ETSI TR 102 154 V1.1.1 (2001-04)

Technical Report

Digital Video Broadcasting (DVB); Implementation guidelines for the use of MPEG-2 Systems, Video and Audio in Contribution and Primary Distribution Applications

Reference
DTR/JTC-DVB-89

Keywords
DVB, audio, broadcasting, data, digital, MPEG,
video, TV, contribution

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://www.etsi.org/ipr).

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 Report (TR) 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 present document is complementary to TR 101 154 [17]. TR 101 154 [17] provides MPEG Implementation Guidelines for the use of MPEG Systems, Video and Audio in Satellite, Cable and Terrestrial Broadcasting Applications. The present document is a companion document which follows the same format and provides MPEG Implementation Guidelines for the use of MPEG Systems, Video and Audio in contribution applications, such as Satellite News Gathering (SNG), and in primary distribution applications such as delivery of network programming to local broadcasting stations.

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 contribution/primary distribution coding and decoding using the MPEG-2 system layer, video coding and audio coding as defined in ISO/IEC 13818-1 [1], ISO/IEC 13818-2 [2], ISO/IEC 11172-3 [10] and ISO/IEC 13818-3 [3] respectively.

Contribution applications are those where the signal from the MPEG decoder is not supplied directly to the ultimate viewer, and where some form of processing (e.g. in a vision mixer) may be expected before the signal reaches the ultimate viewer. Satellite News Gathering (SNG) is a typical example of a contribution application. Primary Distribution applications are those where a completed programme is delivered to a site, such as a local broadcast station, that will then encode the signal for delivery to the final consumer. Some additional processing of the signal may occur. A higher quality of video and audio signal may be required for contribution and primary distribution applications than for the secondary distribution/emission applications. A secondary distribution/emission application is one where the signal from the MPEG decoder is expected to be either supplied directly to the ultimate viewer or else to the ultimate viewer via some intermediate stage that does not involve video or audio processing. Guidelines for secondary distribution/emission applications are provided by TR 101 154 [17].

The guidelines presented in the present document for the Contribution/Primary Distribution Integrated Receiver-Decoder (herewith designated "Contribution IRD") are intended to represent a minimum functionality that all Contribution IRDs of a particular class are required to either meet or exceed.

Contribution IRDs are classified in two dimensions as:

- "25 Hz" or "30 Hz", depending on whether the nominal video frame rates based on 25 Hz or 30 000/1001Hz (approximately 29,97 Hz) are supported. It is expected that 25 Hz Contribution IRDs will be used in those countries where the existing analogue TV transmissions use 25 Hz frame rate and 30 Hz Contribution IRDs will be used in countries where the analogue TV transmissions use 30 000/1001 Hz frame rate. There are also likely to be "dual-standard" Contribution IRDs which have the capabilities of both 25 Hz and 30 Hz Contribution IRDs.

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

To give a complete definition of a Contribution IRD, both dimensions need to be specified, e.g. 25 Hz SDTV Contribution IRD.

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 which the Contribution IRDs are required to decode correctly.

The specification of these baseline features in no way prohibits manufacturers of Contribution IRDs 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 Contribution IRD's up-sampling filter, which affect the quality of the displayed picture rather than whether the Contribution IRD is able to decode pictures at all. Such issues are left to the marketplace.

The guidelines presented for Contribution IRDs observe the following principles:

- a Contribution IRD should be capable of decoding all corresponding contribution bitstreams specified in the mandatory clauses of the present document. For example, a 25 Hz SDTV Contribution IRD should be capable of decoding a 25 Hz SDTV contribution bitstream. However, it would not necessarily be capable of decoding any optional features, such as those specified in the informative annexes of the present document;

- a Contribution IRD should also be capable of decoding all corresponding secondary distribution/emission bitstreams specified in the mandatory clauses of TR 101 154 [17]. For example, a 25 Hz SDTV Contribution IRD should be capable of decoding a 25 Hz SDTV secondary distribution/emission bitstream. However, it would not necessarily be capable of decoding any optional features, such as those specified in the informative annexes of TR 101 154 [17];

- wherever practical, Contribution 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 should be ignored by Contribution 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 [6];
- scrambling may be implemented in accordance with the Basic Interoperable Scrambling System (BISS) as specified in EBU Tech 3290 [27] (see annex F of the present document);
- conditional access uses the MPEG-2 Conditional Access CA_descriptor.

Video:

- MPEG-2 Main Profile at Main Level or 4:2:2 Profile at Main Level is used for SDTV;
- MPEG-2 Main Profile at High Level or 4:2:2 Profile at High Level is used for HDTV;
- the 25 Hz SDTV Contribution IRD supports 24 or 25 Hz frame rate;
- the 25 Hz HDTV Contribution IRD supports frame rates of 24 Hz, 25 Hz or 50 Hz;
- the 30 Hz SDTV Contribution IRD supports frame rates of 24 000/1001, 24, 30 000/1 001 and 30 Hz;
- the 30 Hz HDTV Contribution IRD supports frame rates of 24 000/1001, 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; Contribution IRDs support 4:3 and 16:9 and optionally 2,21:1 aspect ratio;
- HDTV pictures have 4:3, 16:9 or 2,21:1aspect ratio; Contribution IRDs support 4:3, 16:9 and optionally 2.21:1 aspect ratio (the 4:3 aspect ratio applies to SDTV resolution progressive scan pictures with a 50 or 60 Hz frame rate; these are considered HDTV because they require the High Level);
- contribution 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;
- contribution 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:

- sampling rates of 32 kHz, 44,1 kHz and 48 kHz are supported by all Contribution IRDs;
- contribution IRDs support uncompressed audio or SMPTE/AES data via the AES3 interface (see clause 7 of the present document);
- MPEG-1 or MPEG-2 stereo Layer I and Layer II decoding is supported by all Contribution IRDs;
- contribution IRDs support single channel, dual channel, joint stereo, stereo and the extraction of at least a stereo pair from MPEG-2 backwards compatible multi-channel audio;
- contribution IRDs may also optionally support full multi-channel decoding of MPEG-2 Layer II backwards compatible multi-channel audio;
- contribution IRDs may also optionally support Dolby AC-3 audio decoding (refer to annex C of the present document);

- contribution IRDs may also optionally support the decoding of MPEG audio streams which include ancillary data (see annex D of the present document);
- the audio content may be delivered as AES3 linear PCM audio at 48 kHz sampling rate (see clause 7) or may be encoded in one of the following modes: MPEG-1 Layer I, MPEG-1 Layer II, MPEG-2 Layer II backwards compatible audio. When the encoded audio is intended to be decoded by equipment that is known to be compliant with annex C, audio may be encoded with AC-3. By prior arrangement between the transmitting and receiving parties, audio may also be delivered using other types of audio encoding systems that interface to the equipment via the AES ports (see clause 7);
- the encoded bit-stream does not use emphasis;
- the use of Layer II encoding is recommended for MPEG-1 audio bit-streams.

