The source aspect ratio in 30 Hz MPEG-2 HDTV bit-streams shall be 16:9 or 2.21:1.* Note that decoding of 2.21:1 aspect ratio is optional for the 30 Hz MPEG-2 HDTV IRD.

The ***aspect_ratio_information*** *field in the sequence header shall have the value "0011" or "0100".*

Decoding: *The 30 Hz MPEG-2 HDTV IRD shall be able to decode bit-streams with* ***aspect_ratio_information*** *of value "0011", corresponding to 16:9 aspect ratio.* If the IRD has a digital interface, this should be capable of outputting bit-streams with aspect ratios which are not directly supported by the IRD to allow their decoding and display via an external unit.

5.4.4 Luminance resolution

Encoding: *The encoded picture shall have a full-screen luminance resolution within the constraints set by Main Profile at High Level, i.e. it shall not have more than:*

- 1 088 lines per frame;
- 1 920 luminance samples per line;
- 62 668 800 luminance samples per second.

It is recommended that the source video for 30 Hz MPEG-2 HDTV Bitstreams has a luminance resolution of:

- 1 080 lines per frame and 1 920 luminance samples per line, with an associated frame rate of 30 000/1 001 (approximately 29.97) Hz with two interlaced fields per frame.
- The source video may or may not be down-sampled prior to encoding.
- The use of other encoded video resolutions within the constraints of Main Profile at High Level is also permitted. annex A of the present document provides examples of supported full-screen luminance resolutions. In addition, non full-screen pictures may be encoded for display at less than full-size.
- The limit of 62 668 800 luminance samples per second of Main Profile at High Level excludes the use of the maximum allowed picture resolution at 60 Hz and 60 000/1001 frame rates.

NOTE: If the recommended source video format is encoded without down-sampling it gives 62 145 854 luminance sample per second and therefore falls within the allowed range for Main Profile at High Level.

Decoding: *The 30 Hz MPEG-2 HDTV IRD shall be capable of decoding and displaying pictures with luminance resolutions within the constraints set by Main Profile at High Level.*

5.4.5 Chromaticity Parameters

Encoding: *The chromaticity co-ordinates of the ideal display, opto-electronic transfer characteristic of the source picture and matrix coefficients used in deriving luminance and chrominance signals from the red, green and blue primaries shall be explicitly signalled in the encoded HDTV bitstream by setting the appropriate values for each of the following 3 parameters in the* ***sequence_display_extension()****:* ***colour_primaries****,* ***transfer_characteristics****, and* ***matrix_coefficients****.*

It is recommended that ITU-R Recommendation BT.709 [13] colorimetry is used in the 30 Hz HDTV bitstream, which is signalled by setting **colour_primaries** to the value 1, **transfer_characteristics** to the value 1 and **matrix_coefficients** to the value 1.

Decoding: *The 30 Hz HDTV IRD shall be capable of decoding bitstreams with any allowed values of* ***colour_primaries****,* ***transfer_characteristics*** *and* ***matrix_coefficients****.* It is recommended that appropriate processing be included for the accurate representation of pictures using ITU-R Recommendation BT.709 [13] colorimetry.

5.4.6 Chrominance

Encoding: *The operation used to down sample the chrominance information from 4:2:2 to 4:2:0 shall be indicated by the parameter* ***chroma_420_type*** *in the picture coding extension.* A value of zero indicates that the fields have been down sampled independently. A value of one indicates that the two fields have been combined into a single frame before down sampling. It is desirable that the fields are down sampled independently (i.e. **chroma_420_type** = 0) to allow the IRD to use less memory for picture reconstruction.

Decoding: It is desirable that the operation used to up sample the chrominance information from 4:2:0 to 4:2:2 should be dependent on the parameter **chroma_420_type** in the picture coding extension.

5.4.7 Video sequence header

Encoding: It is recommended that a video sequence header, immediately followed by an I-frame, be encoded at least once every 500 ms. If quantizer matrices other than the default are used, the appropriate **intra_quantizer_matrix** and/or **non_intra_quantizer_matrix** are recommended to be included in every sequence header.

NOTE 1: Increasing the frequency of video sequence headers and I-frames will reduce channel hopping time but will reduce the efficiency of the video compression.

NOTE 2: Having a regular interval between I-frames may improve trick mode performance, but may reduce the efficiency of the video compression.

5.4.8 Backwards Compatibility

Decoding: *In addition to the above, a 30 Hz MPEG-2 HDTV IRD shall be capable of decoding any bitstream that a 30 Hz MPEG-2 SDTV IRD is required to decode, as described in clause 5.3.*

5.5 Specifications Common to all H.264/AVC IRDs and Bitstreams

The specification in this clause applies to the following IRDs and Bitstreams:

- 25 Hz H.264/AVC SDTV IRD and Bitstream;
- 30 Hz H.264/AVC SDTV IRD and Bitstream;
- 25 Hz H.264/AVC HDTV IRD and Bitstream;
- 30 Hz H.264/AVC HDTV IRD and Bitstream.

5.5.1 General

*The video encoding and video decoding shall conform to ITU-T Recommendation H.264 | ISO/IEC 14496-10 [*16*]*. Some of the parameters and fields are not used in the DVB System and these restrictions are described below. *H.264/AVC Bitstreams and IRDs shall support some parts of the "Supplemental enhancement information (SEI)" and the "Video usability information (VUI)" syntax elements as specified in ITU-T Recommendation H.264 | ISO/IEC 14496-10 annexes D and E [*16*]*. The H.264/AVC IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16] and the restrictions that are specified for the H.264/AVC IRDs may occur in the broadcast stream even if presently reserved or unused.

NOTE: To improve trick mode it is strongly recommended to disable non-paired fields in H.264/AVC Encoder.

5.5.2 Sequence Parameter Set and Picture Parameter Set

Encoding: More than one picture parameter set can be present in the bitstream between two H.264/AVC RAPs. *Between two H.264/AVC RAPs, the content of a picture parameter set with a particular* ***pic_parameter_set_id*** *shall not change.* I.e. if more than one picture parameter set is present in the bitstream and these picture parameter sets are different from each other, then each picture parameter set shall have a different **pic_parameter_set_id**.

5.5.2.1 pic_width_in_mbs_minus1 and pic_height_in_map_units_minus1

Encoding: *The time interval between two changes in pairs of* ***pic_width_in_mbs_minus1*** *and* ***pic_height_in_map_units_minus1*** *shall be greater than or equal to one second.* Changing the pair **pic_width_in_mbs_minus1** and **pic_height_in_map_units_minus1** requires software processing in the decoder. Limiting the frequency of this change is to constrain the IRD software processing required to support aspect ratio changes.

NOTE: A pair of **pic_width_in_mbs_minus1** and **pic_height_in_map_units_minus1** is distinct from another pair if one or both syntax element values **pic_width_in_mbs_minus1** and **pic_height_in_map_units_minus1** differ.

If the number of samples per row of the luminance component of the source picture is not an integer multiple of 16 and additional samples are padded to make the number of samples per row of the luminance component an integer multiple of 16, it is recommended that these samples are padded at the right side of the picture.

If the number of samples per column of the luminance component of the source picture is not an integer multiple of 16 and additional samples are padded to make the number of samples per column of the luminance component an integer multiple of 16, it is recommended that these samples are padded at the bottom of the picture.

5.5.3 Video Usability Information

The IRD shall support the use of Video Usability Information of the following syntax elements:

- Aspect Ratio Information *(**aspect_ratio_idc**);*
- Colour Parameter Information *(**colour_primaries**, **transfer_characteristics**, and **matrix_coefficients**);*
- Chrominance Information *(**chroma_sample_loc_type_top_field** and **chroma_sample_loc_type_bottom_field**);*
- Timing information *(**time_scale, num_units_in_tick**, and **fixed_frame_rate_flag**).*

Picture Structure Information (***pic_struct_present_flag***)

5.5.3.1 Aspect Ratio Information

The support of **aspect_ratio_idc** values for H.264/AVC SDTV IRDs and Bitstreams is specified in clause 5.6.1.2 and for H.264/AVC HDTV IRDs and Bitstreams is specified in clause 5.7.1.2.

5.5.3.2 Colour Parameter Information

The support of **colour_primaries**, **transfer_characteristics**, and **matrix_coefficients** values for the 25 Hz H.264/AVC SDTV IRD and Bitstream is specified in clause 5.6.2.1, for the 30 Hz H.264/AVC SDTV IRD and Bitstream is specified in clause 5.6.3.1, and for H.264/AVC HDTV IRDs and Bitstreams is specified in clause 5.7.1.3.

5.5.3.3 Chrominance Information

Encoding: It is recommended to specify the chrominance locations using the syntax elements **chroma_sample_loc_type_top_field** and **chroma_sample_loc_type_bottom_field** in the VUI. It is recommended to use chroma sample type equal to 0 for both fields.

Decoding: *H.264/AVC IRDs shall support decoding any allowed values of **chroma_sample_loc_type_top_field** and **chroma_sample_loc_type_bottom_field**.* It is recommended that appropriate processing be included for the display of pictures.

5.5.3.4 Timing Information

The support of **time_scale** and **num_units_in_tick** values for the 25 Hz H.264/AVC SDTV IRD and Bitstream is specified in clause 5.6.2.2, for the 30 Hz H.264/AVC SDTV IRD and Bitstream is specified in clause 5.6.3.2, for the 25 z H.264/AVC HDTV IRD and Bitstream is specified in clause 5.7.2.1, for the 30 Hz H.264/AVC HDTV IRD and Bitstream is specified in clause 5.7.3.1.*In the case of still picture the **fixed_frame_rate_flag** shall be equal to 0. In other cases, the **fixed_frame_rate_flag** shall be equal to 1.* The frame rate can not be changed between two IDR access units.

5.5.3.5 Picture Structure Information

The support of **pic_struct_present_flag** and Bitstream is specified in clause 5.5.4.1 related to use of Picture Structure information in the Picture Timing SEI and is common to all H.264/AVC IRDs and Bitstreams. For bitstreams that carry the picture structure information (such as film mode), it is recommended that the **pic_struct_present_flag** bet set to "1" in the VUI and the picture timing SEI is associated with each access unit in the coded sequence. If the sequence does not require picture structure information, then the **pic_struct_present_flag** should be set to "0" in the VUI. Use of this flag bit in the VUI allows use of picture timing SEI with only the picture structure information without the need to include HRD information (such as CPB and DPB delay or initial values of the delay in the buffering period SEI).

5.5.4 Supplemental Enhancement Information

The IRD shall support the use of Supplemental Enhancement Information of the following message types:

- Picture Timing SEI Message;
- Pan and Scan Rectangle SEI Message.

In addition for IRDs that support AFD (as described in annex B), support for **user_data_registered_itu_t_t35** is required.

5.5.4.1 Picture Timing SEI Message

Encoding: The Picture Timing SEI *message shall be associated with every access unit. If the H.264 bit stream contains picture structure information, then the **pic_struct_present_flag** shall be set to "1" in the VUI and the Picture Timing SEI message shall be associated with every access unit. Otherwise the **pic_struct_present_flag** shall be set to "0".*

Decoding: *H.264/AVC IRDs shall support all values defined in **pic_struct** including all modes requiring field and frame repetition.* The H.264/AVC IRDs need not make use of any other syntax elements (except **pic_struct**) in the Picture Timing SEI message, if these elements are present.

5.5.4.2 Pan-Scan Rectangle SEI Message

Encoding: The **pan_scan_rect** SEI may be used when appropriate.

Decoding: *H.264/AVC IRDs shall support all values specified in **pan_scan_rect**, except **pan_scan_rect_top_offset[i]** and **pan_scan_rect_bottom_offset[i]**.* The IRD need not make use of **pan_scan_rect_top_offset[i]** and **pan_scan_rect_bottom_offset[i]** parameters in the **pan_scan_rect** SEI message.

The support of the use of **pan_scan_rect** for up sampling is specified to allow a 4:3 monitor to give a full-screen display of a selected portion of a 16:9 coded picture with the correct aspect ratio. The support of vertical resampling to obtain the correct aspect ratio for a letterbox display of a 16:9 coded picture on a 4:3 monitor is optional.

5.5.4.3 Still pictures

Encoding: *Still pictures shall comply with "AVC still picture" definition as per ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 / Amd-3 [1]. For Still pictures the frame rate specification for H264 AVC IRDs shall not apply. The **fixed_frame_rate_flag** shall be equal to 0.*

NOTE: For display that requires a fixed frame refresh according to the IRD frequency, the previously decoded picture should be displayed till the next picture is available.

5.5.5 Random Access Point

The definition for H.264/AVC RAP in clause 3 shall apply.

Encoding: The time interval between H.264/AVC RAPs can vary between programs and also within a program. The broadcast requirements should set the time interval between H.264/AVC RAPs. *The maximum time interval between two H.264/AVC RAPs shall be less than or equal to 5 seconds.*

NOTE 1: Decreasing the time interval between H.264/AVC RAPs may reduce channel hopping time and improve trick modes, but may reduce the efficiency of the video compression. For some applications including PVR, the recommended time interval between two H.264/AVC RAPs is less than 1 s.

NOTE 2: Having a regular interval between H.264/AVC RAPs may improve trick mode performance, but may reduce the efficiency of the video compression.

Pictures with Presentation Time Stamp earlier than the Presentation Time Stamp of the picture of the H.264/AVC RAP shall not be reference pictures for inter prediction in pictures with Presentation Time Stamp later than the Presentation Time Stamp of the picture of the H.264/AVC RAP.

Packetization of random access points shall comply with the following additional rule:

A transport packet containing the PES header of a H.264/AVC RAP shall have an adaptation field. The ***payload_unit_start_indicator*** *bit shall be set to "1" in the transport packet header and the* ***adaptation_field_control*** *bits shall be set to "11"(as per ITU-T Recommendation H.222.0 | ISO/IEC 13818-1* [1]*). In addition, the* ***random_access_indicator*** *bit in the adaptation header shall be set to "1". The* ***elementary_stream_priority_indicator*** *bit shall also be set to "1" in the same adaptation header if this transport packet contains the slice start code of the H.264/AVC RAP access unit (see clauses 4.1.5.1 and 4.1.5.2).*

Decoding: *H.264/AVC IRDs shall be able to start decoding and displaying an H.264/AVC Bitstream at an H.264/AVC RAP.*

5.6 H.264/AVC SDTV IRDs and Bitstreams

5.6.1 Specifications Common to all H.264/AVC SDTV IRDs and Bitstreams

The specification in this clause applies to the following IRDs and bitstreams:

- 25 Hz H.264/AVC SDTV IRD and Bitstream;
- 30 Hz H.264/AVC SDTV IRD and Bitstream.

5.6.1.1 Sequence Parameter Set and Picture Parameter Set

Encoding: *In addition to the provisions set forth in ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16], the following restrictions apply for the fields in the sequence parameter set:*

profile_idc	*= 77 (Main Profile)*
profile_idc	= 100 when bitstream complies with High Profile. See clause 5.6.1.2 for details of when the bitstream may optionally comply with High Profile
constraint_set0_flag	*= 0*
constraint_set1_flag	*= 1(when profile_idc = 77) or = 0 (when profile_idc = 100)*
constraint_set2_flag	*= 0*
constraint_set3_flag	*= 0 (when profile_idc = 100)*
gaps_in_frame_num_value_allowed_flag	*= 0 (gaps not allowed)*
vui_parameters_present_flag	*= 1*

5.6.1.2 Profile and level

Encoding: *H.264/AVC SDTV Bitstreams shall comply with Main Profile Level 3 restrictions, as described in ITU-T Recommendation H.264 | ISO/IEC 14496-10 [*16*]*. In addition, in applications where decoders support the High Profile, the encoded bitstream may optionally comply with the High Profile.

The value of level_idc shall be equal to 30.

Decoding: *H.264/AVC SDTV IRDs shall support decoding and displaying of Main Profile Level 3 bitstreams.* Support of the High Profile and other profiles beyond Main Profile is optional. Support of levels beyond Level 3 is optional. *If the H.264/AVC SDTV IRD encounters an extension which it cannot decode, it shall discard the following data until the next start code prefix* (to allow backward compatible extensions to be added in the future*).*

5.6.1.3 Aspect ratio

Encoding: *The source aspect ratio in H.264/AVC SDTV Bitstreams shall be either 4:3 or 16:9.*

The frame cropping information in the Sequence Parameter Set may be used when appropriate.

Decoding: *H.264/AVC SDTV IRDs shall support decoding and displaying H.264/AVC SDTV Bitstreams with the values of aspect_ratio_idc and other constraints that are specified in clause 5.6.2 for the 25 Hz H.264/AVC SDTV IRDs and Bitstreams and 5.6.3 for the 30 Hz H.264/AVC SDTV IRDs and Bitstreams.*

The source aspect ratio information shall be derived from the ***pic_height_in_map_units_minus1*** *and the* ***pic_width_in_mbs_minus1*** *and the frame cropping information coded in the Sequence Parameter Set as well as the sample aspect ratio encoded with the* ***aspect_ratio_idc*** *value in the Video Usability Information (see values of* ***aspect_ratio_idc*** *in ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16], table E-1).*

H.264/AVC SDTV IRDs shall support frame cropping.

5.6.2 25 Hz H.264/AVC SDTV IRD and Bitstream

This clause specifies the 25 Hz H.264/AVC SDTV IRD and Bitstream. All specifications in clauses 5.5 and 5.6.1 shall apply. The specification in the remainder of this clause only applies to the 25 Hz H.264/AVC SDTV IRD and Bitstream.

5.6.2.1 Colour Parameter Information

Encoding: *The chromaticity co-ordinates of the ideal display, opto-electronic transfer characteristic of the source picture and matrix coefficients used in deriving luminance and chrominance signals from the red, green and blue primaries shall be explicitly signalled in the encoded 25 Hz H.264/AVC SDTV Bitstream by setting the appropriate values for each of the following 3 parameters in the* ***VUI****:* ***colour_primaries****,* ***transfer_characteristics****, and* ***matrix_coefficients****.*

It is recommended that BT.470-2 System B, G colorimetry is used in the H.264/AVC bitstream, which is signalled by setting **colour_primaries** to the value 5, **transfer_characteristics** to the value 5 and **matrix_coefficients** to the value 5.

Decoding: *25 Hz H.264/AVC SDTV IRDs shall support decoding bitstreams with any allowed values of* ***colour_primaries****,* ***transfer_characterstics*** *and* ***matrix_coefficients****.* It is recommended that appropriate processing be included for the accurate representation of pictures using BT.470-2 System B, G colorimetry.

5.6.2.2 Frame rate

Encoding: The frame rate shall be 25 Hz in 25 Hz H.264/AVC Bitstreams. This shall be indicated in the VUI by setting **time_scale** and **num_units_in_tick** according to table 7. Time_scale and num_units_in_tick define the picture rate of the video.

Table 7: time_scal and num_units_in_tick for Progressive and Interlace Frame Rates for 25 Hz H.264/AVC SDTV

Frame Rate	Interlaced or Progressive	time_scale	num_units_in_tick
25	P	50	1
25	I	50	1

Decoding: *25 Hz H.264/AVC SDTV IRDs shall support decoding and displaying video with a frame rate of 25 Hz within the constraints of Main Profile at Level 3.* Support of other frame rates is optional.

5.6.2.3 Luminance resolution

Encoding: *25 Hz H.264/AVC SDTV Bitstreams shall represent video with luminance resolutions as shown in table 8.* Non full-screen pictures may be encoded for display at less than full-size (when using one of the standard up-conversion ratios at the 25 Hz H.264/AVC SDTV IRD).