1 Scope

The present document provides implementation guidelines for the use of MPEG-2 audio-visual coding in contribution and primary distribution applications. Both Standard Definition Television (SDTV) and High Definition Television (HDTV) are covered.

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

For the purposes of this Technical Report (TR) the following references apply:

[1] ISO/IEC 13818-1 (1996): "Information technology - Generic coding of moving pictures and associated audio information: Systems".

[2] ISO/IEC 13818-2 (1996): "Information technology - Generic coding of moving pictures and associated audio information: Video".

[3] ISO/IEC 13818-3 (2nd edition, 1998): "Information technology - Generic coding of moving pictures and associated audio information - Part 3: Audio".

[4] ISO/IEC 13818-7 (1997): "Information technology - Generic coding of moving pictures and associated audio information - Part 7: Advanced Audio Coding (AAC)".

[5] 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".

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

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

[8] ETSI TR 101 211: "Digital Video Broadcasting (DVB); Guidelines on implementation and usage of Service Information (SI)".

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

[10] 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".

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

[12] CEI/IEC 61883 (1998): "Consumer audio/video equipment - Digital interface - Part 1: General; Part 4: MPEG2-TS data transmission".

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

[14] ITU-R Recommendation BS.1196 (1995): "Digital Audio Compression (AC-3) Standard (ATSC Standard)" (annex 2).

[15] ETSI ETR 162: "Digital Video Broadcasting (DVB); Allocation of Service Information (SI) codes for DVB systems".

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

[17] ETSI TR 101 154: "Digital Video Broadcasting (DVB); Implementation guidelines for the use of MPEG-2 Systems, Video and Audio in satellite, cable and terrestrial broadcasting applications".

[18] SMPTE 302 Television: "Mapping of AES3 Data into MPEG-2 Transport Stream".

[19] SMPTE 337 M: "Television - Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface".

[20] ANSI S4.40-1992, Digital Audio Engineering: "Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data (AES3)".

[21] SMPTE 276M: "Television - Transmission of AES/EBU Digital Audio Signals Over Coaxial Cable".

[22] IEC 61937: "Interface for non-linear PCM encoded audio bitstreams applying IEC 60958".

[23] IEC 60958-1 (1999): "Digital audio interface - Part 1: General".

[24] IEC 60958-3 (1999): "Digital audio interface - Part 3: Consumer applications".

[25] ITU-R Recommendation BT.1359-1: "Relative timing of sound and vision for broadcasting".

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

[27] EBU Tech 3290 (March 2000): "Basic Interoperable Scrambling System (BISS)".

[28] ISO 639: "Code for the representation of names of languages".

[29] ITU-R Recommendation BT.470-6: "Conventional television systems".

3 Definitions and abbreviations

3.1 Definitions

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

25 Hz SDTV Contribution IRD: Contribution IRD which is capable of decoding and displaying pictures based on a nominal video frame rate of 24 Hz or 25 Hz from MPEG-2 Main Profile at Main Level or 4:2:2 Profile at Main Level bitstreams as specified in the present document.

25 Hz SDTV Contribution Bitstream: Bitstream which contains only Main Profile at Main Level or 4:2:2 Profile at Main Level video based on a nominal video frame rates of 24 Hz or 25 Hz as specified in the present document.

25 Hz HDTV Contribution IRD: Contribution IRD that is capable of decoding and displaying pictures based on a nominal video frame rate of 24 Hz, 25 Hz or 50 Hz from MPEG-2 Main Profile at High Level or 4:2:2 Profile at High Level bitstreams as specified in the present document. In addition, it provides the functionality of a 25 Hz SDTV contribution IRD.

25 Hz HDTV Contribution Bitstream: bitstream which contains only Main Profile at High Level, 4:2:2 Profile at High Level or simpler video based on a nominal video frame rates of 24 Hz, 25 Hz or 50 Hz as specified in the present document

30 Hz SDTV Contribution IRD: Contribution IRD which is capable of decoding and displaying pictures based on a nominal video frame rate of 24 000/1 001 (approximately 23,98) Hz, 24 Hz, 30 000/1 001 (approximately 29,97) Hz or 30 Hz from MPEG-2 Main Profile at Main Level or 4:2:2 Profile at Main Level bitstreams as specified in the present document.

30 Hz SDTV Contribution Bitstream: Bitstream which contains only Main Profile at Main Level or 4:2:2 Profile at Main Level video based on a nominal video frame rates of 24 000/1 001 Hz, 24 Hz, 30 000/1001 Hz or 30 Hz as specified in the present document.

30 Hz HDTV Contribution IRD: Contribution IRD that is capable of decoding and displaying pictures based on nominal video frame rates of 24 000/1 001 Hz, 24 Hz, 30 000/1001 Hz, 30 Hz, 60/1 001 Hz or 60 Hz from MPEG-2 Main Profile at High Level or 4:2:2 Profile at High Level bitstreams as specified in the present document. In addition, it provides the functionality of a 30 Hz SDTV contribution IRD.

30 Hz HDTV Contribution Bitstream: Bitstream which contains only Main Profile at High Level, 4:2:2 Profile at High Level or simpler video based on a nominal video frame rates of 24 000/1 001 Hz, 24 Hz, 30 000/1 001 Hz, 30 Hz, 60/1 001 Hz or 60 Hz as specified in the present document.

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.

3.2 Abbreviations

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

AAC	Advanced Audio Coding according to ISO/IEC 13818-7 [4]
AC-3	Dolby AC-3 audio coding system according to ITU-R Recommendation BS.1196-E [14] (1995)
AFD	Active Format Description
BISS	Basic Interoperable Scrambling System
CA	Conditional Access
CAT	Conditional Access Table
CW	Control Word
DVB	Digital Video Broadcasting
ECM	Entitlement Control Message
EMM	Entitlement Management Message
ES	Elementary Stream
ESCR	Elementary Stream Clock Reference
I-Frame	Intra-coded Frame
IRD	Integrated Receiver-Decoder
HDTV	High Definition Television
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
PSI	Program Specific Information
PSW	Pan and Scan Window
SI	Service Information
SDTV	Standard Definition Television
STD	System Target Decoder
TS	Transport Stream
TSDT	Transport Stream Description Table
T-STD	Transport stream System Target Decoder
VCR	Video Cassette Recorder

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 Contribution IRD. The source bitstream may be carried as a baseband signal via a digital interface or transmitted via a satellite, cable or terrestrial channel.