Decoding: *25 Hz H.264/AVC SDTV IRDs shall be capable of decoding pictures with luminance resolutions as shown in table 8 and applying up sampling to allow the decoded pictures to be displayed at full-screen size. In addition, 25 Hz H.264/AVC SDTV IRDs shall be capable of decoding lower picture resolutions and displaying them at less than full-size after using one of the standard up-conversions, e.g. a horizontal resolution of 704 pixels within the 720 pixel full-screen display.*

Table 8: Resolutions for Full-screen Display from IRD

Coded Picture			Displayed Picture Horizontal up sampling	
Luminance resolution (horizontal × vertical)	**Source Aspect Ratio**	**aspect_ratio_idc**	**4:3 Monitors**	**16:9 Monitors**
720 × 576	4:3 16:9	2 4	× 1 × 4/3 (see note 2)	× 3/4 (see note 1) × 1
544 × 576	4:3 16:9	4 12	× 4/3 × 16/9 (see note 2)	× 1 (see note 1) × 4/3
480 × 576	4:3 16:9	10 6	× 3/2 × 2 (see note 2)	× 9/8 (see note 1) × 3/2
352 × 576	4:3 16:9	6 8	× 2 × 8/3 (see note 2)	× 3/2 (see note 1) × 2
352 × 288	4:3 16:9	2 4	× 2 × 8/3 (see note 2) (and vertical up sampling × 2)	× 3/2 (see note 1) × 2 (and vertical up sampling × 2)

NOTE 1: Up sampling of 4:3 pictures for display on a 16:9 monitor is optional in the IRD, as 16:9 monitors can be switched to operate in 4:3 mode.
NOTE 2: The up sampling with this value is applied to the pixels of the 16:9 picture to be displayed on a 4:3 monitor.

5.6.3 30 Hz H.264/AVC SDTV IRD and Bitstream

This clause specifies the 30 Hz H.264/AVC SDTV IRD and Bitstream. All specifications in clauses 5.5 and 5.6.1 shall apply. The specification in the remainder of this clause only applies to the 30 Hz H.264/AVC SDTV IRD and Bitstream.

5.6.3.1 Colour Parameter Information

Encoding: *The chromaticity co-ordinates of the ideal display, opto-electronic transfer characteristic of the source picture and matrix coefficients used in deriving luminance and chrominance signals from the red, green and blue primaries shall be explicitly signalled in the encoded H.264/AVC bitstream by setting the appropriate values for each of the following 3 parameters in the **VUI**: **colour_primaries**, **transfer_characteristics**, and **matrix_coefficients**.*

It is recommended that SMPTE-170M colorimetry is used for video of all other vertical resolutions in the H.264/AVC bitstream, which is signalled by setting **colour_primaries** to the value 6, **transfer_characteristics** to the value 6 and **matrix_coefficients** to the value 6.

Decoding: *The 30 Hz H.264/AVC SDTV IRD shall be capable of decoding bitstreams with any allowed values of **colour_primaries**, **transfer_characterstics** and **matrix_coefficients**.* It is recommended that appropriate processing be included for the accurate representation of pictures using SMPTE-170M colorimetry.

5.6.3.2 Frame rate

Encoding: *The frame rate shall be 24 000/1 001, 24, 30 000/1 001, 30 Hz. This shall be indicated in the VUI by setting **time_scale** and **num_units_in_tick** according to table 9.* Time_scale and num_units_in_tick define the picture rate of the video.

Table 9: time_scal and num_units_in_tick for Progressive and Interlace Frame Rates for 30 Hz H.264/AVC SDTV

Frame Rate	Interlaced or Progressive	time_scale	num_units_in_tick
24 000/ 1 001	P	48 000	1 001
24	P	48	1
30 000/ 1 001	P	60 000	1 001
30	P	60	1
30 000/ 1 001	I	60 000	1 001
30	I	60	1

Decoding: *The 30 Hz H.264/AVC SDTV IRD shall support decoding and displaying video with a frame rate of 24 000/1 001, 24, 30 000/1 001 or 30 Hz within the constraints of Main Profile at Level 3.* Support of other frame rates is optional.

5.6.3.3 Luminance resolution

Encoding: *30 Hz H.264/AVC SDTV Bitstreams shall represent video with luminance resolutions as shown in table 10.* Non full-screen pictures may be encoded for display at less than full-size (when using one of the standard up-conversion ratios at the 30 Hz H.264/AVC SDTV IRD).

Decoding: *30 Hz H.264/AVC SDTV IRDs shall be capable of decoding pictures with luminance resolutions as shown in table 10 and applying up sampling to allow the decoded pictures to be displayed at full-screen size. In addition, 30 Hz H.264/AVC SDTV IRDs shall be capable of decoding lower picture resolutions and displaying them at less than full-size after using one of the standard up-conversions, e.g. a horizontal resolution of 704 pixels within the 720 pixel full-screen display.*

Table 10: Resolutions for Full-screen Display from IRD

Coded Picture			Displayed Picture Horizontal up sampling	
Luminance resolution (horizontal × vertical)	**Source Aspect Ratio**	**aspect_ratio_idc**	**4:3 Monitors**	**16:9 Monitors**
720 × 480	4:3 16:9	3 5	× 1 × 4/3 (see note 2)	× 3/4 (see note 1) × 1
640 × 480	4:3 16:9	1 11	× 9/8 × 3/2	× 27/32 (see note 1) × 9/8
544 × 480	4:3 16:9	5 13	× 4/3 × 16/9 (see note 2)	× 1 (see note 1) × 4/3
480 × 480	4:3 16:9	11 7	× 3/2 × 2 (see note 2)	× 9/8 (see note 1) × 3/2
352 × 480	4:3 16:9	7 9	× 2 × 8/3 (see note 2)	× 3/2 (see note 1) × 2
352 × 240	4:3 16:9	3 5	× 2 × 8/3 (see note 2) (and vertical up sampling × 2)	× 3/2 (see note 1) × 2 (and vertical up sampling × 2)

NOTE 1: Up sampling of 4:3 pictures for display on a 16:9 monitor is optional in the IRD, as 16:9 monitors can be switched to operate in 4:3 mode.
NOTE 2: The up sampling with this value is applied to the pixels of the 16:9 picture to be displayed on a 4:3 monitor.

5.7 H.264/AVC HDTV IRDs and Bitstreams

5.7.1 Specifications common to all H.264/AVC HDTV IRDs and Bitstreams

The specification in this clause applies to the following IRDs and bitstreams:

- 25 Hz H.264/AVC HDTV IRD and Bitstream;
- 30 Hz H.264/AVC HDTV IRD and Bitstream.

5.7.1.1 Sequence Parameter Set and Picture Parameter Set

Encoding: *In addition to the provisions set forth in ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16], the following restrictions apply for the fields in the sequence parameter set:*

profile_idc	*= 100 (High Profile [21])*
constraint_set0_flag	*= 0*
constraint_set1_flag	*= 0*
constraint_set2_flag	*= 0*
constraint_set3_flag	*= 0*
gaps_in_frame_num_value_allowed_flag	*= 0 (gaps not allowed)*
vui_parameters_present_flag	*= 1*

5.7.1.2 Profile and level

Encoding: *H.264/AVC HDTV Bitstreams shall comply with the High Profile Level 4 restrictions, as described ISO/IEC 14496-10 [16].*

*The value of **level_idc** shall be equal to 30, 31, 32, or 40.*

Decoding: *H.264/AVC HDTV IRDs shall support the decoding of High Profile Level 4 bitstreams.* This requirement includes support for High Profile and levels 3 to 4. Support for profiles and levels other than High Profile, Level 3 to 4 is optional. *If the H.264/AVC HDTV IRD encounters an extension which it cannot decode, it shall discard the following data until the next start code prefix* (to allow backward compatible extensions to be added in the future*).*

5.7.1.3 Aspect ratio

Encoding: *The source aspect ratio in H.264/AVC HDTV Bitstreams shall be 16:9.*

The source aspect ratio information shall be derived from the **aspect_ratio_idc** value in the Video Usability Information (see values of **aspect_ratio_idc** in ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16], table E-1).

The frame cropping information in the Sequence Parameter Set may be used when appropriate.

Decoding: *H.264/AVC HDTV IRDs shall support decoding and displaying H.264/AVC HDTV Bitstreams with the values of **aspect_ratio_idc** and other constraints that are specified in clause 5.7.2 for the 25 Hz H.264/AVC HDTV IRDs and Bitstreams and 5.7.3 for the 30 Hz H.264/AVC HDTV IRDs and Bitstreams.*

The source aspect ratio information shall be derived from the **pic_height_in_map_units_minus1** and the **pic_width_in_mbs_minus1** and the frame cropping information coded in the Sequence Parameter Set as well as the sample aspect ratio encoded with the **aspect_ratio_idc** value in the Video Usability Information (see values of **aspect_ratio_idc** in ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16], table E-1).

H.264/AVC HDTV IRDs shall support frame cropping.

5.7.1.4 Colour Parameter Information

Encoding: *The chromaticity co-ordinates of the ideal display, opto-electronic transfer characteristic of the source picture and matrix coefficients used in deriving luminance and chrominance signals from the red, green and blue primaries shall be explicitly signalled in the encoded H.264/AVC HDTV Bitstream by setting the appropriate values for each of the following 3 parameters in the **VUI**: **colour_primaries**, **transfer_characteristics**, and **matrix_coefficients**.*

It is recommended that ITU-R Recommendation BT.709 [13] colorimetry is used for all H.264/AVC HDTV Bitstreams, which is signalled by setting **colour_primaries** to the value 1, **transfer_characteristics** to the value 1 and **matrix_coefficients** to the value 1.

NOTE: For the 576P/480P video formats, the colorimetry standards recommended for the SDTV IRDs apply, i.e. BT.470-2 System B, G and SMPTE-170M are recommended for respectively the 50 Hz and 60 Hz formats."

Decoding: H.264/AVC HDTV IRDs shall be capable of decoding bitstreams with any allowed values of colour_primaries, transfer_characterstics and matrix_coefficients. It is recommended that appropriate processing be included for the accurate representation of pictures using ITU-R Recommendation BT.709 [13] colorimetry.

5.7.1.5 Luminance resolution

Encoding: *H.264/AVC HDTV Bitstreams shall represent video with luminance resolutions as shown in table 11.* Non full-screen pictures may be encoded for display at less than full-size (when using one of the standard up-conversion ratios at the H.264/AVC HDTV IRD).

Decoding: *H.264/AVC HDTV IRDs shall be capable of decoding pictures with luminance resolutions as shown in table 11 and applying up sampling to allow the decoded pictures to be displayed at full-screen size.*

Table 11: Resolutions for Full-screen Display from IRD

Coded Picture			
Luminance resolution (horizontal × vertical)	**Source Aspect Ratio**	**aspect_ratio_idc**	**16:9 Monitors Horizontal up sampling**
1 920 × 1 080	16:9	1	× 1
1 440 × 1 080	16:9	11	× 4/3
1 280 × 1 080	16:9	4	× 3/2
960 × 1 080	16:9	6	× 2
1 280 × 720	16:9	1	× 1
960 × 720	16:9	11	× 4/3
640 × 720	16:9	6	× 2

5.7.2 25 Hz H.264/AVC HDTV IRD and Bitstream

This clause specifies the 25 Hz H.264/AVC HDTV IRD and Bitstream. All specifications in clauses 5.5 and 5.7.1 shall apply. The specification in the remainder of this clause only applies to the 25 Hz H.264/AVC HDTV IRD and Bitstream.

5.7.2.1 Frame rate

Encoding: *The frame rate shall be 25 or 50 Hz. This shall be indicated in the VUI by setting **time_scale** and **num_units_in_tick** according to table 12.* Time_scale and num_units_in_tick define the picture rate of the video. *The source video format for 50 Hz frame rate material shall be progressive. The source video format for 25 Hz frame rate material shall be interlaced or progressive.*

Table 12: time_scal and num_units_in_tick for Progressive and Interlace Frame Rates for 25 Hz H.264/AVC HDTV

Frame Rate	Interlaced or Progressive	time_scale	num_units_in_tick
25	P	50	1
25	I	50	1
50	P	100	1

Decoding: *25 Hz H.264/AVC HDTV IRDs shall support decoding and displaying video with a frame rate of 25 Hz interlaced or progressive, or 50 Hz progressive within the constraints of High Profile at Level 4.* Support of other frame rates is optional.

5.7.2.2 Backwards Compatibility

Decoding: *25 Hz H.264/AVC HDTV IRDs shall be capable of decoding any bitstream that a 25 Hz H.264/AVC SDTV IRD is required to decode and resulting in the same displayed pictures as the 25 Hz H.264/AVC SDTV IRD, as described in clause 5.6.2.*

5.7.3 30 Hz H.264/AVC HDTV IRDs and Bitstreams

This clause specifies the 30 Hz H.264/AVC HDTV IRD and Bitstream. All specifications in clauses 5.5 and 5.7.1 shall apply. The specification in the remainder of this clause only applies to the 30 Hz H.264/AVC HDTV IRD and Bitstream.

5.7.3.1 Frame rate

Encoding: *The frame rate shall be 24 000/1 001, 24, 30 000/1 001, 30, 60 000/1 001 or 60 Hz. This shall be indicated in the VUI by setting **time_scale** and **num_units_in_tick** according to table 13.* Time_scale and num_units_in_tick define the picture rate of the video. *The source video format for 24 000/1 001, 24, 60 000/1 001 and 60 Hz frame rate material shall be progressive. The source video format for 30 000/1 001 and 30 Hz frame rate material shall be interlaced or progressive.*

Table 13: time_scal and num_units_in_tick for Progressive and Interlace Frame Rates for 30 Hz H.264/AVC HDTV

Frame Rate	Interlaced or Progressive	time_scale	num_units_in_tick
24 000/ 1 001	P	48 000	1 001
24	P	48	1
30 000/ 1 001	P	60 000	1 001
30	P	60	1
30 000/ 1 001	I	60 000	1 001
30	I	60	1
60 000/ 1 001	P	120 000	1 001
60	P	120	1

Decoding: *30 Hz H.264/AVC HDTV IRDs shall support decoding and displaying video with a frame rate of 30 000/1 001, 30 Hz interlaced or progressive, or 24 000/1 001, 24, 60 000/1 001 or 60 Hz progressive within the constraints of High Profile at Level 4.* Support of other frame rates is optional.

5.7.3.2 Backwards Compatibility

Decoding: *30 Hz H.264/AVC HDTV IRDs shall be capable of decoding any bitstream that a 30 Hz H.264/AVC SDTV IRD is required to decode and resulting in the same displayed pictures as the 30 Hz H.264/AVC SDTV IRD, as described in clause 5.7.2.*

6 Audio

This clause describes the guidelines for encoding MPEG-1 or MPEG-2 layer 2 backward compatible audio in DVB broadcast bit-streams, and for decoding this bit-stream in the IRD. Additional optional audio coding systems and ancillary data are described in annexes C, D, F and H.

The recommended level for reference tones for transmission is 18 dB below clipping level, in accordance with EBU Recommendation R.68 [11].

The audio encoding shall conform to either ISO/IEC 11172-3 [9] or ISO/IEC 13818-3 [3], except in systems where IRDs are required to comply with annex C, F or H. Some of the parameters and fields in ISO/IEC 11172-3 [9] and ISO/IEC 13818-3 [3] are not used in the DVB System and these restrictions are described below.

The IRD design should be made under the assumption that any legal structure as permitted by ISO/IEC 11172-3 [9] or ISO/IEC 13818-3 [3] may occur in the broadcast stream even if presently reserved or unused. *To allow full compliance to ISO/IEC 11172-3 [9] and ISO/IEC 13818-3 [3] and upward compatibility with future enhanced versions, a DVB IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the IRD. For example, an IRD which is not designed to make use of the ancillary data field shall skip over that portion of the bit-stream.*

This clause is based on ISO/IEC 11172-3 [9] (MPEG-1 audio) and ISO/IEC 13818-3 [3] (MPEG-2 backwards compatible audio coding).

6.1 Audio mode

Encoding: *The audio shall be encoded in one of the following modes:*

- ISO/IEC 11172-3 [9] single channel;
- ISO/IEC 11172-3 [9] joint stereo;
- ISO/IEC 11172-3 [9] stereo;
- ISO/IEC 13818-3 [3] multi-channel audio, backwards compatible to ISO/IEC 11172-3 [9] (dematrix procedure = 0, 1 or 2).

In addition, audio may be encoded in ISO/IEC 11172-3 [9] dual channel mode, as specified by TR 102 154, in a transmission intended both as a contribution feed and for direct-to-home (DTH) reception. However, this is not recommended. Care needs to be taken to ensure that the optional dual channel decoding mode is supported in the DTH IRD. Furthermore, there may be problems due to the left/right channel selection being performed by different equipment from the decoding unit (e.g. decoding may be by a set-top-box but left/right channel selection and audio balance may be performed by the TV set).

Decoding: *The IRD shall be capable of decoding the following audio modes:*

- ISO/IEC 11172-3 [9] single channel;
- ISO/IEC 11172-3 [9] joint stereo;
- ISO/IEC 11172-3 [9] stereo.

The IRD shall be capable of decoding at least the ISO/IEC 11172-3 [9] compatible basic stereo information from an ISO/IEC 13818-3 [3] multi-channel audio bit-stream. Full decoding of an ISO/IEC 13818-3 [3] multi-channel audio bit-stream is optional.

Support for decoding of ISO/IEC 11172-3 [9] dual channel is optional.

6.2 Layer

Encoding: *An ISO/IEC 11172-3 [9] encoded bit-stream shall use either Layer I or Layer II coding (**layer** = "11" or "10" respectively). Use of Layer II is recommended.*

*An ISO/IEC 13818-3 [3] multi-channel encoded bit-stream shall use Layer II coding (**layer** = "10").*

Decoding: *IRDs shall be capable of decoding Layer I and Layer II.*

6.3 Bit rate

Encoding: *The value of **bitrate_index** in the encoded bit-stream shall be one of the 14 values from "0001" to "1110"(inclusive).*

For Layer I, these correspond to bit rates of: 32 kbits/s, 64 kbits/s, 96 kbits/s, 128 kbits/s, 160 kbits/s, 192 kbits/s, 224 kbits/s, 256 kbits/s, 288 kbits/s, 320 kbits/s, 352 kbits/s, 384 kbits/s, 416 kbits/s or 448 kbits/s.

For Layer II, these correspond to bit rates of: 32 kbits/s, 48 kbits/s, 56 kbits/s, 64 kbits/s, 80 kbits/s, 96 kbits/s, 112 kbits/s, 128 kbits/s, 160 kbits/s, 192 kbits/s, 224 kbits/s, 256 kbits/s, 320 kbits/s, 384 kbits/s.

For ISO/IEC 13818-3 [3] encoded bit-streams with total bit rates greater than 384 kbit/s, an extension bit-stream shall be used. The bit rate of that extension may be in the range of 0 to 682 kbit/s.

Decoding: *IRDs shall be capable of decoding bit-streams with a value of **bitrate_index** from "0001" to "1110"(inclusive).* Support for the free format bit rate (**bitrate_index** = "0000") is optional.

6.4 Sampling frequency

Encoding: *The audio sampling rate of primary sound services shall be 32 kHz, 44,1 kHz or 48 kHz.* Sampling rates of 16 kHz, 22,05 kHz, 24 kHz, 32 kHz, 44,1 kHz or 48 kHz may be used for secondary sound services.

Decoding: *The IRD shall be capable of decoding audio with sampling rates of 32 kHz, 44,1 kHz and 48 kHz.* Support for sampling rates of 16 kHz, 22,05 kHz and 24 kHz is optional.

6.5 Emphasis

Encoding: *The encoded bit-stream shall have no emphasis (**emphasis** = "00").*

Decoding: *The IRD shall be capable of decoding audio with no emphasis.* Support for 50/15 microseconds de-emphasis and ITU-T Recommendation J.17 [10] de-emphasis (**emphasis** = "01" or "11") is optional.

6.6 Cyclic redundancy code

Encoding: *The parity check word (**crc_check**) shall be included in the encoded bit-stream.*

Decoding: It is recommended that the IRD use **crc_check** to detect errors and subsequently invoke suitable concealment or muting mechanisms.

6.7 Prediction

Encoding: *ISO/IEC 13818-3 [3] multichannel encoded bit-streams shall not use mc_prediction (**mc_prediction_on** equals "0").*

Decoding: *The IRD shall be capable of decoding ISO/IEC 13818-3 [3] multichannel encoded bit-streams which do not use mc_prediction.*

6.8 Multilingual

Encoding: *ISO/IEC 13818-3 [3] multichannel encoded bit-streams shall not contain multilingual channels (**no_of_multilingual_channels** equals "0").*

Decoding: *The IRD shall be capable of decoding ISO/IEC 13818-3 [3] multichannel encoded bit-streams which do not contain multilingual channels.*

6.9 Extension Stream

Encoding: When an ISO/IEC 13818-3 [3] encoded bit-stream uses an extension stream, it is recommended that a continuous stream of extension frames is maintained for the duration of a programme, even if a total bit rate of less than 384 kbits/s would be sufficient to encode individual frames. This prevents undesired resets of the audio decoder.