4.1 Contribution bitstreams and Contribution IRDs

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

The application of scrambling to the DVB bitstream is optional. If scrambling is applied, it shall be in accordance with ETR 289 [6]. It may be implemented in accordance with the Basic Interoperable Scrambling System (BISS) as specified in EBU Tech 3290 [27], as described in annex F.

To allow full compliance to ISO/IEC 13818-1 [1] and upward compatibility with future enhanced versions, a DVB contribution IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the contribution 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, contribution IRD design should be made under the assumption that any legal structure as permitted by 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 below in this clause are based on ISO/IEC 13818-1 [1]. The numbers in brackets after the headings are the relevant clause and section headings of ISO/IEC 13818-1 [1].

4.1.1 Introduction (ISO/IEC 13818-1, clause 0)

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 contribution IRDs shall be able to demultiplex the MPEG-2 transport stream. Demultiplexing of program streams (as described in clauses 0.2 and 0.3 of [1]) is optional.

4.1.2 Packetized Elementary Stream (PES) (ISO/IEC 13818-1, clause 0.4)

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 (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 ISO/IEC 13818-1 [1]. It is recommended that the tolerance is within 5 parts per million.

Decoding: The contribution IRD shall operate over the full tolerance range of the system clock frequency specified in clause 2.4.2.1 of ISO/IEC 13818-1 [1].

4.1.4 Transport packet layer (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 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 contribution 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 contribution 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 [6].

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: The transport_scrambling_control bits shall be read by the contribution IRD, and the contribution 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 [7].

4.1.5 Adaptation field (ISO/IEC 13818-1, clause 2.4.3.4)

4.1.5.1 Random_access_indicator

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).

4.1.5.3 elementary_stream_priority_indicator

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

4.1.5.4 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 2.7.3 of ISO/IEC 13818-1 [1]. It is recommended that this interval should be no greater than 40 ms.

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

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

4.1.6 Packetized Elementary Stream (PES) Packet (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 ISO/IEC 13818-1 [1], table 2-18 and table 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 [6].

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 contribution IRD, and the contribution IRD shall respond in accordance with table 2.

4.1.6.3 PES_priority

Decoding: The **PES_priority** bit may be ignored by the contribution 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 contribution IRD need not interpret these bits. *The setting of these bits shall not be altered in any digital output from the contribution 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 contribution IRD may skip over any data which is flagged as being in a trick mode, if it does not support decoding of trick modes.

4.1.6.6 additional_copy_info

Encoding: This field may used as appropriate.

Decoding: The contribution IRD need not interpret this field. *The coding of the field shall not be altered in any digital output from the contribution 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 contribution IRD shall be able to accept bit-streams which contain these fields. The contribution 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 contribution IRD shall be able to accept bit-streams which contain this field. The contribution IRD may ignore the data within the field.

4.1.7 Program Specific Information (PSI) (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 [7]. This document also defines additional tables for service information which use Program Specific Information (PSI) private_section structure defined in ISO/IEC 13818-1 [1].

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

4.1.8 Program and elementary stream descriptors (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 contribution 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 contribution 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 contribution 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 x 576 pixels for a 25 Hz bitstream or is other than 720 x 480 pixels for a 30 Hz bitstream, otherwise its use is optional.

Decoding: If this descriptor is absent, a default grid of 720 x 576 pixels shall be assumed by a 25 Hz contribution IRD, a default grid of 720 x 480 pixels shall be assumed by a 30 Hz contribution IRD. The display of correctly windowed video on background grids other than 720 x 576 pixels is optional for a 25 Hz SDTV contribution IRD, the display of correctly windowed video on background grids other than 720 x 480 pixels is optional for a 30 Hz SDTV contribution IRD. The HDTV contribution 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 contribution 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 [6].

Decoding: The contribution IRD shall interpret this descriptor as defined in ETR 289 [6].

4.1.8.8 ISO_639_Language_descriptor

Encoding: The use of the ISO_639_Language_descriptor is recommended for all audio, video and data streams, especially if multiple audio (or video) streams with different languages are present within a program.

Decoding: The contribution IRD may use the data from this descriptor to assist the selection of the appropriate audio (or video) stream of a program, if multiple audio (or video) streams are 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 contribution 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 contribution 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 contribution 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 contribution 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 contribution IRD need not make use of this descriptor.

4.1.8.14 STD_descriptor

Encoding: The **STD_descriptor** shall be used as specified in ISO/IEC 13818-1 [1].

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

4.1.8.15 IBP_descriptor

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

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

4.1.8.16 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 contribution IRD need not make use of this descriptor, but the information may be of assistance to digital VCRs.

Additional descriptors to those defined in ISO/IEC 13818-1 [1] are defined in EN 300 468 [7], and guidelines for their use are provided in TR 101 211 [8].

4.1.9 Compatibility with ISO/IEC 11172-1 (ISO/IEC 13818-1, clause 2.8)

Decoding: Compatibility with ISO/IEC 11172-1 [9] (MPEG-1 Systems) is optional.

5 Video

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

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

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

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

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

The video encoding shall conform to 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 contribution IRD design should be made under the assumption that any legal structure as permitted by ISO/IEC 13818-2 [2] may occur in the broadcast 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 contribution IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the contribution IRD.

This clause is based on ISO/IEC 13818-2 [2].

5.1 25 Hz SDTV contribution IRDs and Bitstreams

5.1.1 Profile and level

Encoding: Encoded bit-streams shall comply with the 4:2:2 Profile at Main Level restrictions, as described ISO/IEC 13818-2 [2], clause 8.2., amendment 2. The **profile_and_level_indication** is "10000101" for 4:2:2 Profile at Main level or, if appropriate, "01001000" for Main Profile at Main Level or "0nnnnnnn", where "0nnnnnnn" > "01001000", indicating a "simpler" profile or level than Main Profile, Main Level.

Decoding: The 25Hz SDTV contribution IRD shall support the decoding of 4:2:2 Profile at Main Level bitstreams. This requirement includes support for "simpler" profiles and levels, including Main Profile at Main Level, as defined in table 8-15 of ISO/IEC 13818-2 [2]. Support for profiles and levels beyond 4:2:2 Profile, Main Level is optional. If the contribution 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 24 Hz or 25 Hz, i.e. **frame_rate_code** is "0010" or "0011" respectively.

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

Decoding: All 25Hz SDTV contribution IRDs shall support the decoding and display of video material with a frame rate of 24 Hz and 25 Hz (i.e. **frame_rate_code** of "0010" and "0011" respectively). Support of other frame rates is optional.

25 Hz SDTV contribution 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 ISO/IEC 13818-1 [1], clause 2.1.48).

5.1.3 Aspect ratio

Encoding: The source aspect ratio in 25 Hz 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 SDTV contribution 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 SDTV contribution 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 contribution IRD has a digital interface, this should be capable of outputting bit-streams with aspect ratios which are not directly supported by the contribution IRD to allow their decoding and display via an external unit.