6.10 Ancillary Data

Encoding: ISO/IEC 13818-3 [3] stereo or multichannel encoded bitstreams may contain ancillary data as described in annex D. It is recommended to include the data in the bitstream.

Decoding: The IRD may interpret the ancillary data field in an ISO/IEC 13818-3 [3] stereo or multichannel bitstream as described in annex D and it is recommended that the contribution IRD make use of this data.

Annex A (informative): Full screen luminance resolutions for SDTV and HDTV

Table A.1: MPEG-2 screen resolution

vertical_size_ value	horizontal_size_v alue	aspect_ratio_ information	frame_rate_ code (see note)	progressive_ sequence	Decodeable by SDTV IRD
			24, 25	1	
1 080	1 920	16:9	23.976, 24, 29.97, 30	1	
			25	0	
			29.97, 30	0	
1 035	1 920	16:9	25	0	
			29.97, 30	0	
			24, 25, 50	1	
720	1 280	16:9	23.976, 24, 29.97, 30, 59.94, 60	1	
			50	1	
	720	4:3, 16:9	25	1	✔
			25	0	✔
576	544	4:3, 16:9	25	1	✔
			25	0	✔
	480	4:3, 16:9	25	1	✔
			25	0	✔
	352	4:3, 16:9	25	1	✔
			25	0	✔
	720	4:3, 16:9	59.94, 60	1	
			23.976, 24, 29.97, 30	1	✔
			29.97, 30	0	✔
	640	4:3	59.94, 60	1	
480			23.976, 24, 29.97, 30	1	✔
			29.97, 30	0	✔
	544	4:3, 16:9	23.976, 29.97	1	✔
			29.97	0	✔
	480	4:3, 16:9	23.976, 29.97	1	✔
			29.97	0	✔
	352	4:3, 16:9	23.976, 29.97	1	✔
			29.97	0	✔
288	352	4:3, 16:9	25	1	✔
240	352	4:3, 16:9	23.976, 29.97	1	✔

NOTE: Shaded "frame_rate_code" values indicate 30 Hz bitstreams, clear values 25 Hz bitstreams.

Table A.2: AVC screen resolution

Vertical size	Horizontal size	Aspect ratio	Frame rate (see note)	Progressive or Interlaced	H.264/AVC Level
1 080	1 920, 1 440, 1 280, 960	16:9	23.976, 24	P	4
			25	I	4
				P	4
			29.97, 30	I	4
720	1 280, 960, 640	16:9	25, 50	P	4
			23.976, 24, 29.97, 30, 59.94, 60	P	4
576	720	4:3, 16:9	50	P	4
	720	4:3, 16:9	25	P	3
				I	3
	544, 480, 352	4:3, 16:9	25	P	3
				I	3
480	720	4:3, 16:9	59.94, 60	P	4
	720, 640, 544, 480, 352	4:3, 16:9	23.976, 24, 29.97, 30	P	3
			29.97, 30	I	3
288	352	4:3	25, 50	P	3
			25	I	3
240	352	4:3	23.976, 24, 29.97, 30, 59.94, 60	P	3
			29.97, 30	I	3

NOTE: Shaded "frame_rate_code" values indicate 30 Hz bitstreams, clear values 25 Hz bitstreams.

Annex B (informative): Active Format Description

B.1 Overview

The Active Format Description (AFD) describes the portion of the coded video frame that is "of interest". It is intended for use in networks that deliver mixed formats to a heterogeneous receiver population. The format descriptions are informative in nature and are provided to assist receiver systems to optimize their presentation of video.

Transmission of this description, and use of this description by a receiver, are both optional.

The AFD is intended for use where there are compatibility problems between the source format of a programme, the format used for the transmission of that programme, and the format of the target receiver population. For example, a wide-screen production may be transmitted as a 14:9 letter-box within a 4:3 coded frame, thus optimized for the viewer of a 4:3 TV, but causing problems to the viewer of a wide screen TV. The appropriate AFD may be transmitted with the video to indicate to the receiver the "area of interest" of the image, thereby enabling a receiver to present the image in an optimum fashion (which will depend on the format and functionality of the receiving equipment combined with the viewer's preferences). In this example, the functionality provided by the AFD is analogous to that provided by Wide Screen Signalling (WSS) described in EN 300 294 [14].

However, the AFD extends WSS by allowing the "area of interest" of a full-frame 16:9 (anamorphic) image to be described, for example to indicate that the centre 4:3 portion of the image has been protected such that a set-top box connected to a 4:3 set may perform a centre cut-out without removing any essential picture information.

The AFD itself does not describe the aspect ratio of the coded frame (as this is described elsewhere in the MPEG-2 and H264/AVC video syntax).

B.2 AFD and MPEG-2 video

B.2.1 Coding

The AFD is carried in the user data of the video elementary stream. After each sequence start (and repeat sequence start) the default aspect ratio of the area of interest is that signalled by the sequence header and sequence display extension parameters. After introduction, an AFD persists until the next sequence start or until another AFD is introduced.

Encoding: Support for the encoding of AFD is optional.

The AFD may be inserted wherever user data may be inserted in the video elementary stream (after the sequence extension, and/or GOP header, and/or picture coding extension, as specified in ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]). For example, it could be inserted once per sequence after each sequence extension, once per GOP after each GOP header, or once per picture after each picture coding extension. It may be changed for each picture.

Decoding: Support for the decoding of AFD is optional.

A decoder that supports the decoding of AFD shall be capable of decoding it from wherever user data may be inserted in the video stream (i.e. after the sequence extension, and GOP header, and picture coding extension).

B.2.2 Syntax and Semantics

The AFD is carried in the user data of the video elementary stream as defined in ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]. The syntax is illustrated in table B.1.

Table B.1: Active Format Description for MPEG-2 video

Syntax	No. of Bits	Identifier
user_data_start_code	32	bslbf
afd_identifier	32	bslbf
"0"	1	bslbf
active_format_flag	1	bslbf
reserved (set to "00 0001")	6	bslbf
if (active_format_flag == 1) {		
reserved (set to "1111")	4	bslbf
active_format	4	bslbf
}		

afd_identifier: A 32 bit field that identifies that the syntax of the user data is as specified here. Its value is 0x44544731.

active_format_flag: A 1 bit flag. A value of "1" indicates that an active format is described in this data structure.

active_format: A 4 bit field describing the "area of interest" in terms of its aspect ratio within the coded frame as defined in ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2].

The active_format is used by the decoder in conjunction with the "source aspect ratio". The source aspect ratio is derived from the "display aspect ratio"(DAR) signalled in the **aspect_ratio_information**, the **horizontal_size**, **vertical_size**, and **display_horizontal_size** and **display_vertical_size** if present (see ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2]):

- If sequence_display_extension() is not present:

$$\text{source aspect ratio} = \text{DAR}$$

- If sequence_display_extension() is present:

$$\text{source aspect ratio} = \text{DAR} \times \frac{\text{display\_horizontal\_size}}{\text{display\_vertical\_size}} \times \frac{\text{vertical\_size}}{\text{horizontal\_size}}$$

The combination of source aspect ratio and active_format allows the decoder to identify whether the "area of interest" is the whole of the frame (e.g. source aspect ratio 16:9, active_format 16:9 centre), a letterbox within the frame (e.g. source aspect ratio 4:3, active_format 16:9 centre), or a "pillar-box" (see NOTE) within the frame (e.g. source aspect ratio 16:9, active_format 4:3 centre).

NOTE: "Pillar-box" describes a frame that the image fails to fill horizontally, in the same way that a "Letterbox" describes a frame that the image fails to fill vertically.

Table B.2: active_format

Active_format	Aspect ratio of the "area of interest"
0000 - 0001	reserved
0010	box 16:9 (top)
0011	box 14:9 (top)
0100	box > 16:9 (centre)
0101 - 0111	reserved
1000	Active format is the same as the coded frame
1001	4:3 (centre)
1010	16:9 (centre)
1011	14:9 (centre)
1100	reserved
1101	4:3 (with shoot & protect 14:9 centre)
1110	16:9 (with shoot & protect 14:9 centre)
1111	16:9 (with shoot & protect 4:3 centre)

The complete set of Active Formats described in the present document is illustrated in table B.3. Note that for each format two example illustrations have been given, corresponding to the source aspect ratio of the coded frame being 4:3 and 16:9. The AFD may also be used with coded frames of other aspect ratios. For example a coded frame of 2.21:1 with active_format 10 would represent a 16:9 image centred (pillar-box) within a 2.21:1 frame.

The Active Formats are illustrated using the following diagrammatic representation.

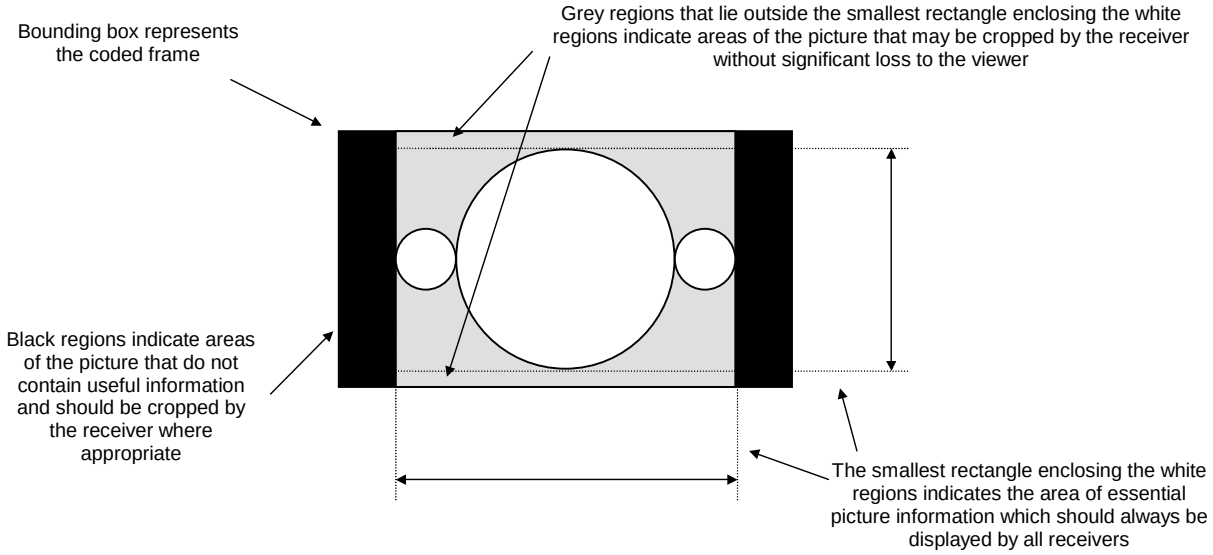
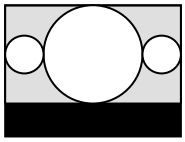
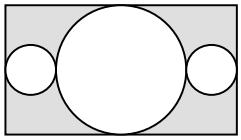
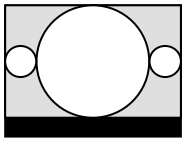
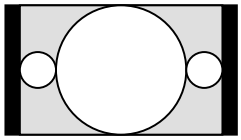
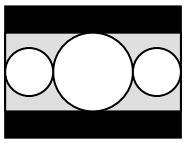
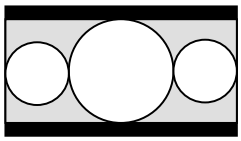
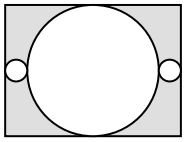
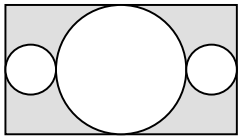
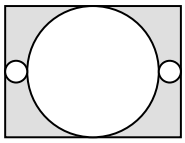
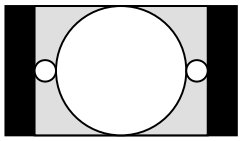
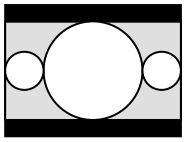
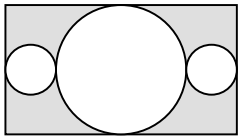
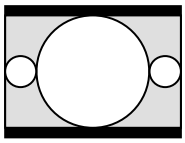
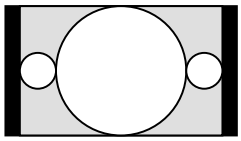
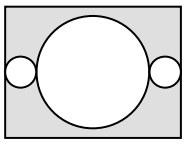
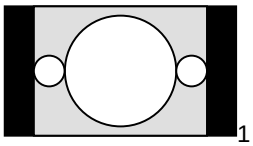
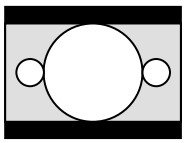
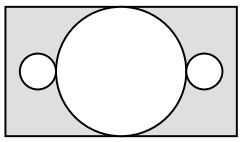
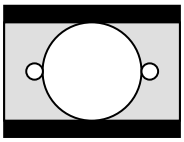
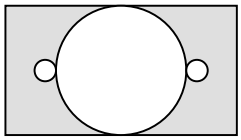


Figure B.1

Table B.3: Active Formats Illustrated

Active_format		Illustration of described format	
value	**description**	**in 4:3 coded frame**	**in 16:9 coded frame**
0000 - 0001	reserved		
0010	box 16:9 (top)		
0011	box 14:9 (top)		
0100	box > 16:9 (centre)		
0101 - 0111	reserved		
1000	As the coded frame		
1001	4:3 (centre)		 (see note)
1010	16:9 (centre)		
1011	14:9 (centre)		
1100	reserved		
1101	4:3 (with shoot & protect 14:9 centre)		 1
1110	16:9 (with shoot & protect 14:9 centre)		

Active_format		Illustration of described format	
value	description	in 4:3 coded frame	in 16:9 coded frame
1111	16:9 (with shoot & protect 4:3 centre)		
NOTE: It is recommended to use the 4:3 coded frame mode to transmit 4:3 source material rather than using a pillar box to transmit it in a 16:9 coded frame. This allows for higher horizontal resolution on both 4:3 and 16:9 sets.			

B.2.3 Relationship with Pan Vectors

Encoding: Encoded bit-streams may optionally include pan vectors and AFDs.

Decoding: The decoder may use the AFD as part of the logic that decides how the IRD processes and positions the reconstructed image for display on a monitor, where the monitor aspect ratio does not match the source aspect ratio (e.g. whether to use pan vectors, or generate a letterbox display).

B.3 AFD and H264/AVC video

B.3.1 Coding

The AFD is carried in the data as Supplemental Enhancement Information in AVC"s "User data registered by ITU-T Recommendation T.35 [20] SEI message" syntactic element (See clauses D.8.5 and D.9.5 of ISO/IEC 14496-10 [16]).

Encoding: Support for the encoding of AFD is optional.

Decoding: Support for the decoding of AFD is optional.

B.3.2 Syntax and Semantics

The AFD is carried in the data as Supplemental Enhancement Information in AVC"s "User data registered by ITU-T Recommendation T.35 SEI message" syntactic element [20]. The syntax is illustrated in table B.2.

Table B.4: Active Format Description for H264/AVC video

user_data_registered_itu_t_t35(payloadSize) {	Descriptor	Notes
itu_t_t35_country_code	b(8)	Registered by DVB
Itu_t_t35_provider_code	u(16)	Registered by DVB
afd_identifier	f(32)	0 x 44544731 ("DTG1")
zero_bit	f(1)	"0"
active_format_flag	u(1)	
alignment_bits	f(6)	"00 0001"
if (active_format_flag == 1) {		
reserved	f(4)	"1111"
active_format	u(4)	
}		
}		

afd_identifier: A 32 bit field that identifies that the syntax of the user data is as specified here. Its value is 0 x 44544731.

itu_t_t35_country_code is a fixed 8-bit field having the value of the country code as registered by DVB. The value is to be a country code as specified by ITU-T Recommendation T.35 [20] annex A.

itu_t_35_provider_code is a fixed 16-bit field having one the value registered by DVB. The value is to be assigned as specified by ITU-T Recommendation T.35 [20].

afd_identifier is a fixed 32-bit field having the value 0x44544731 ("DTG1" in ASCII).

NOTE: In MPEG-2, the only discriminator within **user_data** is this 32-bit value. In the context of AVC, this value is used in addition to country and provider codes to definitively identify this as AFD data.

active_format_flag is a 1-bit flag. A value of "1" indicates that an active format is described in this data structure and that reserved and **active_format** bits immediately follow **alignment_bits**. A value of "0" indicates that no active format is described and that reserved and **active_format** bits are not present in this structure.

active_format is a 4-bit field describing the "area of interest" in terms of its aspect ratio within the coded frame as described in ISO/IEC 14496-10 [16]. The coding of **active_format** is shown in table B.2.

The **active_format** is used by the decoder in conjunction with picture size and shape information as indicated in the sequence parameter set RBSP. In particular, the picture width, picture height, frame cropping information, and sample aspect ratio are important for proper use of **active_format**.

B.4 Relationship with Wide Screen Signalling (WSS)

The AFD provides a super-set of the aspect ratio signalling specified in EN 300 294 [14]. The mapping of source aspect ratio and active_format to WSS Aspect Ratio is given in table B.4.

Table B.5: Support for WSS

Sequence Header	Active Format Description	WSS	
source aspect ratio	value	code (bits 0-3)	description
	1001	0001	full format 4:3
	1011	1000	box 14:9 Centre
	0011	0100	box 14:9 Top
4:3	1010	1101	box 16:9 Centre
	0010	0010	box 16:9 Top
	0100	1011	box > 16:9 Centre
	1101	0111	full format 4:3 (shoot and protect 14:9 Centre)
16:9	1010	1110	full format 16:9 (anamorphic)

B.5 Aspect Ratio Ranges

The labels 4:3, 14:9, 16:9 and > 16:9 used in the AFD shall correspond to the aspect ratio ranges specified in EN 300 294 [14]. (Note that the corresponding active lines specified in EN 300 294 [14] do not, in general, apply.)

Annex C (informative): Guidelines for the Implementation of AC-3 and Enhanced AC-3 Audio in DVB Compliant Transport Streams

C.1 Scope

The inclusion of AC-3 and Enhanced AC-3 audio streams in a DVB multiplex is optional, and IRDs may optionally decode these streams. This annex contains the guidelines to include one or more AC-3 or Enhanced AC-3 elementary streams in a DVB Transport Stream in compliance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1] The coding and decoding of AC-3 and Enhanced AC-3 elementary streams is based upon TS 102 366 [12].

It is recommended that implementations of DVB systems that include AC-3 or Enhanced AC-3 audio streams should comply with this annex.

AC-3 and Enhanced AC-3 packetized elementary streams shall conform to the requirements of a user private stream type 1, as described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1] .

The IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], including private data streams, may occur in the Transport Stream, even if presently reserved or unused. *To allow full compliance to the MPEG-2 standard and upward compatibility with future enhanced versions, a DVB IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the IRD.*

This clause is based on ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1] and TS 102 366 [12].

C.2 Introduction

AC-3 and Enhanced AC-3 elementary bit streams may be multiplexed into an MPEG-2 transport stream in much the same way an MPEG-1 audio stream would be included. The elementary stream is packetized into PES packets with a structure similar to an MPEG audio PES. An MPEG-2 transport stream containing AC-3 or Enhanced AC-3 elementary stream(s) must meet the constraints described in the STD model in clause C.4.5.

It is necessary to unambiguously indicate that an MPEG private stream is, in fact, an AC-3 or an Enhanced AC-3 stream. Two public DVB descriptors, the AC-3_descriptor and the Enhanced_AC-3_Descriptor have been specified for this purpose. The syntactical elements that need to be specified in order to include AC-3 within an MPEG-2 transport stream are: the MPEG stream_type, stream_id and the DVB AC - 3_descriptor. The syntactical elements that need to be specified in order to include Enhanced AC-3 within an MPEG-2 transport stream are: the MPEG stream_type, stream_id and the DVB Enhanced_AC-3_descriptor.

The ISO 639 language descriptor may be used to indicate the language of the content of the AC-3 or Enhanced AC-3 stream.

IRDs compatible with AC-3 shall decode all bit rates and sample rates listed in TS 102 366 [12] (not including annex E) .