All 25 Hz SDTV contribution IRDs shall support the use of pan vectors and upsampling 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. contribution IRDs implementing the 2,21:1 aspect ratio should support the use of pan vectors and upsampling 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 contribution IRD).

Decoding: The 25 Hz SDTV contribution IRD shall be capable of decoding pictures with luminance resolutions as shown in table 4 and applying upsampling to allow the decoded pictures to be displayed at full-screen size. In addition, contribution 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 contribution IRD

Coded Picture		**Displayed Picture Horizontal upsampling**	
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)	× ¾ (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 upsampling × 2)	× 3/2 (see note 1) × 2 × 5/2 (see note 4) (and vertical upsampling × 2)

NOTE 1: Upsampling of 4:3 pictures for display on a 16:9 monitor is optional in the contribution IRD, as 16:9 monitors can be switched to operate in 4:3 mode.
NOTE 2: The upsampling with this value is applied to the pixels of the 16:9 picture to be displayed on a 4:3 monitor.
NOTE 3: The upsampling with this value is applied to the pixels of the 2,21:1 picture to be displayed on a 4:3 monitor. Upsampling from 2,21:1 pictures for display on a 4:3 monitor is optional in the contribution IRD.
NOTE 4: The upsampling with this value is applied to the pixels of the 2,21:1 picture to be displayed on a 16:9 monitor. Upsampling from 2,21:1 pictures for display on a 16:9 monitor is optional in the contribution 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 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_characterstics** 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.470-6 [29] System B,G,I.

5.1.6 Chrominance

Encoding: If chrominance downsampling is performed, the operation used to downsample 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 downsampled independently. A value of one indicates that the two fields have been combined into a single frame before downsampling. It is desirable that the fields are downsampled independently (i.e. **chroma_420_type** = 0) to allow the contribution IRD to use less memory for picture reconstruction.

Decoding: It is desirable that the operation used to upsample 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 HDTV contribution IRDs and Bitstreams

5.2.1 Profile and level

Encoding: Encoded 25Hz HDTV bit-streams shall comply with the 4:2:2 Profile at High Level restrictions, as described ISO/IEC 13818-2 [2], clause 8.2, amendment 2. The **profile_and_level_indication** is "xxxxxxxx" for 4:2:2 Profile at High Level or, if appropriate, "01000100" for Main Profile at High Level or "0nnnnnnn", where "0nnnnnnn" > "01000100", indicating a "simpler" profile or level than Main Profile at High Level.

Decoding: The 25Hz HDTV contribution IRD shall support the decoding of 4:2:2 Profile at High Level bitstreams. This requirement includes support for "simpler" profiles and levels, including Main Profile at High Level, and 4:2:2 Profile at Main Level, as defined in ISO/IEC 13818-2 [2] table 8-15. Support for profiles and levels beyond 4:2:2 Profile at High Level is optional. If the contribution 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 24 Hz, 25 Hz or 50 Hz, i.e. **frame_rate_code** is "0010", "0011" or "0110" respectively.

The source video format for 24 Hz and 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 ISO/IEC 13818-1 [1], clause 2.1.48).

Decoding: All 25 Hz HDTV contribution IRDs shall support the decoding and display of video material with a frame rate of 24 Hz or 25 Hz progressive, 25 Hz interlaced or 50 Hz progressive (i.e. **frame_rate_code** of "0010", "0011", "0011" or "0110" respectively) within the constraints of 4:2:2 Profile at High Level. Support of other frame and field rates is optional.

25 Hz HDTV contribution 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 ISO/IEC 13818-1 [1], clause 2.1.48).

5.2.3 Aspect ratio

Encoding: The source aspect ratio in 25 Hz HDTV bit-streams shall be 4:3, 16:9 or 2,21:1. Note that decoding of 2,21:1 aspect ratio is optional for the 25 Hz HDTV contribution IRD. Note that the 4:3 aspect ratio only applies to SDTV resolution progressively scanned pictures transmitted at a 50 Hz frame rate.

The **aspect_ratio_information** in the sequence header shall have the value "0010", "0011" or "0100".

Decoding: The 25 Hz HDTV contribution IRD shall be able to decode bit-streams with **aspect_ratio_information** of value "0010", corresponding to a 4:3 aspect ration, or "0011", corresponding to 16:9 aspect ratio. The support of the aspect ratio 2,21:1 is optional. If the contribution IRD has a digital interface, this should be capable of outputting bit-streams with aspect ratios which are not directly supported by the contribution 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 4:2:2 Profile at High Level, i.e. it shall not have more than:

- 1 152 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 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 24 Hz or 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 4:2:2 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 4:2:2 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 4:2:2 Profile at High Level.

Decoding: The 25 Hz HDTV contribution IRD shall be capable of decoding and displaying pictures with luminance resolutions within the constraints set by 4:2:2 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 colorimetry is used in the 25 Hz HDTV bitstream, which is signalled by setting **colour_primaries** to the value 1, **transfer_characterstics** to the value 1 and **matrix_coefficients** to the value 1.

Decoding: The 25 Hz HDTV contribution 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 colorimetry.

5.2.6 Chrominance

Encoding: If chrominance downsampling is performed, the operation used to downsample 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 downsampled independently. A value of one indicates that the two fields have been combined into a single frame before downsampling. It is desirable that the fields are downsampled independently (i.e. **chroma_420_type** = 0) to allow the contribution IRD to use less memory for picture reconstruction.

Decoding: It is desirable that the operation used to upsample 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 HDTV contribution IRD shall be capable of decoding any bitstream that a 25 Hz SDTV contribution IRD is required to decode, as described in 5.1.

5.3 30 Hz SDTV Contribution IRDs and Bitstreams

5.3.1 Profile and level

Encoding: Encoded bit-streams shall comply with the 4:2:2 Profile at Main Level restrictions, as described ISO/IEC 13818-2 [2], clause 8.2, amendment 2. The **profile_and_level_indication** is "10000101" for 4:2:2 Profile at Main Level or, if appropriate, "01001000" for Main Profile at Main Level or "0nnnnnnn", where "0nnnnnnn" > "01001000", indicating a "simpler" profile or level than Main Profile at Main Level.

Decoding: The contribution IRD shall support the decoding of 4:2:2 Profile at Main Level bitstreams. This requirement includes support for "simpler" profiles and levels, including Main Profile at Main Level, as defined in table 8-15 of ISO/IEC 13818-2 [2]. Support for profiles and levels beyond 4:2:2 Profile at Main Level is optional. If the contribution 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 24000/1001 Hz, 24 Hz, 30000/1001 Hz 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 ISO/IEC 13818-1 [1], clause 2.1.48).