IRDs compatible with Enhanced AC-3 shall additionally decode Enhanced AC-3 streams with data rates from 32 kbps to 3 024 kbps and support all sample rates listed in TS 102 366 [12] annex E.

Enhanced AC-3 bit streams are similar in nature to standard AC-3 bit streams, but are not backwards compatible (i.e., they are not decodable by standard AC-3 decoders). Some constraints are placed on the PES layer for the case of multiple audio streams intended to be reproduced in exact sample synchronism as described in clause C.5.

C.3 DVB Compliant Streams

AC-3 and Enhanced AC-3 PES shall be carried as an MPEG private data stream type, conforming to the structure of a private_stream_1 as described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1[1], table 2-18 (stream_id) and table 2-29 (stream_type).

When an AC-3 stream is included in a DVB transport stream, the AC-3_descriptor shall be included. The AC-3_descriptor is defined in EN 300 468 [6] annex D, but for information a description is included here in clause C.4.3. The AC-3_descriptor is located in the PMT and the Selection Information Table of the DVB SI Tables defined in EN 300 468 [6].

When an Enhanced AC-3 stream is included in a DVB transport stream, the Enhanced_AC-3_descriptor shall be included. The Enhanced_AC-3_descriptor is defined in EN 300 468 [6], but for information a description is included here in clause C.4.4. The Enhanced_AC-3_descriptor is located in the PMT and the Selection Information Table of the DVB SI Tables defined in EN 300 468 [6].

Certain other of the DVB Service Information descriptors defined in EN 300 468 [6] can provide additional means of identifying the existence of an AC-3 or Enhanced AC-3 stream without accessing the PMT. The component_descriptor (see clause C.4.2) may have values assigned to its syntactical elements, which indicate both the presence and type of AC-3 or Enhanced AC-3 stream(s) in the DVB-SI.

C.4 Detailed specification

C.4.1 MPEG Transport Stream Compliance

C.4.1.1 Stream_id

Semantics: The semantics of the stream_id field are described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], table 2-18. Multiple AC-3 or Enhanced AC-3 streams may share the same value of stream_id since each stream is carried with a unique PID value. The mapping of values of PID to stream_type is indicated in the transport stream Programme Map Table (PMT).

Encoding: *The value of the stream_id field for an AC-3 or Enhanced AC-3 elementary stream shall be 0xBD (indicating private_stream_1).*

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with MPEG systems syntax.*

C.4.1.2 Stream_type

Semantics: The semantics of the stream_type field are described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], table 2-29.

Encoding: *The recommended value of stream_type for an AC-3 or Enhanced AC-3 elementary stream shall be 0x06 (indicating PES packets containing private data).*

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with MPEG systems syntax.*

C.4.2 Use of the DVB-SI component_descriptor and multilingual_component_descriptor

Semantics: The semantics of the component_descriptor and multilingual_component_descriptor are defined in EN 300 468 [6]. The stream_content and component_type assigned values for DVB AC-3 and Enhanced AC-3 audio streams are listed in EN 300 468 [6], table 26.

Encoding: *The values for the elements of the component_descriptor and multilingual_component_descriptor shall be set in accordance with EN 300 468 [6].*

Decoding: *These fields shall be read by the IRD, and the IRD shall interpret these fields to indicate the type of audio service present.*

C.4.3 AC-3_descriptor

The syntax of the AC-3_descriptor is described in table C.1.

NOTE: Horizontal lines in the table indicate allowable termination points for the descriptor.

The AC-3_descriptor syntax provides information about individual AC-3 elementary streams to be identified in the PSI PMT sections. The descriptor is located in the PSI PMT, and used once in a program map section following the relevant ES_info_length field for any stream containing AC-3 audio coded in accordance with TS 102 366 [12] (not including annex E).

Table C.1: AC-3 descriptor Syntax

Syntax	No. of Bits	Identifier
ac-3_ descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
component_type_flag	1	bslbf
bsid_flag	1	bslbf
mainid_flag	1	bslbf
asvc_flag	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
reserved	1	bslbf

if (component_type_flag)==1{
component_type
}
	8	uimsbf

if (bsid_flag)==1{
bsid
{
	8	uimsbf

if (mainid_flag)==1{
mainid
}
	8	uimsbf

if (asvc_flag)==1{
asvc
}
	8	bslbf

for (i=0;i<n;i++){
additional_info [i]
}
	N x 8	uimsbf

}		

C.4.3.1 descriptor_tag

Encoding: The descriptor tag is an 8-bit field, which identifies each descriptor. The value assigned to the AC-3 descriptor_tag is 0x6A (see EN 300 468 [6], table 12).

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1]*

C.4.3.2 descriptor_length

Semantics: This 8-bit field specifies the total number of bytes of the data portion of the descriptor following the byte defining the value of this field. The AC-3 descriptor has a minimum length of one byte but may be longer depending on the use of the optional flags and the additional_info_loop.

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

C.4.3.3 component_type_flag

Semantics: This 1-bit field is mandatory for AC-3 streams. If set to "1" the optional component_type field is included in the descriptor.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended that IRDs decode this field.*

C.4.3.4 bsid_flag

Semantics: This 1-bit field is mandatory for AC-3 streams. If set to "1" the optional bsid field is included in the descriptor.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended that IRDs decode this field.*

C.4.3.5 mainid_flag

Semantics: This 1-bit field is mandatory for AC-3 streams. If set to "1" the optional mainid field is included in the descriptor.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended that IRDs decode this field.*

C.4.3.6 asvc_flag

Semantics: This 1-bit field is mandatory for AC-3 streams. If set to "1" the optional asvc field is included in the descriptor.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended that IRDs decode this field.

C.4.3.7 reserved flags

Semantics: These 1-bit fields are reserved for future use. They should always be set to "0".

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.3.8 component_type

Semantics: This optional 8-bit field indicates the type of audio carried in the AC-3 elementary stream.

Encoding: This field is set to the same value as the component_type field of the component descriptor (see EN 300 468 [6], table 12).

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.3.9 bsid

Semantics: This optional 8-bit field indicates the AC-3 coding version.

Encoding: The three MSBs should always be set to "0". The five LSBs are set to the same value as the bsid field in the AC-3 elementary stream.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.3.10 mainid

Semantics: This 8-bit field is optional. It contains a number in the range 0 to 7 which identifies a main audio service. Each main service should be tagged with a unique number. This value is used as an identifier to link associated services with particular main services.

Encoding: Each main service should be tagged with a unique number in the range 0 to 7.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.3.11 asvc

Semantics: This 8-bit field is optional.

Encoding: Each bit (0 to 7) indicates to which main service(s) this associated service belongs. The left most bit, bit 7, indicates whether this associated service may be reproduced along with main service number 7. If the bit has a value of 1, the service is associated with main service number 7. If the bit has a value of 0, the service is not associated with main service number 7.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.3.12 additional_info

Semantics: These optional bytes are reserved for future use.

Decoding: *IRDs shall be able to accept bit-streams, which contain these bytes.* IRDs may ignore the data within these bytes.

C.4.4 Enhanced_AC-3_Descriptor

The syntax of the Enhanced_AC-3_descriptor is described in table C.2.

NOTE: Horizontal lines in the table indicate allowable termination points for the descriptor.

The Enhanced_AC-3_descriptor syntax provides information about individual Enhanced AC-3 elementary streams to be identified in the PSI PMT sections. The descriptor is located in the PSI PMT, and used once in a program map section following the relevant ES_info_length field for any stream containing Enhanced AC-3 audio coded in accordance with TS 102 366 [12], annex E.

Table C.2: Enhanced_AC-3 descriptor Syntax

Syntax	No.of Bits	Identifier
Enhanced AC-3_ descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
component_type_flag	1	bslbf
bsid_flag	1	bslbf
mainid_flag	1	bslbf
asvc_flag	1	bslbf
mixinfoexists	1	bslbf
substream1_flag	1	bslbf
substream2_flag	1	bslbf
substream3_flag	1	bslbf

If (component_type_flag)==1{
component_type
}
	8	uimsbf

If (bsid_flag)==1{
bsid
{
	8	uimsbf

If (mainid_flag)==1{
mainid
}
	8	uimsbf

If (asvc_flag)==1{
asvc
}
	8	bslbf

If (substream1_flag)==1{
substream1
}
	8	uimsbf

If (substream2_flag)==1{
substream2
}
	8	uimsbf

If (substream3_flag)==1{
substream3
}
	8	uimsbf

For (I=0;I<N;I++){
additional_info [I]
}
	N x 8	uimsbf

}		

C.4.4.1 descriptor_tag

Encoding: The descriptor tag is an 8-bit field, which identifies each descriptor. The value assigned to the Enhanced_AC-3 descriptor_tag is 0x7A (see EN 300 468 [6], table 12).

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

C.4.4.2 descriptor_length

Semantics: This 8-bit field specifies the total number of bytes of the data portion of the descriptor following the byte defining the value of this field. The Enhanced AC-3 descriptor has a minimum length of one byte but may be longer depending on the use of the optional flags and the additional_info_loop.

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

C.4.4.3 component_type_flag

Semantics: This 1-bit field is mandatory for Enhanced AC-3 streams. If set to "1" the optional component_type field is included in the descriptor.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended that IRDs decode this field.*

C.4.4.4 bsid_flag

Semantics: This 1-bit field is mandatory for Enhanced AC-3 streams. If set to "1" the optional bsid field is included in the descriptor.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended that IRDs decode this field.*

C.4.4.5 mainid_flag

Semantics: This 1-bit field is mandatory for Enhanced AC-3 streams. If set to "1" the optional mainid field is included in the descriptor.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended that IRDs decode this field.*

C.4.4.6 asvc_flag

Semantics: This 1-bit field is mandatory for Enhanced AC-3 streams. If set to "1" the optional asvc field is included in the descriptor.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended that IRDs decode this field.

C.4.4.7 mixinfoexists

Semantics: This 1-bit field is mandatory for Enhanced AC-3 streams. If set to "1" the Enhanced AC-3 stream contains metadata in independent substream 0 to control mixing with another AC-3 or Enhanced AC-3 stream.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended that IRDs decode this field.

C.4.4.8 substream1_flag

Semantics: This 1-bit field is mandatory. It should be set to "1" to include the optional substream1 field in the descriptor. This flag should be set to "1" when the Enhanced AC-3 stream contains an additional programme carried in independent substream 1.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended that IRDs decode this field.

C.4.4.9 substream2_flag

Semantics: This 1-bit field is mandatory. It should be set to "1" to include the optional substream2 field in the descriptor. This flag should be set to "1" when the Enhanced AC-3 stream contains an additional programme carried in independent substream 2.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended that IRDs decode this field.

C.4.4.10 substream3_flag

Semantics: This 1-bit field is mandatory. It should be set to "1" to include the optional substream3 field in the descriptor. This flag should be set to "1" when the Enhanced AC-3 stream contains an additional programme carried in independent substream 3.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended that IRDs decode this field.

C.4.4.11 component_type

Semantics: This optional 8-bit field indicates the type of audio carried in the Enhanced AC-3 elementary stream.

Encoding: This field is set to the same value as the component_type field of the component descriptor (see EN 300 468 [6] annex D, table D.1).

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.4.12 bsid

Semantics: This optional 8-bit field indicates the Enhanced AC-3 coding version.

Encoding: The three MSBs should always be set to "0". The five LSBs are set to the same value as the bsid field in the Enhanced AC-3 elementary stream.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.4.13 mainid

Semantics: This 8-bit field is optional. It contains a number in the range 0 to 7 which identifies a main audio service. Each main service should be tagged with a unique number. This value is used as an identifier to link associated services with particular main services.

Encoding: Each main service should be tagged with a unique number in the range 0 to 7.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.4.14 asvc

Semantics: This 8-bit field is optional.

Encoding: Each bit (0 to 7) indicates to which main service(s) this associated service belongs. The left most bit, bit 7, indicates whether this associated service may be reproduced along with main service number 7. If the bit has a value of 1, the service is associated with main service number 7. If the bit has a value of 0, the service is not associated with main service number 7.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.4.15 substream1

Semantics: This optional 8-bit field indicates the type of audio carried in independent substream 1 of the Enhanced AC-3 elementary stream. The value assignments of each bit are indicated in table C.3.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.4.16 substream2

Semantics: This optional 8-bit field indicates the type of audio carried in independent substream 2 of the Enhanced AC-3 elementary stream. The value assignments of each bit are indicated in table C.3

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

C.4.4.17 substream3

Semantics: This optional 8-bit field indicates the type of audio carried in independent substream 3 of the Enhanced AC-3 elementary stream. The value assignments of each bit are indicated in table C.3.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

Table C.3: Substream field byte value assignments

Substream1 - 3 bit values								Description
mixing metadata flag	full service flag	Service type flags			number of channels flags			
B7	B6	B5	B4	B3	B2	B1	B0	
1	X	X	X	X	X	X	X	Mixing metadata present in substream
0								No mixing metadata present in substream
	1	X	X	X	X	X	X	Main Service
X	0							Associated Service
	X	X	X	X	0	0	0	Mono
					0	0	1	1+1 Mode
					0	1	0	2 channel (stereo)
					0	1	1	2 channel Dolby Surround encoded (stereo)
					1	0	0	Multichannel audio (> 2 channels)
					1	0	1	Multichannel audio (> 5.1 channels)
					1	1	0	Reserved
					1	1	1	Reserved
0	1	0	0	0	X	X	X	Complete Main (CM)
X	0	0	0	1				Music and Effects (ME)
	X	0	1	0				Visually Impaired (VI)
	X	0	1	1				Hearing Impaired (HI)
	0	1	0	0				Dialogue (D)
	X	1	0	1	0	0	0	Commentary (C)
0	1	1	1	0				Emergency (E)
X	0	1	1	1				Voiceover (VO)
X	1	1	1	1	X	X	X	Karaoke (mono and "1+1" prohibited)

C.4.4.18 additional_info

Semantics: These optional bytes are reserved for future use.

Decoding: *IRDs shall be able to accept bit-streams, which contain these bytes.* IRDs may ignore the data within these bytes.

C.4.5 STD audio buffer size

It is recommended that for AC-3 and Enhanced AC-3 audio in a DVB system, the main audio buffer size (BS_n) has a fixed value of 5 696 bytes. Refer to ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1] for the derivation of (BS_n) for audio elementary streams.

C.5 AC-3 and Enhanced AC-3 PES constraints

C.5.1 Encoding

In some applications, the audio decoder may be capable of simultaneously decoding two elementary streams containing different programme elements, and then combining the programme elements into a complete programme.

Most of the programme elements are found in the main audio service. Another programme element (such as a spoken narration of the picture content intended for the visually impaired listener, a specially created dialogue based audio service for the hearing impaired listener, or additional audio services such as a spoken director's commentary or alternative languages) may be found in an associated audio service.

In order to have the audio from the two elementary streams reproduced in exact sample synchronism, it is necessary for the original audio elementary stream encoders to have encoded the two audio programme elements frame synchronously; i.e., if audio stream 1 has sample 0 of frame *n* taken at time *t* 0, then audio stream 2 should also have frame *n* beginning with its sample 0 taken the identical time *t* 0. *If the encoding of multiple audio services is done frame and sample synchronous, and decoding is intended to be frame and sample synchronous, then the PES packets of these audio services shall contain identical values of PTS, which refer to the audio access units intended for synchronous decoding.*

Audio services intended to be combined together for reproduction according to the mixing process defined in TS 102 366 [12](annex E) shall meet the following constraints:

- *Audio services intended to be combined together for reproduction shall be encoded at an identical sample rate.*
- *The main programme audio shall be encoded as either an AC-3 or an Enhanced AC-3 elementary stream. The associated audio service shall be encoded as an Enhanced AC-3 elementary stream.*
- *The Enhanced AC-3 elementary stream carrying the associated audio service shall contain mixing metadata for use by the decoder to control the mixing process.*
- *The main programme shall contain from 1 to 5.1 channels of audio. The Enhanced AC-3 elementary stream that carries the associated audio services to be mixed with the main programme audio shall contain no more than two audio channels, and shall not contain more audio channels than the main audio programme.*
- *Dual-mono coding mode is not supported for either the main programme or associated audio service.*
- *The encoding of the associated audio service and subsequent creation of the associated audio service elementary stream shall be done with knowledge of the encoding of the main programme stream.*
- *The pgmscl field in the associated programme bitstream should be set to a positive value. It is recommended this be positive 12 dB to match the default user volume adjustment setting in the decoder.*

C.5.2 Decoding

If audio access units from two audio services which are to be simultaneously decoded have identical values of PTS indicated in their corresponding PES headers, then the corresponding audio access units shall be presented to the audio decoder for simultaneous synchronous decoding. Synchronous decoding means that for corresponding audio frames (access units), corresponding audio samples are presented at the identical time.

If the PTS values do not match (indicating that the audio encoding was not frame synchronous) then the audio frames (access units) of the main audio service may be presented to the audio decoder for decoding and presentation at the time indicated by the PTS. An associated service, which is being simultaneously decoded, may have its audio frames (access units), which are in closest time alignment (as indicated by the PTS) to those of the main service being decoded, presented to the audio decoder for simultaneous decoding. In this case the associated service may be reproduced out of sync by as much as 1/2 of a frame time. (This is typically satisfactory; a visually impaired narration does not require highly precise timing.)

A minimum functionality mixer is described in clause E.4 of TS 102 366 [12]. IRDs that implement this mixing method shall set the default user volume adjustment of the associated programme level to minus 12 dB.

The IRD may use the ISO 639 language descriptor to indicate the language of the content of the associated programme. As the associated services are carried in separate elementary streams to the main service different languages may be indicated for each programme stream.

C.5.3 Byte-alignment

The AC-3 and Enhanced AC-3 elementary stream shall be byte-aligned within the MPEG-2 data stream. This means that the initial 8 bits of an AC-3 or Enhanced AC-3 frame shall reside in a single byte, which is carried by the MPEG-2 data stream.

C.6 Enhanced AC-3 with multiple independent substreams PES constraints

C.6.1 Encoding

In some applications, the audio decoder may be capable of simultaneously decoding two different programme elements, carried as separate independent substreams within a single Enhanced AC-3 elementary stream, and then combining the programme elements into a complete programme.

Most of the programme elements are found in the main audio service. Another programme element (such as a spoken narration of the picture content intended for the visually impaired listener, a specially created dialogue based audio service for the hearing impaired listener or additional audio services such as a spoken director's commentary) may be found in one or more independent substreams carried in the same Enhanced AC-3 bitstream as the main programme.

The Enhanced AC-3 elementary stream shall contain no more than three independent substreams in addition to the independent substream containing the main audio programme. The main audio programme shall only be delivered in independent substream 0.

In order to have the independent substreams containing audio from the main programme and the associated audio service reproduced in exact sample synchronism, it is necessary for the Enhanced AC-3 encoder to have encoded all of the audio programme elements frame synchronously; i.e., if the independent substream 0 has sample 0 of frame *n* taken at time t_0, then independent substream 1 should also have frame *n* beginning with its sample 0 taken the identical time t_0.

Independent substreams intended to be combined together for reproduction according to the mixing process defined in TS 102 366 [12] (annex E) shall meet the following constraints:

- *Independent substreams intended to be combined together for reproduction shall be encoded at an identical sample rate.*
- *The independent substream carrying the associated audio service shall contain mixing metadata for use by the decoder to control the mixing process.*
- *The independent substream that carries the main programme shall contain from 1 to 5.1 channels of audio. The independent substream that carries the associated audio services to be mixed with the main programme audio shall contain no more than two audio channels, and shall not contain more audio channels than the main audio programme.*
- *Dual-mono coding mode is not supported for either the main programme or associated audio service.*
- *The encoding of the associated audio service and subsequent creation of the associated audio service substream shall be done with knowledge of the encoding of the main programme substream.*
- *The pgmscl field in the associated programme substream should be set to a positive value. It is recommended this be positive 12 dB to match the default user volume adjustment setting in the decoder.*

C.6.2 Decoding

IRDs shall be able to accept Enhanced AC-3 elementary streams that contain more than one independent substream.