Decoding: All 30 Hz SDTV contribution IRDs shall support the decoding and display of 4:2:2 Profile @ Main Level video with a frame rate of 24000/1001 Hz, 24 Hz, 30000/1001 Hz or 30 Hz. Support of other frame rates is optional.

Contribution 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 ISO/IEC 13818-1 [1], clause 2.1.48)

5.3.3 Aspect ratio

Encoding: The source aspect ratio in 30 Hz 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 contribution 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 SDTV contribution 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 contribution IRD has a digital interface, this should be capable of outputting bit-streams with aspect ratios which are not directly supported by the contribution IRD to allow their decoding and display via an external unit.

All 30 Hz SDTV contribution IRDs shall support the use of pan vectors and upsampling 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. contribution IRDs implementing the 2,21:1 aspect ratio should support the use of pan vectors and upsampling 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 contribution IRD).

Decoding: The 30 Hz SDTV contribution IRD shall be capable of decoding pictures with luminance resolutions as shown in table 6 below and applying upsampling to allow the decoded pictures to be displayed at full-screen size. In addition, contribution 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 contribution IRD

Coded Picture		**Displayed Picture Horizontal upsampling**	
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 upsampling × 2)	× 3/2 (see note 1) × 2 × 5/2 (see note 4) (and vertical upsampling × 2)

NOTE 1: Upsampling of 4:3 pictures for display on a 16:9 monitor is optional in the contribution IRD, as 16:9 monitors can be switched to operate in 4:3 mode.
NOTE 2: The upsampling with this value is applied to the pixels of the 16:9 picture to be displayed on a 4:3 monitor.
NOTE 3: The upsampling with this value is applied to the pixels of the 2,21:1 picture to be displayed on a 4:3 monitor. Upsampling from 2,21:1 pictures for display on a 4:3 monitor is optional in the contribution IRD.
NOTE 4: The upsampling with this value is applied to the pixels of the 2,21:1 picture to be displayed on a 16:9 monitor. Upsampling from 2,21:1 pictures for display on a 16:9 monitor is optional in the contribution 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: If chrominance downsampling is performed, the operation used to downsample 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 downsampled independently. A value of one indicates that the two fields have been combined into a single frame before downsampling. It is desirable that the fields are downsampled independently (i.e. **chroma_420_type** = 0) to allow the contribution IRD to use less memory for picture reconstruction.

Decoding: It is desirable that the operation used to upsample 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 HDTV contribution IRDs and Bitstreams

5.4.1 Profile and level

Encoding: Encoded 30 Hz HDTV bit-streams shall comply with the 4:2:2 Profile High Level restrictions, as described ISO/IEC 13818-2 [2], clause 8.2, amendment 2.

The **profile_and_level_indication** is "xxxxxxxx" for 4:2:2 Profile at High Level or, if appropriate, "01000100" for Main Profile at High Level or "0nnnnnnn", where "0nnnnnnn" > "01000100", indicating a "simpler" profile or level than Main Profile at High Level.

Decoding: The 30 Hz HDTV contribution IRD shall support the decoding of 4:2:2 Profile at High Level bitstreams. This requirement includes support for "simpler" profiles and levels, including Main Profile at High Level and 4:2:2 Profile at Main Level, as defined in table 8-15 of ISO/IEC 13818-2 [2]. Support for profiles and levels beyond 4:2:2 Profile at High Level is optional. If the contribution 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 24 000/1001 Hz, 24 Hz, 30 000/1001 Hz, 30 Hz, 60 000/1 001Hz or 60 Hz, i.e. **frame_rate_code** is "0001", "0010", "0100", "0101", "0111" or "1000".

The source video format for 24 000/1001 Hz, 24 Hz, 60 000/1 001 Hz 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 ISO/IEC 13818-1 [1], clause 2.1.48).

Decoding: All 30 Hz HDTV contribution IRDs shall support the decoding of video material with a frame rate of 24 000/1 001 Hz, 24 Hz, 30 000/1 001 Hz, 30 Hz, 60 000/1 001 Hz or 60 Hz (i.e. **frame_rate_code** of "0001", "0010", "0100", "0101", "0111" or "1000") within the constraints of 4:2:2 Profile at High Level. Support of other frame rates is optional.

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

30 Hz HDTV contribution 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 ISO/IEC 13818-1 [1], clause 2.1.48).

5.4.3 Aspect ratio

Encoding: The source aspect ratio in 30 Hz HDTV bit-streams shall be 4:3, 16:9 or 2,21:1. Note that decoding of 2,21:1 aspect ratio is optional for the 30Hz HDTV contribution IRD. Note that the 4:3 aspect ratio only applies for SDTV resolution progressively scanned pictures transmitted at a 60/1.001 or 60 Hz frame rate.

The **aspect_ratio_information** field in the sequence header shall have the value "0010", "0011" or "0100".

Decoding: The 30 Hz HDTV contribution IRD shall be able to decode bit-streams with **aspect_ratio_information** of value "0010" corresponding to 4:3, or "0011", corresponding to 16:9 aspect ratio. If the contribution IRD has a digital interface, this should be capable of outputting bit-streams with aspect ratios which are not directly supported by the contribution 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 4:2:2 Profile at High Level, i.e. it shall not have more than:

- 1152 lines per frame;
- 1920 luminance samples per line;
- 62 668 800 luminance samples per second.

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

- 1080 lines per frame; and
- 1920 luminance samples per line,

with an associated frame rate of 30000/1001 (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 4:2:2 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 4:2:2 Profile at High Level excludes the use of the maximum allowed picture resolution at 60 Hz and 60 000/1 001 frame rates.

NOTE 2: 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 4:2:2 Profile at High Level.

Decoding: The 30 Hz HDTV contribution IRD shall be capable of decoding and displaying pictures with luminance resolutions within the constraints set by 4:2:2 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 BT.709 colorimetry is used in the 30 Hz HDTV bitstream, which is signalled by setting **colour_primaries** to the value 1, **transfer_characterstics** to the value 1 and **matrix_coefficients** to the value 1.

Decoding: The 30 Hz HDTV contribution 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 BT.709 colorimetry.

5.4.6 Chrominance

Encoding: If chrominance downsampling is performed, the operation used to downsample 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 downsampled independently. A value of one indicates that the two fields have been combined into a single frame before downsampling. It is desirable that the fields are downsampled independently (i.e. **chroma_420_type** = 0) to allow the contribution IRD to use less memory for picture reconstruction.

Decoding: It is desirable that the operation used to upsample 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 HDTV contribution IRD shall be capable of decoding any bitstream that a 30 Hz SDTV contribution IRD is required to decode, as described in 5.3.