For TV-broadcasting applications, noticeably public service broadcasting, there is often a requirement for commentary or narration audio services to provide for different languages or Visually Impaired or Hearing Impaired audiences. To allow cost effective transmission and reproduction of these services it is strongly recommended that IRDs be able to select additional independent substreams carried in an Enhanced AC-3 elementary stream and mix the selected independent substream with the main audio programme. A minimum functionality mixer is described in clause E.4 of TS 102 366 [12]. IRDs that include this mixing capability shall set the default user volume adjustment of the associated programme level to minus 12 dB.

The IRD may use the ISO 639 language descriptor to indicate the language of the content of the main programme. As the associated programmes are carried in the same elementary stream as the main programme, the IRD shall assume that the language of associated programmes carried in independent substreams is the same as that of the main programme. To deploy associated programmes with different languages than the main programme, separate Enhanced AC-3 elementary streams shall be used, as described in clauses C.5.1 and C.5.2.

IRDs that support multiple different output-interfaces, for example headphone output or baseband analogue outputs, may optionally support separate mixes for each output created by multiple Enhanced AC-3 decoders.

Annex D (informative): Implementation of Ancillary Data for MPEG Audio

D.1 Scope

This annex contains the guidelines required to include ancillary data in the MPEG Audio elementary stream.

The IRD design should be made under the assumption that any structure as permitted by this annex may occur in the broadcast stream. The IRD is not required to make use of this data but its use is recommended.

D.2 Introduction

An MPEG audio elementary stream provides for the inclusion of ancillary data. This data can be used to convey specific information about the audio content to the decoder, allowing the broadcaster to control rendering of the content to a greater extent. The data includes dynamic range control information and dialogue normalization information.

In case of MPEG1 streams or MPEG2 streams without an extension stream (MPEG audio format1), ancillary data described in this annex is placed at the end of each base frame.

In case of MPEG2 streams with extension stream (MPEG audio format 2), the ancillary data described in this annex is placed at the end of each base frame.

In case of MPEG4 streams in LATM/LOAS format, the ancillary data described in this annex is placed into **data_stream_element()** (see ISO/IEC 14496-3 [17], table 4.10).

D.3 DVB Compliance

The ancillary data format described in this annex does not introduce any additional elements to the DVB transport stream. It is compliant with the current specification and compatible with all MPEG audio decoders.

Presence and type of ancillary data in audio elementary streams is signalled in DVB SI Program Map Table by the "Ancillary data descriptor" (see EN 300 468 [6], clause 6.2.2).

D.4 Detailed specification for MPEG1 and MPEG2

D.4.1 DVD-Video Ancillary Data

The transmission of "dynamic_range_control" in MPEG1 Layer I/II and MPEG2 Layer I audio is optional. *If applied, 16 bits of ancillary data [b15.b0] (situated at the end of each MPEG audio base frame) shall be used.*

Table D.1: DVD-Video ancillary data syntax

Syntax	No. of bits	Mnemonic
dvd_ancillary_data( ) {		
dynamic_range_control	8	bslbf
dynamic_range_control_on	1	bslbf
reserved (set to "000 0000b")	7	bslbf
}		

Semantics: The 8-bit dynamic_range_control field leads to the following gain control value by considering the upper 3 bits as unsigned integer X and the binary value of the lower 5 bits as unsigned integer Y:

- linear: $G = 2^{4-(X + Y/30)}$

  $(0 \le X \le 7, 0 \le Y \le 29)$

- in dB: $G = 24.082 - 6.0206\ X - 0.2007\ Y$

  $(0 \le X \le 7, 0 \le Y \le 29)$

If the dynamic_range_control_on field is set to "0b", the dynamic range_range_control field does not convey useful information.

Encoding: *When dynamic range control is temporarily not applied, that value of dynamic_range_control shall be set to "1000 0000b" or dynamic_range_control_on shall be set to "0b".*

Decoding: *The decoder shall read this field, and the decoder shall interpret the value G as a gain value applied to all sub band samples, before the reconstruction filter.* This value may be scaled in the decoder to allow user control of the amount of dynamic range compression that is applied.

D.4.2 Extended ancillary data syntax

The syntax of the extended ancillary data field is described in table D.2.

The extended ancillary data is inserted beginning from the end of the base frame. It is recommended that it be parsed from the end. The description in table D.2 is in the reverse order of the transmission. The bit order in each byte is, however, such that the msb comes first in the transmission.

Table D.2: Extended ancillary data syntax

Syntax	No. of bits	Mnemonic
extended ancillary_data( ) {		
dvd_ancillary_data	16	bslfb
extended_ancillary_data_sync (set to 0xBC)	8	bslfb
bs_info	8	bslbf
ancillary_data_status	8	bslbf
if(advanced_dynamic_range_control_status == 1)		
advanced_dynamic_range_control	24	bslbf
if(dialog_normalization_status == 1)		
dialog_normalization	8	bslbf
if(reproduction_level_status == 1)		
reproduction_level	8	bslbf
if(downmixing_levels_MPEG2_status == 1)		
downmixing_levels_MPEG2	8	bslbf
if(audio_coding_mode_and_compression_status == 1) {		
audio_coding_mode	8	bslbf
compression	8	bslbf
}		
if(coarse_grain_timecode_status == 1)		
coarse_grain_timecode	16	bslbf
if(fine_grain_timecode_status == 1)		
fine_grain_timecode	16	bslbf
if(scale_factor_CRC_status == 1)		
scale_factor_CRC	16 - 32	bslbf
}		

The elements of the ancillary data structure are described in the following clauses. The order of the bits is in transmission order, msb first.

D.4.2.1 ancillary_data_sync

Encoding: *This field shall be set to 0xBC.*

Decoding: The decoder may use this field to verify the availability of the extended ancillary data. If the IRD indicates that this information is present, this takes precedence.

D.4.2.2 bs_info

The detailed syntax is described in table D.3.

Table D.3: bs_info syntax

Syntax	No. of bits	Mnemonic
bs_info( ) {		
mpeg_audio_type	2	bslbf
dolby_surround_mode	2	bslbf
ancillary_data_bytes	4	uimsbf
}		

D.4.2.3 mpeg_audio_type

Table D.4: MPEG audio type Table

mpeg_audio_type	Description
"00"	Reserved
"01"	Only MPEG1 audio data
"10"	MPEG2 audio data
"11"	Reserved

Decoding: The decoder may ignore this field.

D.4.2.4 dolby_surround_mode

Table D.5: Dolby surround mode Table

mpeg_audio_type	Description
"00"	Reserved
"01"	MPEG1 part is not Dolby surround encoded
"10"	MPEG1 part is Dolby surround encoded
"11"	Reserved

Decoding: It is recommended that the decoder parse this field and provides this information to the reproduction set-up.

D.4.2.5 ancillary_data_bytes

This field indicates the amount of ancillary data bytes that precede this byte in the transmission. This field may be used by the decoder as an indication of how many bytes it needs to buffer.

D.4.2.6 ancillary_data_status

The detailed syntax is described on table D.6.

Table D.6: ancillary_data_status syntax

Syntax	No. of bits	Mnemonic
ancillary_data_status( ) {		
advanced_dynamic_range_control_status	1	bslbf
dialog_normalization_status	1	bslbf
reproduction_level_status	1	bslbf
downmix_levels_MPEG2_status	1	bslbf
scale_factor_CRC_status	1	bslbf
audio_coding_mode_and_compression status	1	bslbf
coarse_grain_timecode_status	1	bslbf
fine_grain_timecode_status	1	bslbf
}		

Semantics: The bits in this field indicate the presence of the associated fields in the ancillary data.

Encoding: *A bit in this field shall be set to "1" if the associated field is present in the bitstream.*

Decoding: It is recommended that the decoder parse this field to allow parsing of the following fields in the ancillary data section.

D.4.2.7 advanced_dynamic_range_control

The detailed syntax is described on table D.7.

Table D.7: advanced_dynamic_range_control syntax

Syntax	No. of bits	Mnemonic
advanced_dynamic_range_control( ) {		
advanced_drc_part_0	8	bslbf
advanced_drc_part_1	8	bslbf
advanced_drc_part_2	8	bslbf
}		

Semantics: Each field consists of an unsigned integer value X in the three msb's and an unsigned integer value Y in the five lsb's. The actual value is 24.082 - 6.0206 X - 0.2007 Y dB. The 1152 samples of an MPEG2 frame are divided in 3 parts of 384 samples. The advanced_drc values are applicable for the corresponding part of the audio frame.

Decoding: *If this field is present and the decoder supports this type of dynamic range control, these values shall be used rather than the DVD-Video ancillary data. The decoder shall apply these values to the sub band samples, before the reconstruction filter.* These values may be scaled in the decoder to allow user control of the amount of dynamic range compression that is applied.

D.4.2.8 dialog_normalization

The detailed syntax is described on table D.8.

Table D.8: dialog_normalization syntax

Syntax	No. of bits	Mnemonic
dialog_normalization( ) {		
dialog_normalization_on	2	bslbf
dialog_normalization_value	6	uimsbf
}		

D.4.2.9 dialog_normalization_on

Table D.9: Dialog normalization Table

dialog_normalization_on	Description
"00"	dialog_normalization_value is not valid
"01"	reserved
"10"	dialog_normalization_value is valid
"11"	Reserved

D.4.2.10 dialog_normalization_value

Semantics: This field represents the headroom in dB of the dialogue component in the MPEG1 compatible part, relative to full-scale sine wave. Values 41 through 63 are reserved. *When dialogue normalization is temporarily not applied, "Dialogue_Normalization_on" shall be set to "00" and "Dialog_Normalization_value" shall be set to "000000".*

Decoding: It is recommended that the decoder parse this field. The decoder should apply these values to the sub band samples, before the reconstruction filter, in order to allow reproduction of different programmes with the same dialogue level.

D.4.2.11 reproduction_level

The detailed syntax is described on table D.10.

Table D.10: reproduction_level syntax

Syntax	No. of bits	Mnemonic
reproduction_level ( ) {		
surround_reproduction_level	1	bslbf
production_roomtype	2	bslbf
reproduction_level_value	5	uimsbf
}		

D.4.2.12 surround_reproduction_level

Table D.11: Surround reproduction level Table

surround_reproduction_level	Description
"0"	The surround channels have the correct level for reproduction
"1"	The surround channels should be attenuated by 3 dB during reproduction

Decoding: It is recommended that the decoder parse this filed and pass the value to the reproduction unit to allow correct adjustment of the surround levels.

D.4.2.13 production_roomtype

Table D.12: Production room type Table

production_roomtype	Description
"00"	not indicated
"01"	large room
"10"	small room
"11"	reserved

Decoding: It is recommended that the decoder parse this field and pass the value to the reproduction unit to allow correct adjustment of the monitoring equipment.

D.4.2.14 reproduction_level_value

Semantics: This field represents the absolute acoustic sound pressure level in dB SPL during the final audio mixing session.

Decoding: The decoder may ignore this field.

D.4.2.15 downmixing_levels_MPEG2

The detailed syntax is described on table D.13. The down mixing levels describe the down mix in the decoder for stereo reproduction.

Table D.13: downmixing_levels_MPEG2 syntax

Syntax	No. of bits	Mnemonic
downmixing_levels_MPEG2 ( ) {		
center_mix_level_on	1	bslbf
center_mix_level_value	3	bslbf
surround_mix_level_on	1	bslbf
surround_mix_level_value	3	bslbf
}		

D.4.2.16 center_mix_level_on

Semantics: If this field is set to "1" the center_mix_value field indicates nominal down mix level of the centre channel with respect to the left and right front channels. *If this field is set to "0" the center_mix_value field shall be set to "000".*

Decoding: It is recommended that the decoder parse this field.

D.4.2.17 surround_mix_level_on

Semantics: If this field is set to "1" the surround_mix_value field indicates nominal down mix level of the surround channels with respect to the left and right front channels. *If this field is set to "0" the surround_mix_value field shall be set to "000".*

Decoding: It is recommended that the decoder parse this field.

D.4.2.18 mix_level_value

Table D.14: Mix level value Table

mix_level_value	Multiplication factor
"000"	1.000 (0.0 dB)
"001"	0.841 (-1.5 dB)
"010"	0.707 (-3.0 dB)
"011"	0.596 (-4.5 dB)
"100"	0.500 (-6.0 dB)
"101"	0.422 (-7.5 dB)
"110"	0.355 (-9.0 dB)
"111"	0.000 (-∞ dB)

Decoding: The multi-channel decoder may apply these values as gain factors to the individual channels when a down mix for stereo listening has to be created. The values need to be scaled to avoid overload after the mixing process.

D.4.2.19 audio_coding_mode

The detailed syntax is described in table D.15.

Table D.15: audio coding mode syntax

Syntax	No. of bits	Mnemonic
audio_coding_mode ( ) {		
MPEG2_extension_stream_present	1	bslbf
MPEG2_center	2	bslbf
MPEG2_surround	2	bslbf
MPEG2_lfeon	1	bslbf
MPEG2_copyright_ident_present	1	bslbf
compression_on	1	bslbf
}		

Semantics: The semantics of the fields MPEG2_extension_stream_present, MPEG2_center, MPEG2_surround and MPEG2_lfeon is as defined in the mc_header field in [3].

If MPEG2_copyright_ident_present is set to "0" the copyright identification in the MPEG 2 mc_header is not filled in. If MPEG2_copyright_ident_present is set to "1" the copyright identification in the MPEG 2 mc_header is used.

Decoding: The decoder may ignore this field. It may be parsed be multiplexers and bitstream monitors to simplify extraction of these parameters from a bitstream.

D.4.2.20 compression_on

Semantics: If this field is set to "1" the compression_value field indicates the heavy compression factor used for monophonic down mix reproduction. *If this field is set to "0" the compression_value field shall be "0000 0000".*

Decoding: It is recommended that the decoder parse this field.

D.4.2.21 compression_value

Semantics: This field consists of a value X in the four msb's and a value Y in the four lsb's. The actual value is 48.164 - 6.0206 X - 0.4014 Y dB.

Decoding: *These values shall be applied to the sub band samples, before the reconstruction filter when the decoder has to create a mix for monophonic listening where overloading of a subsequent analog transmission is highly undesirable.*

D.4.2.22 coarse_grain_timecode

The detailed syntax is described on table D.16.

Table D.16: coarse grain time code syntax

Syntax	No. of bits	Mnemonic
coarse_grain_timecode ( ) {		
coarse_grain_timecode_on	2	bslbf
coarse_grain_timecode_value	14	bslbf
}		

Semantics: If coarse_grain_timecode_on is set to "10" the five msb's of this value represents the time in hours, the next six bits represent time in minutes, and the final three bits represent the time in eight second increments. *If coarse_grain_timecode_on is not set to "10" all the bits of coarse_grain_timecode_value shall be set to "0".*

Decoding: The decoder may ignore this field.

D.4.2.23 fine_grain_timecode

The detailed syntax is described in table D.17.

Table D.17: fine grain time code syntax

Syntax	No. of bits	Mnemonic
fine_grain_timecode ( ) {		
fine_grain_timecode_on	2	bslbf
fine_grain_timecode_value	14	bslbf
}		

Semantics: If fine_grain_timecode_on is set to "10" the three msb's of this value represents the time in seconds, the next five bits represent time in video frames, and the final six bits represent the time in fractions of 1/64 of a video frame. *If fine_grain_timecode_on is not set to "10" all the bits of fine_grain_timecode_value shall be set to "0".*

Decoding: The decoder may ignore this field.

D.4.2.24 scale_factor_CRC

Semantics: The scale_factor CRC permits to verify the integrity of the MPEG Audio scale factors. The coding is according to [20].

Encoding: It recommended that scale_factor_CRC be included for mobile applications.

Decoding: It is recommended to parse the data from the end. The length of the field depends on the bit rate index of the MPEG 1 header of the following frame. It is recommended to always parse the full 32 possible bits.

D.4.3 Announcement Switching Data

The transmission of announcement switching data in the ancillary data field of MPEG audio frames is optional. The syntax of the announcement switching data field is described in table D.18. Note that the description in table D.18 is in the reverse order of the transmission. The bit order in each byte is, however, such that the msb comes first in the transmission. The data field length gives the number of bytes following this byte within this data field.

Table D.18: Announcement switching data field

Syntax	No. of bits	Mnemonic
announcement_switching_data( ) {		
announcement_switching_data_sync	8	bslbf
data_field_length	8	bslbf
announcement_switching_flag_field_1	16	bslbf
announcement_switching_flag_field_2	16	bslbf
}		

Semantics: The announcement_switching_data_sync should be set to 0 x AD.

The announcement_switching_flag_fields are 16-bit flag fields specifying which type of announcements are actually running. *The association between the bits of the flag field and the announcement types shall be according to the announcement_support_indicator [6]. A bit shall be set to "1" if the announcement is running and it shall be set to "0" if the announcement is not running.*

The announcement_switching_flag_field_1 shall be used for announcements within the audio elementary stream that is actually decoded.

The announcement_switching_flag_field_2 shall be used for announcements within other audio elementary streams. Corresponding links shall be provided by means of the announcement_support_descriptor [6].

Encoding: The announcement_switching_data_field is allowed to be embedded at the end of a MPEG audio packet, between the end of the audio data and another data field that is part of the ancillary data field or between two other data fields that are part of the ancillary data field.

If data fields according to DVD-video, extended ancillary data or ancillary data according to the DAB specification [19] are used, then the announcement_switching_data_field is not allowed to be inserted at the end of an audio packet.

Decoding: It is recommended to parse the data from the end.

D.4.4 Scale Factor Error Check

The transmission of a scale factor error check in the ancillary data field of MPEG audio frames is optional. The syntax of the corresponding data field is described in table D.19. Note that the description in table D.19 is in the reverse order of the transmission. The bit order in each byte is, however, such that the msb comes first in the transmission. The data_field_length gives the number of bytes following this byte within this data field.

Table D.19: Scale factor error check data field

Syntax	No. of bits	Mnemonic
scale_factor_error_check_data( ) {		
scale_factor_error_check data_sync	8	bslbf
data_field_length	8	bslbf
scale factor CRC	32	bslbf
}		

Semantics: The scale_factor_error_check data_sync should be set to 0 x FE.

The scale_factor CRC permits to verify the integrity of the MPEG Audio scale factors.

Encoding: The scale_factor_error_check is allowed to be embedded at the end of a MPEG audio packet, between the end of the audio packet and another data field that is part of the ancillary data field or between two other data fields that are part of the ancillary data field.

If data fields according to DVD-video, extended ancillary data (as described in annex D) or ancillary data according to the DAB specification EN 300 401[19] are used, then the scale_factor_error_check_data_field is not allowed to be inserted at the end of an audio packet.

Decoding: It is recommended to parse the data from the end.

D.5 Detailed specification for MPEG4

D.5.1 Transmission of MPEG4 ancillary data

*Presence of MPEG4 ancillary data shall be signalled in DVB SI by setting b_5 in **ancillary_data_identifier** to "1" (see EN 300 468 [6], table 16).*

*MPEG4 ancillary data as defined in this annex shall be placed into a single **data_stream_element()** as defined in ISO/IEC 14496-3, table 4.10 [17].*

*The **data_stream_element()** <DSE> shall follow any combination of related <SCE>, <CPE>, <LFE>, and <FIL <EXT-SBR_DATA>> audio elements, to which the ancillary data applies.*

*The **element_instance_tag** of this **data_stream_element()** shall have the same value as the **element_instance_tag** of the first audio element to which the ancillary data applies.*

Examples of possible streams are:

for a 2-channel program:
<CPE><DSE><FIL><TERM><CPE><DSE><FIL><TERM>…

for a 2-channel program with SBR:
<CPE><SBR(CPE)><DSE><FIL><TERM><CPE><SBR(CPE)><DSE><FIL><TERM>…

for a 5.1-channel program
<SCE><CPE><CPE><LFE><DSE><FIL><TERM><SCE><CPE><CPE><LFE><DSE><FIL><TERM>…

For further reference see clauses 4.5.2.1.2 and 4.5.2.9.2 in ISO/IEC 14496-3 [17].

D.5.2 MPEG4 ancillary data syntax

The syntax of the ancillary data field is described in table D.20. Data are transmitted in the order as given in the table.