6 Audio

This clause describes the guidelines for encoding MPEG-1 or MPEG-2 backward compatible audio in DVB broadcast bit-streams, and for decoding this bit-stream in the contribution IRD. The guidelines to include AES3 PCM audio are given in clause 7 and additional optional audio coding systems and ancillary data are described annexes C and D.

The recommended level for reference tones for transmission is 18 dB below clipping level, in accordance with EBU Recommendation R.68 [13].

The audio encoding shall conform to either ISO/IEC 11172-3 [10] or ISO/IEC 13818-3 [3], except in systems where contribution IRDs are required to comply with annex C. Some of the parameters and fields in ISO/IEC 11172-3 [10] and ISO/IEC 13818-3 [3] are not used in the DVB System and these restrictions are described below.

The contribution IRD design should be made under the assumption that any legal structure as permitted by ISO/IEC 11172-3 [10] 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 [10] and ISO/IEC 13818-3 [3] and upward compatibility with future enhanced versions, a DVB contribution IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the contribution IRD. For example, an contribution 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 [10] (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 [10] single channel;
- ISO/IEC 11172-3 [10] dual channel;
- ISO/IEC 11172-3 [10] joint stereo;
- ISO/IEC 11172-3 [10] stereo;
- ISO/IEC 13818-3 [3] multi-channel audio, backwards compatible to ISO/IEC 11172-3 [10] (dematrix procedure = 0, 1, 2 or 3).

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

- ISO/IEC 11172-3 [10] single channel;
- ISO/IEC 11172-3 [10] dual channel;
- ISO/IEC 11172-3 [10] joint stereo;
- ISO/IEC 11172-3 [10] stereo.

The contribution IRD shall be capable of decoding at least the ISO/IEC 11172-3 [10] 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.

6.2 Layer

Encoding: An ISO/IEC 11172-3 [10] 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: Contribution IRDs shall be capable of decoding MPEG-1 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, 64, 96, 128, 160, 192, 224, 256, 288, 320, 352, 384, 416 or 448 kbits/s.

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

For ISO/IEC 13818-3 [3] encoded bit-streams with total bitrates 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 kbit/s to 682 kbit/s.

Decoding: contribution 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 contribution 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 contribution IRD shall be capable of decoding audio with no emphasis. Support for 50/15 microseconds de-emphasis and ITU-T Recommendation J.17 [11] 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 contribution 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 at least the ISO/IEC 11172-3 [10] compatible basic stereo information from ISO/IEC 13818-3 [3] multi-channel audio 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_ch** equals "0").

Decoding: The IRD shall be capable of decoding at least the ISO/IEC 11172-3 [10] compatible basic stereo information from ISO/IEC 13818-3 [3] multi-channel audio 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 contribution IRD may interpret the ancillary data field in an ISO/IEC 131818-3 [3] stereo or multichannel bitstream as described in annex D and it is recommended that the contribution IRD make use of this data.

7 AES3 PCM Audio and SMPTE/AES Data

7.1 Scope

This clause contains the guidelines to include one or more AES3 PCM audio (or data) streams in a DVB Transport Stream by means of the SMPTE 302M [18] standard.

The inclusion of AES3 audio/data streams in a DVB multiplex is optional. *Contribution IRDs shall support uncompressed audio or SMPTE/AES data via the AES3 interface.*

7.2 Introduction

The AES3 standard [20] is widely used to convey one or two channels of linear PCM audio over a balanced twisted pair cable. SMPTE 276M [21] specifies the use of an unbalanced 75 ohm interconnect for the AES3 signal.
IEC 60958-1 [23] and 60958-3 [24] together specify a similar digital audio interface for consumer applications. SMPTE standard 337M [22] specifies the delivery of non-PCM (i.e. compressed) audio and/or data over the AES3 interface.
IEC 61937 [22] specifies the delivery of non-PCM (i.e. compressed) audio over the IEC 60958-3 [24] consumer interface. SMPTE standard 302M [18] specifies the transport of 1, 2, 3, or 4 AES3 signals through an MPEG-2 transport stream. (SMPTE 302M [18] will transport IEC 60958-3 [24] signals in an identical manner.)

This clause specifies the use of SMPTE 302M [18] to carry AES3 streams (or IEC 60958-3 [24] streams) through the DVB multiplex. A Contribution IRD has the capability to receive DVB Transport streams that can carry linear PCM (uncompressed) audio, or any type of data that is formatted onto an AES3 or IEC 60958-3 [24] signal. Data types that can be carried include metadata, captioning data, as well as several types of compressed audio. If data (which could be compressed audio) is carried, additional equipment may be needed to decode this data into some useful representation.

7.3 Specification

The Contribution IRD design should be made under the assumption that any legal structure as permitted by 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 Contribution IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the Contribution IRD.*

AES3 signals may be multiplexed into the DVB contribution transport stream by means of SMPTE 302M [18]. SMPTE 302M [18] allows one, two, three or four AES3 signals to be multiplexed into one elementary stream (ES) and PES packet, and transported through an MPEG-2 multiplex. It is possible to deliver multiple instances of a 302M multiplex through the MPEG-2 multiplex. It is recommended that encoding equipment support carriage of at least two AES3 signals. This may be done either by a single instance of 302M containing two AES3 signals, or by two instances of 302M each containing a single AES3 signal. *The Contribution IRD shall support both a single instance of 302M containing two AES3 signals and also two instances of 302M each containing a single AES3 signal.* Equipment may optionally support multiplexing up to four AES3 signals into a single 302M stream, and may support a large number of separate 302M streams.

7.3.1 Encoding and Decoding

Encoding: It is recommended that contribution encoders should allow at least one or two AES3 signals containing linear PCM audio to be included in either one or two 302M elementary streams that are carried as part of a program by the Contribution transport stream. Contribution encoders may optionally allow three or four AES3 signals to be included in a single 302M ES. Contribution encoders may optionally allow multiple 302M elementary streams, each containing between one and four AES3 streams, to be included in a program in the MPEG-2 multiplex.

Decoding: *Contribution IRD's shall accept transport streams with programs that include one or more 302M elementary streams that each contain one to four AES3 streams. The Contribution IRD shall allow the selection and decoding of any two AES3 streams included in the set of 302M streams that are part of a single program within the transport stream. In the absence of user selection, the Contribution IRD shall default to decoding the first two AES3 streams contained within the first 302M stream that is listed in the PMT. If the first 302M stream listed in the PMT contains a single AES3 stream, then that AES3 shall be selected by default, along with the first AES3 stream contained in the second 302M stream (if present) listed in the PMT.* Contribution IRD's may optionally allow the reproduction of a larger number of AES3 streams carried by single or multiple 302M streams.

7.3.2 MPEG-2 T-STD

Semantics: The MPEG-2 transport stream system target decoder (T-STD) is specified in ISO/IEC 13818-1 [1].

Encoding: *Contribution encoders shall multiplex the transport packets so as not to underflow or overflow the 302M T-STD buffer specified for decoding.*

Decoding: *Contribution IRDs shall accept bitstreams compliant with the T-STD model specified in ISO/IEC 13818-1 [1].* The transport buffer TBn is specified to be 512 bytes. The 302M elementary stream buffer Bn shall have a size of 65 024 (2^{16}-512) bytes.

7.3.3 stream_id

Semantics: The semantics of the stream_id field are described in ISO/IEC 13818-1 [1], table 2-18. Multiple 302M 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 302M elementary stream shall be 0xBD (indicating private_stream_1).*

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

7.3.4 stream_type

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

Encoding: *The value of stream_type for a 302M elementary stream shall be 0x06 (indicating PES packets containing private data).*

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

7.3.5 registration_descriptor

Semantics: The registration descriptor provides a method to uniquely and unambiguously identify that a private data stream is a SMPTE 302M [18] elementary stream.

Encoding: *The SMPTE 302M registration descriptor shall be included in the PMT for every program that includes AES PCM audio or SMPTE/AES data. The format_identifier shall have a value of 0x42535344 as specified by SMPTE 302M. The additional_identification_info bytes shall not be included unless they are defined in a revision of SMPTE 302M, in which case they shall be compliant with the specifications in 302M.*

Decoding: *This descriptor shall be read by the Contribution IRD and shall be used to determine that the data stream with the associated PID value is in fact a 302M stream.* Decoders are not required to interpret information in the additional_identification_info bytes.

7.3.6 Presentation time stamp

Semantics: The semantics of the presentation time stamp are described in ISO/IEC 13818-1 [1]. SMPTE 302M [18] specifies that each PES packet contains one audio access unit. The audio access unit is specified as those AES3 words that are associated with a corresponding video frame. ITU-R Recommendation BT.1359-1 [25] specifies that segments of the broadcast chain that are not under the control of the broadcaster shall not introduce any timing error in excess of ±2 ms.

Encoding: *Each PES header for AES PCM audio or SMPTE/AES data shall contain a PTS. The value of the PTS shall be identical to the value of the PTS that applies to a corresponding video frame. The accuracy of the PTS with respect to the actual payload shall be within ±1 ms.*

Decoding: *This field shall be read by the Contribution IRD and shall be interpreted in accordance with MPEG systems syntax. The Contribution IRD shall reproduce the 302M AES3 payload with a time accuracy, relative to the associated video, of within ±1 msec.*

Annex A (informative): Examples of full screen luminance resolutions for SDTV and HDTV

vertical_size_ value	horizontal_size_ value	aspect_ratio_ information	frame_rate_ code (see note)	progressive_ sequence	*Decodeable by SDTV contribution IRD*
1152	1440	16:9	25	0	
1080	1920	16:9	24, 25	1	
			23,976, 24, 29,97, 30	1	
			25	0	
			29,97, 30	0	
1035	1920	16:9	25	0	
			29,97, 30	0	
720	1280	16:9	24, 25, 50	1	
			23,976, 24, 29,97, 30, 59,94, 60	1	
576	720	4:3, 16:9	50	1	
			24, 25	1	✔
			25	0	✔
	544	4:3, 16:9	24, 25	1	✔
			25	0	✔
	480	4:3, 16:9	24, 25	1	✔
			25	0	✔
	352	4:3, 16:9	24, 25	1	✔
			25	0	✔
480	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	
			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	24, 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.

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 widescreen 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 [16].

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 video syntax).

B.2 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 EC 13818-2 [2]). For example, it could be inserted once per sequence after each sequence exte the sequence extension, and/or GOP header, and/or picture coding extension, as specified in ISO/Insion, 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.3 Syntax and Semantics

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

Table B.1: Active Format Description

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 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 ISO/IEC 13818-2 [2]):

- If sequence_display_extension() is not present:
  - source aspect ratio = 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" 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 this specification 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 diagramatic representation:

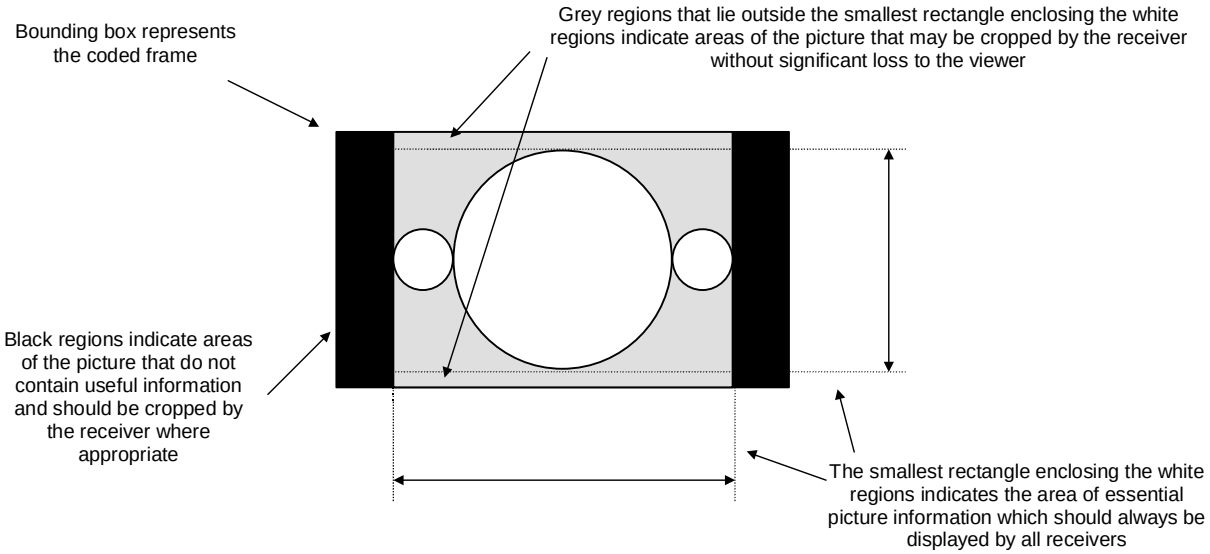
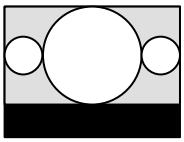
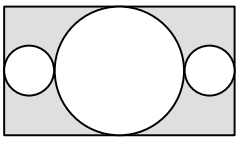
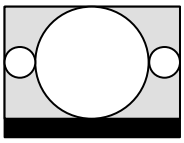
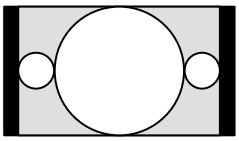
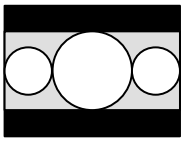
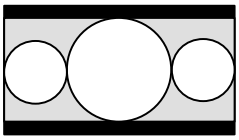
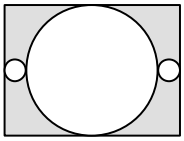
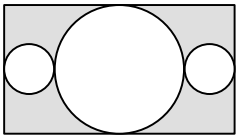
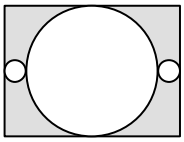
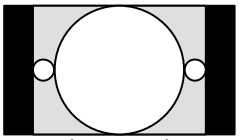
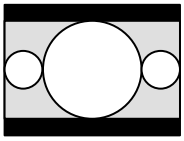
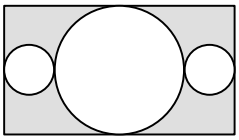
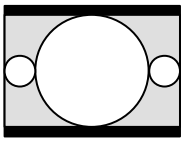
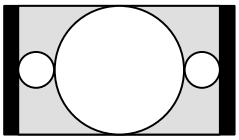
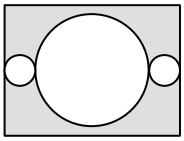
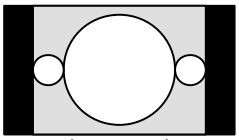
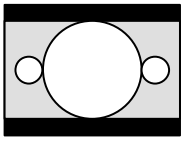
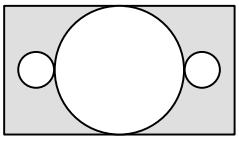
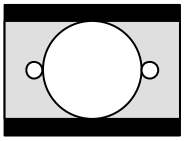
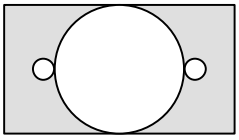