Table D.20: MPEG4 ancillary data syntax

Syntax	No. of bits	Mnemonic
MPEG4 ancillary_data( ) {		
ancillary_data_sync	8	bslfb
bs_info	8	bslbf
ancillary_data_status	8	bslbf
If (downmixing_levels_MPEG4_status == 1)		
downmixing_levels_MPEG4	8	bslbf
If (audio_coding_mode_and_compression_status == 1) {		
audio_coding_mode	8	bslbf
Compression_value	8	bslbf
}		
if(coarse_grain_timecode_status == 1)		
coarse_grain_timecode	16	bslbf
if(fine_grain_timecode_status == 1)		
fine_grain_timecode	16	bslbf
}		

D.5.2.1 ancillary_data_sync

Encoding: *This field shall be set to 0xBC.*

Decoding: The decoder may use this field to verify the availability of the MPEG4 ancillary data.

D.5.2.2 bs_info

The detailed syntax is described in table D.21.

Table D.21: bs_info syntax

Syntax	No. of bits	Mnemonic
bs_info( ) {		
mpeg_audio_type	2	bslbf
dolby_surround_mode	2	bslbf
reserved, set to "0000"	4	bslbf
}		

D.5.2.2.1 mpeg_audio_type

Table D.22: MPEG audio type Table

mpeg_audio_type	Description
"00"	Reserved
"01"	Reserved
"10"	Reserved
"11"	MPEG4 Audio data

Encoding: *This field shall be set according to table D.22.*

Decoding: The decoder may ignore this field.

D.5.2.2.2 dolby_surround_mode

Table D.23: Dolby surround mode Table

mpeg_audio_type	Description
"00"	Dolby surround mode not indicated
"01"	2-ch audio part is not Dolby surround encoded
"10"	2-ch audio part is Dolby surround encoded
"11"	Reserved

Semantics: In case of 2-channel audio streams it can be indicated, whether the audio signal is encoded in Dolby surround mode.Encoding: This field may be provided by encoders when the audio stream is in 2-channel (stereo) format. *It shall be set to "00" for other than 2-channel audio streams.*

Decoding: It is recommended that the decoder parses this field and provides this information to the reproduction set-up.

D.5.2.3 ancillary_data_status

The detailed syntax is described on table D.24.

Table D.24: ancillary_data_status syntax

Syntax	No. of bits	Mnemonic
ancillary_data_status( ) {		
Reserved, set to "0"	1	bslbf
Reserved, set to "0"	1	bslbf
Reserved, set to "0"	1	bslbf
downmixing_levels_MPEG4_status	1	bslbf
Reserved, set to "0"	1	bslbf
audio_coding_mode_and_compression status	1	bslbf
coarse_grain_timecode_status	1	bslbf
fine_grain_timecode_status	1	bslbf
}		

Semantics: The bits in this field indicate the presence of the associated fields in the ancillary data.

Encoding: *A bit in this field shall be set to "1" if the associated field is present in the bitstream.*

Decoding: It is recommended that the decoder parse this field to allow parsing of the following fields in the ancillary data section.

D.5.2.4 downmixing_levels_MPEG4

When multichannel audio streams are decoded by an IRD and only 2-channel audio output is required, then matrix mix down has to be applied. For MPEG-4 AAC and MPEG-4 HE AAC matrix mix down is described in ISO/IEC 14496-3 [17].

This part of MPEG-4 ancillary data gives a possibility to transmit matrix mix down coefficients with higher resolution than defined in ISO/IEC 14496-3 [17]. The detailed syntax is described in table D.25.

Table D.25: downmixing_levels_MPEG4 syntax

Syntax	No. of bits	Mnemonic
downmixing_levels_MPEG4 ( ) {		
center_mix_level_on	1	bslbf
center_mix_level_value	3	bslbf
surround_mix_level_on	1	bslbf
surround_mix_level_value	3	bslbf
}		

Encoding: This matrix mix down information may be supplied by the encoder.

Decoding: It is recommended that the decoder parses this field and uses the information in cases matrix mix down is needed.

D.5.2.4.1 center_mix_level_on

Semantics: This field indicates, whether the **center_mix_value** field carries information for matrix mix down.

Encoding: *If this field is set to "1" the **center_mix_value** field shall indicate the matrix mix down level of the centre channel with respect to the left and right front channels. If this field is set to "0" the **center_mix_value** field shall be set to "000".*

Decoding: It is recommended that the decoder parse this field.

D.5.2.4.2 surround_mix_level_on

Semantics: This field indicates, whether the **surround_mix_value** field carries information for matrix mix down.

Encoding: *If this field is set to "1" the surround_mix_value shall indicate the matrix mix down level of the surround channels with respect to the left and right front channels. If this field is set to "0" the **surround_mix_value** field shall be set to "000".*

Decoding: It is recommended that the decoder parse this field.

D.5.2.4.3 mix_level_value

Table D.26: Mix level value Table

mix_level_value	Multiplication factor
"000"	1.000 (0.0 dB)
"001"	0.841 (-1.5 dB)
"010"	0.707 (-3.0 dB)
"011"	0.596 (-4.5 dB)
"100"	0.500 (-6.0 dB)
"101"	0.422 (-7.5 dB)
"110"	0.355 (-9.0 dB)
"111"	0.000 (-∞ dB)

Encoding: When provided, the values of **center_mix_level**_value and **surround_mix_level**_value shall be set to indicate the multiplication factors for 2-channel matrix mix down.

Decoding: The multi-channel decoder may apply these values as gain factors to the individual channels when a down mix for 2-channel stereo listening has to be created. The values need to be scaled to avoid overload after the mixing process.

D.5.2.5 audio_coding_mode

The detailed syntax is described in table D.27.

Table D.27: audio coding mode syntax

Syntax	No. of bits	Mnemonic
audio_coding_mode ( ) {		
reserved, set to "000 0000"	7	bslbf
compression_on	1	bslbf
}		

Decoding: It is recommended that the decoder parse this field.

D.5.2.5.1 compression_on

Semantics: This field indicates, whether the **compression_value** field carries information.

Encoding: If this field is set to "1" the **compression_value** field indicates the heavy compression factor used for monophonic down mix reproduction. *If this field is set to "0" the **compression_value** field shall be "0000 0000".*

Decoding: It is recommended that the decoder parse this field.

D.5.2.5.2 compression_value

Semantics: This field consists of a value X in the four msb's and a value Y in the four lsb's. The actual value is 48.164 - 6.0206 X - 0.4014 Y dB.

Encoding: The encoder may provide this information.

Decoding: *When available, the IRD shall apply these values to the spectral samples, before the reconstruction transform, when the decoder has to create a mix for monophonic listening where overloading of a subsequent analog transmission is highly undesirable.*

D.5.2.6 coarse_grain_timecode

See clause D.4.2.22.

D.5.2.7 fine_grain_timecode

See clause D.4.2.23.

D.5.3 Announcement Switching Data

The transmission of announcement switching data in MPEG4 ancillary data is optional. The syntax of the announcement switching data field is described in table D.28.

Table D.28: Announcement switching data field

Syntax	No. of bits	Mnemonic
announcement_switching_data( ) {		
announcement_switching_data_sync	8	bslbf
data_field_length	8	bslbf
announcement_switching_flag_field_1	16	bslbf
announcement_switching_flag_field_2	16	bslbf
}		

Semantics: The **announcement_switching_data_sync** should be set to 0xAD.
The **data_field_length** gives the number of bytes following this byte within this data field.

The **announcement_switching_flag_fields** are 16-bit flag fields specifying which type of announcements are actually running. *The association between the bits of the flag field and the announcement types shall be according to* ***the announcement_support_indicator [***6***]****. A bit shall be set to "1" if the announcement is running and it shall be set to "0" if the announcement is not running.*

The announcement_switching_flag_field_1 *shall be used for announcements within the audio elementary stream that is actually decoded.*

The announcement_switching_flag_field_2 *shall be used for announcements within other audio elementary streams. Corresponding links shall be provided by means of the* ***announcement_support_descriptor [***6***]****.*

Decoding: It is recommended that the decoder parse this field.

Annex E (informative): Coding of Data Fields in the Private Data Bytes of the Adaptation Field

E.1 Introduction

This annex contains the guidelines required to include and to decode data fields in the private data bytes of the adaptation field [1].

E.2 Detailed specification

Transport stream (TS) packets coded according to ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1] may include an adaptation field. The presence of an adaptation field is indicated by means of the adaptation_field_control, i.e. a 2-bit field in the header of the TS packet. The adaptation field itself may contain private_data_bytes. The presence of private data bytes is signalled by means of the transport_private_data_flag coded at the beginning of the adaptation field. If private data bytes exist the total number of private data bytes is specified by means of the transport_private_data_length, an 8-bit field that is directly followed by the private data bytes. The private data bytes may be composed of one or more data fields as shown in figure E.1. Gaps are not allowed between two data fields.

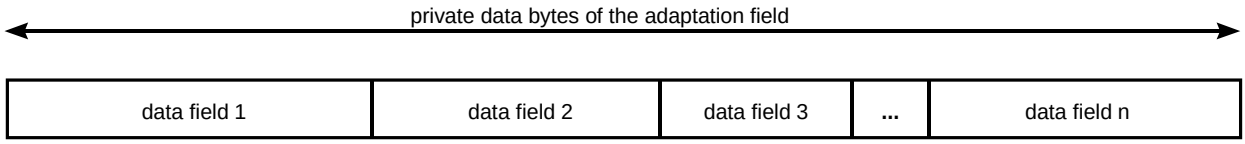


Figure E.1: Coding scheme for private data bytes within the adaptation field

Encoding: *The support of data fields that are specified in this annex shall be indicated by means of the adaptation_field_data_descriptor [7]. This descriptor shall be inserted in the corresponding ES_info loop.*

Moreover, the following semantics apply to all data fields specified in this annex.

data_field_tag: The data field tag is an 8-bit field which identifies the type of each data field. The values of data_field_tag are defined in table E.1.

data_field_length: The data field length is an 8-bit field specifying the total number of bytes of the data portion of the data field following the byte defining the value of this field.

Table E.1: Allocation of data_field_tags

data_field_tag	Description
0x00	Reserved
0x01	Announcement switching data field
0x02	AU_information data field
0 x 03 to 0 x 9F	Reserved for future use
0 x A0 to 0 x FF	User defined

Decoding: The IRD design should be made under the assumption that any structure as permitted by this annex may occur in the broadcast stream. The IRD is not required to make use of this data.

E.2.1 Announcement Switching Data

The announcement switching data field is used to indicate whether spoken announcements are actually running or not. In comparison with that, the general support of announcements is indicated by means of the announcement_support_descriptor [7].

The transmission of the announcement switching data field is optional but it shall be continuously provided in those audio streams that may carry announcements at some point in time. The announcement switching data field shall be present at least every 100 ms. The syntax of the announcement switching data field is described in table E.2.

Table E.2: Announcement switching data field

Syntax	No. of bits	Mnemonic
announcement_switching_data( ) {		
data_field_tag	8	uimsbf
data_field_length	8	uimsbf
announcement_switching_flag_field	16	bslbf
}		

Semantics: announcement_switching_flag_field: This 16-bit flag field specifies which type of announcements are actually running. *The association between the bits of the flag field and the announcement types shall be according to the announcement_support_indicator that is specified for the announcement_support_descriptor [7]. A bit shall be set to "1" if the announcement is running and it shall be set to "0" if the announcement is not running.*

E.2.2 AU_information

The AU_information data field is used to signal the presence of the start of an access unit in the payload of the transport packet containing the data field, and to convey information about that access unit that is of use to PVR applications. All the information provided in this descriptor should be considered "helper" information rather than definitive information. *Thus, if there are any conflicts between the information signalled in this descriptor and the actual stream, then the information in the stream shall take precedence over the information in this descriptor.* However, such a conflict should be considered an error condition and as such should not occur. It is recommended that the AU_information data field is present at the start of each access unit of an H.264 | ISO/IEC 14496-10 [16] video streams.

Where multiple access units occur in a transport packet, then multiple AU_information data fields may be used. *Each descriptor shall apply to the corresponding access unit in the transport packet. I.e. the first data field shall apply to the first access unit starting in the transport packet, the second data field shall apply to the second access unit starting in the transport packet, etc.*

The AU_information data field(s), when present, shall be the first data field(s) in the adaptation field.

There shall not be more descriptors than there are access units starting in the packet.

The presence of AU_information data fields shall be indicated via bit b_1 of the adaptation_field_data_identifier in the adaptation field descriptor.

Table E.3: AU_information data field

Syntax	No. of Bits	Mnemonic
AU_information () {		
data_field_tag	8	uimsbf
data_field_length	8	uimsbf
AU_coding_format	4	uimsbf
AU_coding_type_information	4	bslbf
AU_ref_pic_idc	2	uimsbf
AU_pic_struct	2	bsblf
AU_PTS_present_flag	1	bslbf
AU_profile_info_present_flag	1	bslbf
AU_stream_info_present_flag	1	bslbf
AU_trick_mode_info_present_flag	1	bslbf
if (AU_PTS_flag == "1") {		
AU_PTS_32	32	uimsbf
}		
if (AU_steam_info_flag == "1") {		
Reserved	4	"0000"
AU_frame_rate_code	4	uismbf
}		
if (AU_profile_info_flag == "1") {		
AU_profile_idc	8	uismbf
AU_constraint_set0_flag	1	bslbf
AU_constraint_set1_flag	1	bslbf
AU_constraint_set2_flag	1	bslbf
AU_AVC_compatible_flags	5	bslbf
AU_level_idc	8	uismbf
}		
if (AU_trick_mode_info_present_flag == "1") {		
AU_max_I_picture_size	12	uismbf
AU_nominal_I_period	8	uismbf
AU_max_I_period	8	uismbf
Reserved	4	"0000"
}		
for(i=0; i<n; i++) {		
AU_reserved_byte	8	bslbf
}		
}		

Semantics: ***data_field_tag****: this shall have the value 0x02.*

data_field_length: this indicates the length of the descriptor. The values 0 and 1 may be used to signal short versions of the descriptor. The value 0 means that no fields after the data_field_length are sent, and is used as a dummy descriptor. The value 1 means that only the fields AU_coding_format and AU_coding_type_information are present.

AU_coding_format: This shall signal the coding format used by the elementary stream carried on this packet. The values are as show in table E.4.

Table E.4: AU_coding_format values

Value	Stream Type
0	Undefined
1	ITU-T Rec H.262 | ISO/IEC 13818-2 [2] Video or ISO/IEC 11172-1 [8]constrained parameter video stream
2	AVC video stream as defined in ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16] Video
3-0xF	reserved

AU_coding_type_information: i*ndicates the elementary stream types present in the immediately following access unit. For ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16] video, this field shall be interpreted as a four bit field with the syntax shown in table E.5.*

Table E.5: AU_coding_type_information for ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16] video

Syntax	No. of Bits	Mnemonic
AU_IDR_slice_present_flag	1	bslbf
AU_I_slice_present_flag	1	bslbf
AU_P_slice_present_flag	1	bslbf
AU_B_slice_present_flag	1	bslbf

For ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] Video, this field shall be interpreted according to table E.6. These values are identical to (but one bit longer than) the values in table 6-12 of ISO/IEC 13818-2 [2].

Table E.6: AU_coding_type_information for ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] video

Value	AU_coding_type_information
0	Undefined
1	I
2	P
3	B
4-0xF	Reserved

AU_ref_pic_idc: This field indicates if any of the access unit is required in the reconstruction of other access units. The value "00" means that it is not used by other access units. *In the case of ITU-T Recommendation H.264 | ISO/IEC 14496-10 [*16*], the value shall be the nal_ref_idc field in the NAL header used for any slice that makes up the access unit.*

AU_pic_struct: *This field shall be set to "01" if the access unit is a top field picture, "10" if it is a bottom field. Otherwise, it shall be set to "00". "11" value is reserved.*

AU_PTS_present_flag: *This field shall be set to "1" when the AU_PTS_32 value is present in the descriptor, otherwise it shall take the value "0".*

AU_profile_info_present_flag*: This field shall be set to "1" when the AU_profile_idc and AU_level_idc values are present in the descriptor, otherwise it shall take the value "0".*

AU_stream_info_present_flag*: This field shall be set to "1" when the AU_frame_rate_code value is present in the descriptor, otherwise it shall take the value "0".*

AU_trick_mode_info_present_flag*: This field shall be set to "1" when the AU_max_I_picture_size and AU_max_I_period are present in the descriptor.*

AU_PTS_32: the 32 most significant bits of the 33-bit PTS encoded in the PES header immediately following this adaptation field, or of the value that applies to the access unit to which this descriptor applies, if no PES header is present.

AU_profile_idc: this field conveys the profile used to which the access unit conforms. For ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16] video this carries the profile_idc value as defined ISO/IEC 14496-10 [16], annex A. For ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] video the least significant 3 bits of this field carry the profile as defined in clause 8 of ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2].

AU_level_idc: this field conveys the level used to which the access unit conforms. For ITU-T Recommendation H.264 | ISO/IEC 14496-10 [16] video this carries the level_idc value as defined ISO/IEC 14496-10 [16], annex A. For ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2] video the least significant 4 bits of this field carry the level as defined in clause 8 of ITU-T Recommendation H.262 | ISO/IEC 13818-2 [2].

Constraint_set0_flag, constraint_set1_flag, constraints_set2_flag, AVC compatible flags: These fields carry the same semantics as the fields of the same name in the AVC_video_descriptor in clause 2.6.54 of ISO/IEC 13818-1: [1] 2000 (AMD3), which in turn have semantics defined in ISO/IEC 14496-10 [21], clause 7.4.2.1. Note that with High profile, the first bit in AVC_compatible_flags carries constraint_set3_flag.

AU_frame_rate_code: this field indicates the video frame rate in the stream carried on the current PID. In the case of video, this is encoded as in clause 6.3.3 of ISO/IEC 13818-2 [2]:2000, as shown in table 6-4 of the same. The values in this table are informatively replicated on table E.8.

Table E.7: Informative Frame Rate values taken from table 6-4 of 13818-2:2000

AU_frame_rate_code	Corresponding Frame Rate (Hz)
0	Forbidden
1	23.976
2	24
3	25
4	29.97
5	30
6	50
7	59.94
8	60
9 to 0 x F	Reserved

AU_max_I_picture_size: this value indicates the buffer size, in units of 16x1024 bits, that is implemented by the encoder rate control, and thus the maximum intra picture size that can be found in the current bitstream. This value, according to profile and level, shall comply with ISO/IEC 14496-10 [21] and ISO/IEC 13818-2 [2] limits. The value 0 is forbidden.

AU_nominal_I_period: this value indicates the nominal distance between two consecutive I/IDR pictures, on a frame picture count basis. The value 0 is forbidden.

AU_max_I_period: this value indicates the maximum distance that can be found in the stream between two consecutive I/IDR pictures, on a frame picture count basis. The value 0 is forbidden.

Annex F (informative): Guidelines for the Implementation of DTS Coded Audio in DVB Compliant Transport Streams

F.1 Scope

The inclusion of DTS coded audio streams in a DVB multiplex is optional, and IRDs may optionally decode these streams. This annex contains the guidelines to include one or more DTS coded elementary streams in a DVB Transport Stream in compliance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1]. The coding and decoding of a DTS coded elementary stream is based upon TS 102 114 [15].

It is recommended that implementations of DVB systems that include DTS coded audio streams should comply with this annex.

The DTS packetized elementary stream shall conform to the requirements of a user private stream type 1, as described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].

The IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], including private data streams, may occur in the Transport Stream, even if presently reserved or unused. *To allow full compliance to the MPEG-2 standard and upward compatibility with future enhanced versions, a DVB IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the IRD.*

This clause is based on ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].

F.2 Introduction

A DTS coded elementary bitstream may be multiplexed into an MPEG-2 transport stream in much the same way an MPEG-1 or AC-3 audio stream would be included. The DTS coded elementary stream is packetized into PES packets with a structure similar to an MPEG audio PES.

It is necessary to unambiguously indicate that an MPEG private stream is, in fact, a DTS coded stream. A public DVB descriptor, the DTS_audio_descriptor will be specified for this purpose and is defined as 0x73. The DTS registration_descriptor outlined in table 1 must also be specified. The syntactical elements that need to be specified in order to include DTS within an MPEG-2 transport stream are: the MPEG stream_type, stream_id and the DVB DTS_audio_descriptor.

IRDs shall decode all bit rates and sample rates listed herein.

Some constraints are placed on the PES layer for the case of multiple audio streams intended to be reproduced in exact sample synchronism as described in clause 5.

F.3 DVB Compliant Streams

The DTS PES shall be carried as an MPEG private data stream type, conforming to the structure of a private_stream_1 as described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], table 2-18 (stream_id) and table 2-29 (stream_type).

When a DTS stream is included in a DVB transport stream, the DTS Audio descriptor (DTS_audio_descriptor) shall also be included. The DTS Audio descriptor is defined in annex F of EN 300 468 [6], but for information a description is included here in clause 4.3. Either the DTS Audio Descriptor or the DTS registration descriptor must be located in the PMT to identify the DTS stream as such; similarly one of DTS Audio Descriptor or DTS registration descriptor must be located in the SIT. The DTS Audio descriptor is located in the PMT and the Selection Information Table of the DVB SI Tables in annex F of EN 300 468 [6].

DTS streams may also be signalled by the presence of a component_descriptor where the stream_content value is 0x05 (see EN 300 468 [6], clause 6.2.7) in the relevant service information tables.

Certain other of the DVB Service Information in EN 300 468 [6] can provide additional means of identifying the existence of a DTS stream without accessing the PMT.

F.4 Detailed specification

F.4.1 MPEG Transport Stream Compliance

F.4.1.1 stream_id

Semantics: The semantics of the stream_id field are described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], table 2-18. Multiple DTS streams may share the same value of stream_id since each stream is carried with a unique PID value. The mapping of values of PID to stream_type is indicated in the transport stream programme map Table (PMT).

Encoding: *The value of the stream_id field for a DTS elementary stream shall be 0xBD (indicating private_stream_1). If multiple DTS elementary streams are carried in a program stream the stream_id shall use values 110x xxxx where x xxxx indicates a stream number.* Confusion may be avoided by use of a Program Stream Map, which associates values of a stream_id with a stream_type.

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with MPEG systems syntax.*

F.4.1.2 stream_type

Semantics: The semantics of the stream_type field are described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].

Encoding: *The recommended value of stream_type for a DTS elementary stream shall be 0x06 (indicating PES packets containing private data) or any value which the MPEG-2 specification has assigned as "user private".*

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with MPEG systems syntax.*

F.4.2 DTS Registration descriptor

The DTS registration descriptor is shown in table F.1. It is mandatory that the IRD decodes of the registration descriptor so that the stream is clearly identified as carrying DTS data.

Table F.1: DTS registration descriptor

Syntax	Number of Bits	Mnemonic
Registration_descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
format_identifier	32	uimsbf
}		

F.4.2.1 descriptor_tag

Encoding: The registration descriptor tag is an 8-bit field, which identifies each descriptor. The value assigned to the DTS descriptor_tag is 0x05.

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

F.4.2.2 descriptor_length

Semantics: This 8-bit field specifies the total number of byes of the data portion of the registration descriptor following the byte defining the value of this field. The value assigned to the DTS registration descriptor_length is 0x04.

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

F.4.2.3 format_identifier

Encoding: The SMPTE registered format identifier sets the frame size for the DTS coded stream and is set according the values as follows;

- DTS format identifier is 0x44545331 ("DTS1") for frame size 512;
- DTS format identifier is 0x44545332 ("DTS2") for frame size 1 024;
- DTS format identifier is 0x44545333 ("DTS3") for frame size 2 048.

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

F.4.3 DTS Audio Descriptor

The DTS audio descriptor is shown in table F.2. It is optional that the IRD decodes the DTS audio descriptor.

Table F.2: DTS Audio Descriptor

Syntax	Number of Bits	Mnemonic
DTS_audio_stream_descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
sample_rate_code	4	bslbf
bit_rate_code	6	bslbf
nblks	7	bslbf
fsize	14	uimsbf
surround_mode	6	bslbf
lfe_flag	1	uimsbf
extended_surround_flag	2	uimsbf
for(i=0;i<N;i++)	8*N	
{		bslbf
additional_info[N]		
}		
}		

F.4.3.1 descriptor_tag

Encoding: The audio descriptor tag is an 8-bit field, which identifies each descriptor. The proposed value assigned to the audio descriptor_tag is defined as 0x73.

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

F.4.3.2 descriptor_length

Semantics: This 8-bit field specifies the total number of byes of the data portion of the audio descriptor following the byte defining the value of this field.

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].*

F.4.3.3 sample_rate_code

Semantics: This 4-bit field is equivalent to SFREQ in DTS Coherent Acoustics. Specification and details are listed in table F.3. While broadcasters may use only a subset of these the complete table is given for consistency with the DTS Coherent Acoustics specification as defined in TS 102 114 [15].

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended IRDs decode this field.*

Table F.3: Sample Rate Code

sample_rate_code	Sample Rate
0b0000	Invalid
0b0001	8 kHz
0b0010	16 kHz
0b0011	32 kHz
0b0100	64 kHz
0b0101	128 kHz
0b0110	11,025 kHz
0b0111	22,05 kHz
0b1000	44,.1 kHz
0b1001	88,02 kHz
0b1010	176,4 kHz
0b1011	12 kHz
0b1100	24 kHz
0b1101	48 kHz
0b1110	96 kHz
0b1111	192 kHz

F.4.3.4 bit_rate_code

Semantics: The specification and details of typical broadcast bit_rate_code are listed in table F.4. While broadcasters may use only a subset of these, the complete table of fixed transmission bit rate values is given for consistency with the DTS Coherent Acoustics specification as defined in TS 102 114 [15]. Note, it is recommended that DTS 5.1 compressed audio streams be transmitted at data rate of 384 kpbs or above.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended IRDs decode this field.*

Table F.4: Bit Rate Table

bit_rate_code	Transmission bit rate
0bx00101	128 kbps
0bx00110	192 kbps
0bx00111	224 kbps
0bx01000	256 kbps
0bx01001	320 kbps
0bx01010	384 kbps
0bx01011	448 kbps
0bx01100	512 kbps
0bx01101	576 kbps
0bx01110	640 kbps
0bx01111	768 kbps
0bx10000	960 kbps
0bx10001	1 024 kbps
0bx10010	1 152 kbps
0bx10011	1 280 kbps
0bx10100	1 344 kbps
0bx10101	1 408 kbps
0bx10110	1 411,2 kbps
0bx10111	1 472 kbps
0bx11000	1 536 kbps
0bx11001	1 920 kbps
0bx11010	2 048 kbps
0bx11011	3 072 kbps
0bx11100	3 840 kbps
0bx11101	open
0bx11110	variable
0bx11111	lossless
NOTE: "x" indicated the bit is reserved and should be ignored.	

F.4.3.5 nblks

Semantics: This 7-bit word is equivalent to NBLKS in listed in TS 102 114 [15]. This equals the number of PCM Sample Blocks. It indicates that there are (NBLKS + 1) blocks (a block = 32 PCM samples per channel, corresponding to the number of PCM samples that are fed to the filterbank to generate one subband sample for each subband) in the current frame. The actual encoding window size is 32*(NBLKS + 1) PCM samples per channel. Valid range: 5 to 127. Invalid range: 0 to 4. For normal frames, this indicates a window size of either 2 048, 1 024, or 512 samples per channel. For termination frames, NBLKS can take any value in its valid range.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended IRDs decode this field.*

F.4.3.6 Fsize

Semantics: This 14-bit word is equivalent to FSIZE listed in TS 102 114 [15]. (FSIZE + 1) is the byte size of the current primary audio frame. The valid range for fsize is 95 - 8192. The invalid range for fsize is 0 - 94, 8193 - 16384.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended IRDs decode this field.*

F.4.3.7 surround_mode

Semantics: This 6-bit word is equivalent to AMODE in DTS Coherent Acoustics Specification. The values for surround_mode are given in table F.5. While broadcasters may use only a subset of these the complete table is given for consistency in TS 102 114 [15], table 5.4.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field. It is recommended IRDs decode this field.*

Table F.5: Surround Mode

Surround_mode	Number of Channels/Channel Layout
0b000000	1 / mono
0b000010	2 / L + R (stereo)
0b000011	2 / (L+R) + (L-R) (sum-difference)
0b000100	2 / LT +RT (left and right total)
0b000101	3 / C + L + R
0b000110	3 / L + R+ S
0b000111	4 / C + L + R+ S
0b001000	4 / L + R+ SL+SR
0b001001	5 / C + L + R+ SL+SR
0b001010	User defined
0b001011	User defined
0b001100	User defined
0b001101	User defined
0b001110	User defined
0b001111	User defined
0b010000 - 0b111111	User defined
NOTE: Legends: L =left, R = right, C =centre, SL = surround left, SR = surround right, T = total.	

F.4.3.8 lfe_flag

Semantics: *The lfe flag shall be set to 0 when the LFE (Low Frequency Effects) audio channel is OFF. The flag shall be set to 1 when the LFE audio channel is ON.*

Decoding: *IRDs shall be able to accept bit-streams, which contain this field*. It is recommended IRDs decode this field.

F.4.3.9 extended_surround_flag

Semantics: The extended_surround_flag indicates the presence of DTS ES rear centre audio as defined in TS 102 114 [15]. Its values are given in table F.6.

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended IRDs decode this field.

Table F.6: extended_surround_flag values

Value	Description
00	No Extended Surround
01	Matrixed Extended Surround
10	Discrete Extended Surround
11	undefined

F.4.4 Use of the DVB-SI component_descriptor

Semantics: The semantics of the component_descriptor is defined in EN 300 468 [6]. The stream_content and component_type assigned values for DVB DTS audio stream are listed in annex F of EN 300 468 [6].

Encoding: *The values for the elements of the component_descriptor shall be set in accordance with annex F of EN 300 468 [6].*

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field to indicate the type of audio service present.*

F.5 PES Constraints

F.5.1 Encoding

In some applications, the audio decoder may be capable of simultaneously decoding two elementary streams containing different programme elements, and then combining the programme elements into a complete programme.

Most of the programme elements are found in the main audio service. Another programme element (such as a narration of the picture content intended for the visually impaired listener) may be found in the associated audio service.

In order to have the audio from the two elementary streams reproduced in exact sample synchronism, it is necessary for the original audio elementary stream encoders to have encoded the two audio programme elements frame synchronously; i.e., if audio stream 1 has sample 0 of frame *n* taken at time *t* 0, then audio stream 2 should also have frame *n* beginning with its sample 0 taken the identical time *t* 0. *If the encoding of multiple audio services is done frame and sample synchronous, and decoding is intended to be frame and sample synchronous, then the PES packets of these audio services shall contain identical values of PTS, which refer to the audio access units intended for synchronous decoding.*

Audio services intended to be combined together for reproduction shall be encoded at an identical sample rate.

F.5.2 Decoding

If audio access units from two audio services which are to be simultaneously decoded have identical values of PTS indicated in their corresponding PES headers, then the corresponding audio access units shall be presented to the audio decoder for simultaneous synchronous decoding. Synchronous decoding means that for corresponding audio frames (access units), corresponding audio samples are presented at the identical time.

If the PTS values do not match (indicating that the audio encoding was not frame synchronous) then the audio frames (access units) of the main audio service may be presented to the audio decoder for decoding and presentation at the time indicated by the PTS. An associated service, which is being simultaneously decoded, may have its audio frames (access units), which are in closest time alignment (as indicated by the PTS) to those of the main service being decoded, presented to the audio decoder for simultaneous decoding. In this case the associated service may be reproduced out of sync by as much as 1/2 of a frame time. (This is typically satisfactory; a visually impaired narration does not require highly precise timing.)

F.5.3 DTS PES Field constraints

The DTS Audio format PES packet is defined according to ISO/IEC 13818-1 [1] with the following exceptions.

F.5.3.1 stream_id

In Program Streams, the stream_id for DTS is "private_stream_1" = 1011 1101 = 0xBD.

F.5.3.2 data_alignment_indicator

This is a 1 bit flag. When set to a value of "1" it indicates that the PES packet header is immediately followed by the DTS audio syncword.

F.5.3.3 PTS_flags

This is a 2 bit field. *If the PTS_flags field equals "10", the PTS fields shall be present in the PES packet header. If the PTS_flags field equals "00" no PTS fields shall be present in the PES packet header.* The value "01" is forbidden and 11 is invalid for audio PES streams.

F.5.3.4 DSM_trick_mode_flag

A 1 bit flag, which when set to "1" it indicates the presence of an 8 bit trick mode field. This has no meaning for DTS audio and is hence 0.

F.5.3.5 PES_extension_flag

A 1 bit flag, which when set to "1" indicates that an extension field exists in this PES packet header. When set to a value of "0" it indicates that this field is not present. It is always set to zero for DTS audio packets.

F.5.3.6 stuffing_byte

fixed 8-bit value equal to "1111 1111" that can be inserted. It should not be sent to the decoder unless it is placed at the end of the DTS data prior to the next sync word. A maximum of 32 stuffing bytes may be inserted.

F.5.4 Byte-alignment

The DTS elementary stream shall be byte-aligned within the MPEG-2 data stream. This means that the initial 8 bits of a DTS frame shall reside in a single byte, which is carried by the MPEG-2 data stream.

Annex G (informative): Receiver-Mixed Audio Description and other supplementary Audio Services

G.1 Overview

Audio description (AD) delivers a description of the scene as an ancillary component associated with a TV service. It is intended to aid understanding and enjoyment particularly, but not exclusively, for viewers who have visual impairments.

Loud sound effects or music could make added description hard to discern so an important requirement is to adjust, on a passage-by-passage basis, the relative level of programme sound in the mix which the AD user hears. The programme maker is best able to determine the level under controlled conditions when authoring the AD - information to modulate the level of programme sound in the AD-capable receiver is thus transmitted within the AD stream.

Individual AD users will have different aural acuity, describers will have different styles of delivery (voice pitch and timbre), several voices may be used to describe one programme and there are, in practice, differences in audio signal level for different home receivers. An essential requirement is for the user to be able to adjust the volume of the description signal to suit his/her condition.

The ability to optionally mix one or more supplementary additional audio channels with the main programme sound can have other applications, including multi-language commentaries, use for interactivity, and educational purposes.

G.2 Coding

Description content is voice only and is conveyed as a mono signal coded in accordance with ISO/IEC 11172-3 [9].

The principles of processing in a basic AD decoder are shown diagrammatically below in figure G.1.

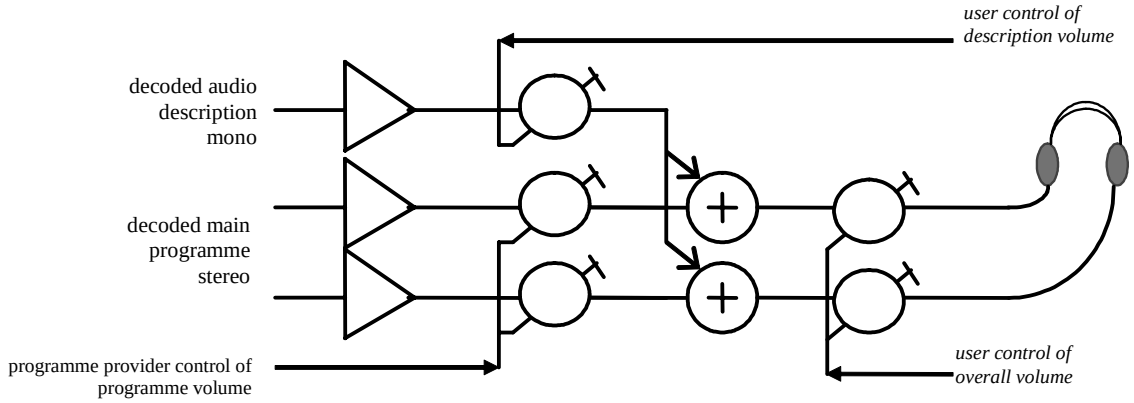


Figure G.1: Functionality of AD decoder processing

The level by which the programme sound should be attenuated during a description passage is signalled in PES_private_data within the PES encapsulation of the coded AD component (as specified in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].

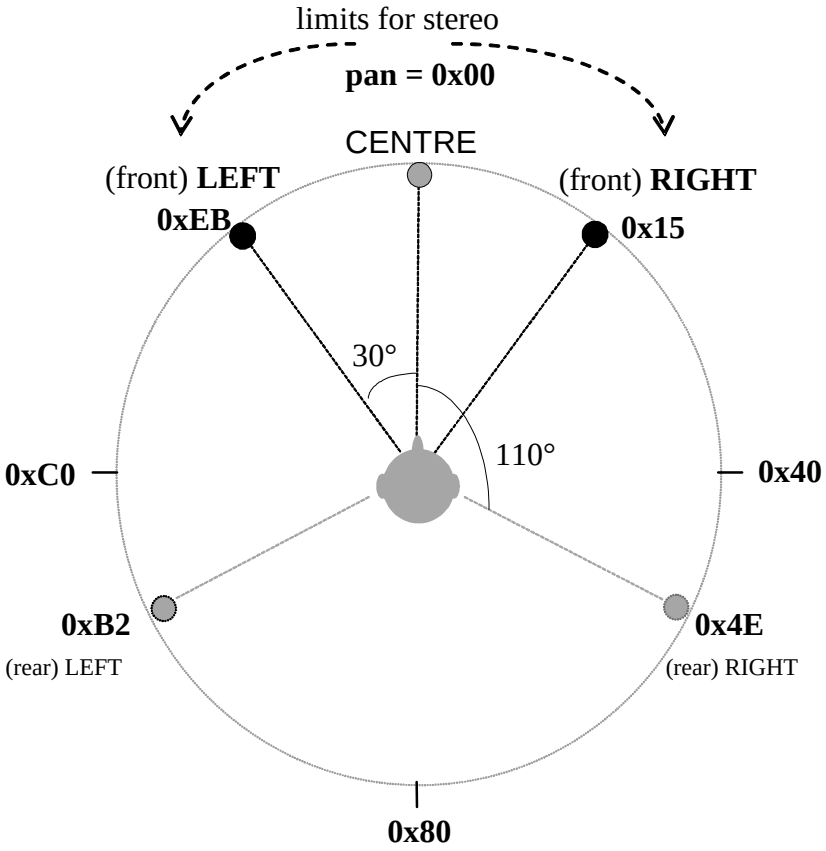
Coding : Support for the encoding of AD is optional.

Decoding : Support for the encoding of AD is optional.

The signalled fade value is an unsigned byte value, 0x00 representing 0 dB, each increment representing a nominal 0,3 dB, 0xFE representing approximately -77 dB whilst the fade value 0xFF represents completely mute programme sound.

A pan control value is also included within the transmitted data structure, enabling the decoded AD signal to be panned around the sound stage of the main programme sound and thus allowing the programme maker to place the "describer" at any preferred position within the sound field. As with fade, transmitted pan is a byte value, 0x00 representing centre front where each increment represents about 1.4° clockwise looking down on the listener (see figure G.2 below). For stereo the pan value will be restricted to ±30° of the centre front (i.e. to the range 0xEB..0xFF & 0x00..0x15) but the syntax of the signalling allows for any future use in which an AD component might be provided with a surround-sound main programme audio.

The values of fade & pan signalled in a PES packet apply to each access unit of AD sound contained within that same PES packet. This allows a fade (and a pan) to be relatively gradual or to be abrupt as the programme material allows.



NOTE: (Seen from above the listener; includes mapping onto multi-channel sound presentation).

Figure G.2: interpretation of audio description pan value

G.3 Syntax and Semantics

AD fade & pan control information is coded in PES_private_data within the PES encapsulation of the coded AD component in accordance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].

Table G.1: AD_descriptor

Syntax	value	No.of bits	Identifier
AD_descriptor {			
Reserved	1111	4	bslbf
AD_descriptor_length	1000	4	bslbf
AD_text_tag	0x4454474144	40	bslbf
revision_text_tag	0x31	8	bslbf
AD_fade_byte	0xXX	8	bslbf
AD_pan_byte	0xYY	8	bslbf
Reserved	0xFFFFFFFFFFFFFF	56	bslbf
}			

AD_descriptor_length: the number of significant bytes following the length field (i.e. 8).

AD_text_tag: a string of 5 bytes forming a simple and unambiguous means of distinguishing this from any other PES_private_data. A receiver which fails to recognize this tag should not interpret this audio stream as audio description.

revision_text_tag: the AD_text_tag is extended by a single ASCII character version designator (here "1" indicates revision 1). *Descriptors with the same **AD_text_tag** but a higher revision number shall be backwards compatible with this specification* - the syntax and semantics of the fade & pan fields will be identical but some of the reserved bytes may be used for additional signalling.

AD_fade_byte: takes values between 0x00 (representing no fade of the main programme sound) and 0xFF (representing a full fade). Over the range 0x00 to 0xFE one lsb represents a step in attenuation of the programme sound of approximately 0.3 dB giving a range of about 77 dB. The fade value of 0xFF represents no programme sound at all (i.e. mute). The rate of signalling and the expected behaviour of a decoder to changes in fade byte are described below.

AD_pan_byte: takes values between 0x00 representing a central forward presentation of the audio description and 0xFF, each increment representing a $^{360}/_{256}$ degree step clockwise looking down on the listener (i.e. just over 1.4 degrees, see figure G.2 above). The rate of signalling and the expected behaviour of a decoder are described below.

reserved: the remaining 7 bytes are set to 0xFF and reserved for future developments if and when required.

The maximum rate of signalling of fade & pan values is determined by the number of audio PES packets per second for that AD stream. For efficiency several access units (AUs) of audio are typically encapsulated within one PES packet and the fade & pan values in each AD_descriptor are deemed to apply to each AU encapsulated within, and which commences in, that PES packet. In typical efficient encapsulation fade & pan values are transmitted every 120 ms to 200 ms. This allows the control over the attack and decay of a fade where a particular gap in the narrative permits.

An AD decoder must maintain the relative timing between the decoded description signal and the decoded programme sound signal and between the appropriate fade & pan values and the decoded description signal.

During programmes for which there is no description there is little reason to transmit an AD stream of continual silence; in these cases the bit-rate accorded to AD may be reassigned for other purposes. Decoders should therefore be able to respond promptly to the restoration of the AD component at the start of a described programme.

The streams for programme sound and for AD are distinguished in the PSI by the use of the ISO_639_language descriptor. The audio_type field within the descriptor associated with programme sound is typically assigned the value 0x00 ("undefined") whilst the equivalent descriptor associated with AD has its audio_type field assigned the value 0x03 ("visual impaired commentary"). If a service has AD in several languages the PMT reference to each stream will have the appropriate ISO_639_language_code and the AD-capable decoder should discriminate between them on the basis of the preferred language chosen in the user settings.

G.4 Decoder behaviour

If there is a valid AD descriptor in the encoded description signal for the selected service, the AD decoder should present the appropriate mix of programme sound and description signal to the user, attenuating the programme sound by 0.3 dB per fade value increment. If the AD decoder cannot support such small steps then the implemented attenuation should match the intended attenuation as closely as possible. For example if only 1 dB steps are possible then fade values of 0x00 and 0x01 should map to 0 db, 0x02, 0x03 and 0x04 should map to1dB, 0x05, 0x06, 0x07 & 0x08 to -2 db etc.

When the fade value is 0x00 (or in the absence of an AD stream) the programme sound level should be unattenuated. Care should be taken to ensure that the default levels of programme sound and description are consistent when fed with streams coding standard level signals. It is also important that the mono description is matrixed to the stereo output so as to achieve a constant perceived description volume as the description is panned from stereo left through stereo centre to stereo right.

NOTE 1: E.g. using a model based on constant power as the description is panned across the stereo sound stage.

NOTE 2: The perceived loudness level of the main programme audio may well vary between different broadcast services. If the main programme audio is derived from a system using gain control metadata, for example AC-3, then the perceived loudness of the programme dialogue should be constant but it is likely to be different to that of a service for which the programme sound is delivered as MPEG-1 layer II. For any receiver which can decode main audio sources other than MPEG-1 layer II, the manufacturer may need to consider implementing different default gain levels for the audio description signal to provide a reasonable match of loudness to that of the programme dialogue. The ability of the user to adjust the relative level of description should nevertheless be retained.

In a stereo environment the AD decoder should interpret any pan values outside the ranges 0xEB..0xFF and 0x00..0x15 in the following manner. Pan values from 0x16 to 0x7F inclusive should be mapped to the value 0x15 (i.e. stereo hard right). Pan values from 0x80 to 0xEA should be mapped to the value 0xEB (i.e. stereo hard left).

When the user selects a new service or if the AD decoder detects an error in, or absence of, the AD descriptor in the encoded AD signal, the AD decoder should have a strategy which leads to muting the decoded description signal, restoring the programme sound to its default unfaded amplitude and setting the effective fade & pan values to 0x00. This restoration should not be abrupt - it is recommended that under such conditions the value of fade and of pan are ramped to the default values (0x00) over a period of at least 1 second. Equally, if the AD stream component is suddenly regained the implemented value of fade and of pan should be ramped to the signalled values from the default values (0x00) over a similar period.

G.5 Decoder user indicators

Description is typically confined to gaps in the programme narrative; these opportunities are therefore dependent on the programme. Some programmes are more suited to description than others; one may be effectively self-describing whilst another (e.g. news or a studio interview) might offer no opportunity for descriptive interpolation. Receiver implementations of AD should therefore allow the user to confirm that, in what may be extended gaps between description passages, description silence does not necessarily imply failure in delivery of the service or in the receiving equipment.

Many potential users of AD will be visually impaired. The user interface should not, therefore, rely solely on visual clues (lights or on-screen display logos) to indicate status (e.g. presence or absence of description). Audible indications are desirable and designers should consider how to distinguish different states using, for example, contrasting tones.

Annex H (informative): Guidelines for the Implementation of MPEG-4 High Efficiency AAC and High Efficiency AAC v2 Audio in DVB Compliant Transport Streams

H.1 Scope

The inclusion of MPEG-4 High Efficiency AAC v2 (HE AAC v2) audio streams in a DVB multiplex is optional, and IRDs may optionally decode these streams. This annex contains the guidelines to include one or more MPEG-4 HE AAC and HE AAC v2 elementary streams in a DVB Transport Stream in compliance with ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1]. The coding and decoding of an MPEG-4 HE AAC and HE AAC v2 elementary stream is based upon ISO/IEC 14496-3 [17].

It is recommended that implementations of DVB systems that include MPEG-4 HE AAC and HE AAC v2 audio streams should comply with this annex.

The MPEG-4 AAC and the MPEG-4 HE AAC profiles are subsets of the MPEG-4 HE AAC v2 profile. The MPEG-4 HE AAC adds the AOT SBR to the MPEG-4 AAC profile. The MPEG-4 HE AAC v2 Profile adds the AOT PS to the MPEG-4 HE AAC profile to improve the audio quality at low bit rates. Every HE AAC decoder can decode an HE AAC v2 bitstream, but will not be able to use the parametric stereo information and will therefore replay on a mono signal.

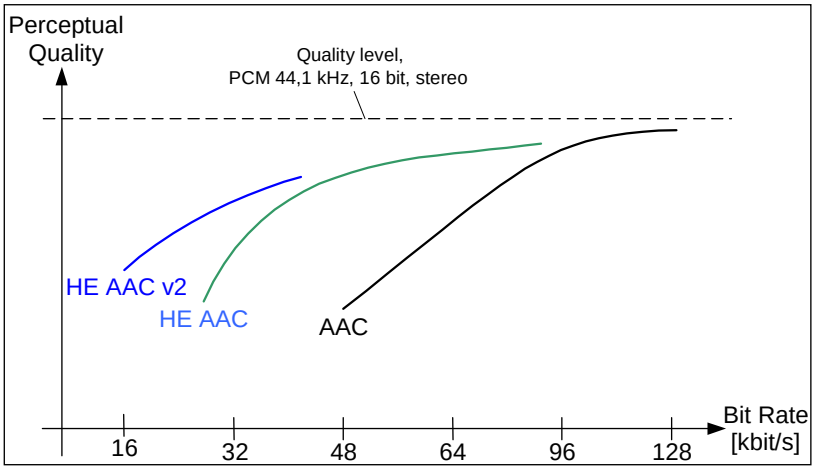


Figure H.1: Typical bit rate range of the HE AAC v2, HE AAC and AAC for stereo

Figure H.1 indicates the typical bit rate ranges for the use of HE AAC v2, HE AAC and AAC on the encoder side for stereo. The actual bit rates for the use of the different tools is dependent from the encoder implementation.

The IRD design should be made under the assumption that any legal structure as permitted by ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], including private data streams, may occur in the Transport Stream, even if presently reserved or unused. *To allow full compliance to the MPEG-2 standard and upward compatibility with future enhanced versions, a DVB IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the IRD.*

H.2 Introduction

An MPEG-4 HE AAC or HE AAC v2 elementary bitstream may be multiplexed into an MPEG-2 transport stream in much the same way an MPEG-1 audio stream would be included. The MPEG-4 HE AAC or HE AAC v2 elementary stream is packetized into PES packets with a structure similar to an MPEG audio PES.

It is necessary to unambiguously indicate that an MPEG stream is, in fact, an MPEG-4 HE AAC or an HE AAC v2 stream. A public DVB descriptor, the **AAC_descriptor** has been specified for this purpose. The syntactical elements that need to be specified in order to include MPEG-4 HE AAC and HE AAC v2 within an MPEG-2 transport stream are: the MPEG **stream_type, stream_id** and the DVB **AAC_descriptor**.

The ISO 639 language descriptor may be used to indicate the language of the content of the HE AAC or HE AAC v2 stream.

H.3 DVB Compliant Streams

*The MPEG-4 HE AAC or HE AAC v2 elementary stream data shall be first encapsulated in the LATM multiplex format according to ISO/IEC 14496-3 [17]. The **AudioMuxElement()** multiplex element format shall be used.*

*The LATM formatted MPEG-4 HE AAC or HE AAC v2 elementary stream data shall be encapsulated in the LOAS transmission format according to ISO/IEC 14496-3 [17]. The **AudioSyncStream()** version shall be used. **AudioSyncStream()** adds a sync word to the audio stream to allow for synchronization.*

*The LATM/LOAS formatted MPEG-4 HE AAC or HE AAC v2 elementary stream data shall be encapsulated in PES packets. The MPEG-4 HE AAC PES shall be carried with an MPEG **stream_id** = 110x xxxx and a stream type assignment of 0x11 as described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1].* No alignment is required. More than one audio unit is allowed per PES packet. *If a PTS is present in the PES header it shall refer to the first audio frame that follows the first syncword that commences in the PES packet.*

When an MPEG-4 HE AAC or HE AAC v2 stream is included in a DVB transport stream, the AAC_descriptor shall also be included. The AAC_descriptor is located in the PMT and the Selection Information Table of the DVB SI Tables defined in EN 300 468 [6].

H.4 Profiles and Levels

MPEG-4 HE AAC and HE AAC v2 is defined in the HE AAC and the HE AAC v2 profile. For Monaural, Parametric Stereo and Stereo, MPEG-4 HE AAC v2 bit-streams will comply with level 2. For Monaural and Stereo, MPEG-4 HE AAC bit-streams will comply with level 2. For multichannel, up to 5.1 channels, MPEG-4 HE AAC and HE AAC v2 bit-streams will comply with level 4.

Encoding: *The encoder shall use either the MPEG-4 AAC LC Profile, the MPEG-4 HE AAC Profile or the MPEG-4 HE AAC v2 Profile.* Use of the MPEG-4 HE AAC Profile is recommended.

Bit-streams including support for MPEG-4 HE AAC v2 monaural, parametric stereo and stereo shall comply with the HE AAC v2 Profile Level 2 restrictions.

Bit-streams including support for MPEG-4 HE AAC monaural and stereo shall comply with the HE AAC Profile Level 2 restrictions.

Bit-streams including support for MPEG-4 HE AAC or HE AAC v2 multichannel shall comply with the HE AAC or HE AAC v2 Profile Level 4 restrictions respectively.

Decoding: *The IRD shall be capable of decoding the MPEG-4 HE AAC or theMPEG-4 HE AAC v2 Profile.*

A MPEG-4 HE AAC v2 monaural, parametric stereo and stereo enabled decoder shall support MPEG-4 HE AAC v2 Level 2 bitstreams. This requirement does include support for lower levels, but not other profiles. Support for other profiles and for levels beyond Level 2 is optional.

A MPEG-4 HE AAC monaural and stereo enabled decoder shall support MPEG-4 HE AAC Level 2 bitstreams. This requirement does include support for lower levels, but not other profiles. Support for other profiles and for levels beyond Level 2 is optional.

MPEG-4 HE AAC or HE AAC v2 multi-channel enabled decoder shall support MPEG-4 HE AAC or HE AAC v2 Level 4 bitstreams respectively. This requirement does include support for lower levels, but not other profiles. Support for other profiles and for levels beyond Level 4 is optional.

If an IRD supports more than Level 2 then it shall also support Matrix-Mixdown. It shall further support the application of downmixing_levels_MPEG4 in ancillary data (annex D).

H.5 Dynamic Range Control

The MPEG-4 AAC Dynamic Range Control (DRC) tool is defined in ISO/IEC 14496-3 [17], clause 4.5.2.7. The default level for the program reference level as referred to in clause 4.5.2.7.3 shall be -31.75 dB, which corresponds to **prog_ref_level =127**. For more detailed information on the MPEG-4 AAC Dynamic Range Control tool see ISO/IEC 14496-3 [17].

Encoding: The encoder may use the MPEG-4 AAC Dynamic Range Control (DRC) tool.

Decoding: *Each IRD shall support the MPEG-4 AAC Dynamic Range Control (DRC) too. In case no DRC data is transmitted by the encoder, the decoder shall not apply the DRC tool.*

H.6 Detailed specification

H.6.1 MPEG Transport Stream Compliance

H.6.1.1 Stream_id

Semantics: MPEG-4 HE AAC and HE AAC v2 streams will use the **stream_id** 110x xxxx as shown in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], table 2-18. The mapping of values of PID to **stream_type** is indicated in the Transport Stream (TS) Programme Map Table (PMT).

Encoding: *The value of the **stream_id** field for an MPEG-4 HE AAC and HE AAC v2 elementary streams shall be 110x xxxx, where each x can be either 0, or 1.*

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with MPEG systems syntax.*

H.6.1.2 Stream_type

Semantics: The semantics of the **stream_type** field are described in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], table 2-29.

Encoding: *The value of **stream_type** for an MPEG-4 HE AAC and HE AAC v2 elementary streams shall be 0x11 (indicating ISO/IEC 14496-3 [17] Audio with the LATM transport syntax).*

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with MPEG systems syntax.*

H.6.1.3 LATM/LOAS formatting

Semantics: The semantics of the **AudioMuxElement()** and **AudioSyncStream()** formatting are described in ISO/IEC 14496-3 [17].

Encoding: *The MPEG-4 HE AAC and HE AAC v2 elementary streams shall be formatted with **AudioMuxElement()** LATM multiplex format, and **AudioSyncStream()** LOAS transmission format.*

The following limitations to the LATM multiplex shall apply;

- **numLayer** shall be "0", as no scalable profile is used;
- **numProgram** shall be "0", as there is only one audio program per LATM multiplex;
- **numSubFrames** shall be "0", as there is only one **PayloadMux()** (access unit) per LATM **AudioMuxElement();**
- **allStreamsSameTimeFraming** shall be "1", as all payloads belong to the same access unit.

Decoding: *These formats shall be read by the IRD, and the IRD shall interpret these formats in accordance with MPEG-4 audio syntax.*

H.6.2 Use of the DVB-SI component_descriptor and multilingual_component_descriptor

Semantics: The semantics of the **component_descriptor** and **multilingual_component_descriptor** are defined in EN 300 468 [6]. The **stream_content** and **component_type** assigned values for DVB MPEG-4 HE AAC and HE AAC v2 audio streams are listed in table 26 of EN 300 468 [6].

Encoding: *The values for the elements of the **component_descriptor** and **multilingual_component_descriptor** shall be set in accordance with EN 300 468 [6], clauses 6.2.8 and 6.2.21.*

Decoding: *These fields shall be read by the IRD, and the IRD shall interpret these fields to indicate the type of audio service present.*

H.6.3 AAC_descriptor

The syntax of the **AAC_descriptor** is described in table H.1.

The **AAC_descriptor** syntax provides information about individual MPEG-4 AAC, MPEG-4 HE AAC and MPEG-4 HE AAC v2 elementary streams to be identified in the PSI PMT sections. The descriptor is located in the PSI PMT, and used once in a program map section following the relevant **ES_info_length** field for any stream containing MPEG-4 AAC, MPEG-4 HE AAC or MPEG-4 HE AAC v2 audio.

Table H.1: AAC descriptor Syntax

Syntax	No.of Bits	Identifier
AAC_ descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
Profile_and_level	8	uimsbf
AAC_type_flag	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
reserved	1	bslbf
if (AAC_type_flag == 1)		
AAC_type	8	uimsbf
for(i=0;i<N;i++) {		
additional_info[N]	8*N	uimsbf
}		
}		
NOTE: Horizontal lines in the Table indicate allowable termination points for the descriptor.		

H.6.3.1 descriptor_tag

Semantics: The descriptor tag is an 8-bit field, which identifies each descriptor.

Encoding: *The value of the AAC **descriptor_tag** shall be set to 0x79 (see table 12 in EN 300 468 [6]).*

Decoding: *The IRD shall use this field to identify the descriptor.*

H.6.3.2 descriptor_length

Semantics: This 8-bit field specifies the total number of bytes of the data portion of the descriptor. The **AAC_descriptor** has a minimum length of four bytes but may be longer depending on the use of the **AAC_type_flag** and the **additional_info_loop**.

Encoding: *This field shall be set to the total number of bytes of the data portion of the descriptor following the byte defining the value of this field.*

Decoding: *This field shall be read by the IRD, and the IRD shall interpret this field in accordance with clause 2.6.1 of ITU-T Recommendation H.222.0 | ISO/IEC 13818-1:2000 [1].*

H.6.3.3 Profile_and_level

Semantics: This 8-bit field specifies the Profile and Level used in MPEG-4 AAC, MPEG-4 HE AAC or MPEG-4 HE AAC v2.

Encoding: *This field shall be set to the Profile and Level according to table 2-62 in ISO/IEC 13818-1:2000/FPDAM 5 [1].*

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended that IRDs decode this field.

H.6.3.4 AAC_type_flag

Semantics: This 1-bit field indicates the presence of the **AAC_type** field.

Encoding: *This bit shall be set to "1" if the optional AAC_type field is included in the descriptor.*

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* It is recommended that IRDs decode this field.

H.6.3.5

Void.

H.6.3.6

Void.

H.6.3.7 reserved flags

Semantics: These 1-bit fields are reserved for future use.

Encoding: *These bits shall all be set to "0".*

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

H.6.3.8 AAC_type

Semantics: This optional 8-bit field indicates the type of audio carried in the MPEG-4 AAC, MPEG-4 HE AAC or MPEG-4 HE AAC v2 elementary stream.

Encoding: *This field shall be set to the same value as the component_type field of the component descriptor (see table 26 in EN 300 468 [6]).*

Decoding: *IRDs shall be able to accept bit-streams, which contain this field.* IRDs may ignore the data within this field.

H.6.4 STD audio buffer size

It is recommended that for MPEG-4 HE_AAC v2 audio in a DVB system, the main audio buffer size *(BS_n)* has a value of 3 584 bytes for level 2 decoders and 8 976 bytes for level 4 decoders as defined in ITU-T Recommendation H.222.0 | ISO/IEC 13818-1 [1], clause 2.11.2.2.

Annex I (informative): Bibliography

ITU-R Recommendation BT.470-3: "System B,G,I".

ETSI TR 101 162: "Digital broadcasting systems for television, sound and data services; Allocation of Service Information (SI) codes for Digital Video Broadcasting (DVB) systems".

CEI/IEC 61883-1: "Consumer audio/video equipment - Digital interface - Part 1: General".

CEI/IEC 61883-2: "Consumer audio/video equipment - Digital interface - Part 2: SD-SVCR data transmission".

CEI/IEC 61883-3: "Consumer audio/video equipment - Digital interface - Part 3: HD-DVCR data transmission".

CEI/IEC 61883-4: "Consumer audio/video equipment - Digital interface - Part 4: MPEG2-TS data transmission".

History

Document history		
Edition 1	January 1996	Publication as ETR 154
Edition 2	October 1996	Publication as ETR 154
Edition 3	September 1997	Publication as ETR 154
V1.4.1	July 2000	Publication as TR 101 154
V1.5.1	May 2004	Publication
V1.6.1	January 2005	Publication
V1.7.1	June 2005	Publication