Table B.3: Active Formats Illustrated

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)		 (see note)
1110	16:9 (with shoot & protect 14:9 centre)		
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.4 Relationship with Wide Screen Signalling (WSS)

The AFD provides a super-set of the aspect ratio signalling specified in EN 300 294 [16]. The mapping of source aspect ratio and active_format to WSS Aspect Ratio is given in table B.4.

Table B.4: Support for WSS

Sequence Header	Active Format Description	WSS	
source aspect ratio	**value**	**code (bits 0-3)**	**description**
4:3	1001	0001	full format 4:3
	1011	1000	box 14:9 Centre
	0011	0100	box 14:9 Top
	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 [16]. (Note that the corresponding active lines specified in EN 300 294 [16] do not, in general, apply).

B.6 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 contribution 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).

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

C.1 Scope

The inclusion of AC-3 audio streams in a DVB multiplex is optional, and contribution IRDs may optionally decode these streams. This annex contains the guidelines to include one or more AC-3 elementary streams in a DVB Transport Stream in compliance with ISO/IEC 13818-1 [1] . The coding and decoding of an AC-3 elementary stream is based upon Recommendation ITU-R Recommendation BS.1196-E [14] - annex 2. However, Appendix 1 to annex 2 of ITU-R Recommendation BS.1196-E [14] should be disregarded as it is not applicable to the DVB system.

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

The AC-3 packetized elementary stream shall conform to the requirements of a user private stream type 1, as described in ISO/IEC 13818-1 [1].

The contribution IRD design should be made under the assumption that any legal structure as permitted by 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 contribution IRD shall be able to skip over data structures which are currently "reserved", or which correspond to functions not implemented by the contribution IRD.

This clause is based on ISO/IEC 13818-1 [1] and Recommendation ITU-R Recommendation BS.1196-E [14], annex 2.

C.2 Introduction

An AC-3 elementary bit stream may be multiplexed into an MPEG-2 transport stream in much the same way an MPEG audio stream would be included. The AC-3 elementary stream is packetized into PES packets with a structure similar to an MPEG audio PES. An MPEG-2 transport stream containing 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 stream. A public DVB descriptor, the AC-3_descriptor has 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 ISO 639 [28] language descriptor may be used to indicate the language of the content of the AC-3 stream.

Contribution IRDs shall decode all bit rates and sample rates listed in ITU-R Recommendation BS.1196-E [14], annex 2.

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

The 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 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 also be included. The AC-3_descriptor is defined in EN 300 468 [7], but for information a description is included here in clause C.4.4. The AC-3_descriptor is located in the PMT and the Selection Information Table of the DVB SI tables defined in EN 300 468 [7].

Certain other of the DVB Service Information descriptors defined in EN 300 468 [7] can provide additional means of identifying the existence of an AC-3 stream without accessing the PMT. The component_descriptor (see clause C.4.3) may have values assigned to its syntactical elements which indicate both the presence and type of 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 ISO/IEC 13818-1 [1], table 2-18. Multiple 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 elementary stream shall be 0xBD (indicating private_stream_1).

Decoding: This field shall be read by the contribution IRD, and the contribution 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 ISO/IEC 13818-1 [1], table 2-29.

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

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

C.4.2 Use of the DVB-SI component_descriptor

Semantics: The semantics of the component_descriptor are described in EN 300 468 [7]. The stream_content and component_type assigned values for DVB AC-3 audio streams are listed in annex E, tables E.2 and E.3 of EN 300 468 [7].

Encoding: The values for the elements of the component_descriptor shall be set in accordance with annex E, tables E.2 and E.3 of EN 300 468 [7].

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

C.4.3 Use of the DVB-SI multilingual_component_descriptor

Semantics: The semantics of the multilingual_component_descriptor are described in EN 300 468 [7].

Encoding: The use of multilingual_component_descriptor is optional and should only be used if multiple audio streams are present.

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

C.4.4 AC-3_descriptor

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

Note that 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 clauses. 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 ITU-R Recommendation BS.1196-E [14] (1995) - annex 2.

Table C.1: AC-3 descriptor Syntax

Syntax	No. of Bits	Identifier
AC-3_ descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
AC-3_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 (AC-3_type_flag)==1{		
AC-3_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.4.4.1 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 [7], annex E, table E.1).

Decoding: This field shall be read by the contribution IRD, and the contribution IRD shall interpret this field in accordance with 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 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 contribution IRD, and the contribution IRD shall interpret this field in accordance with ISO/IEC 13818-1 [1].

C.4.4.3 AC-3_type_flag

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

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

C.4.4.4 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: contribution IRDs shall be able to accept bit-streams which contain this field. It is recommended that contribution IRDs decode this field.

C.4.4.5 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: contribution IRDs shall be able to accept bit-streams which contain this field. It is recommended that contribution IRDs decode this field.

C.4.4.6 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: contribution IRDs shall be able to accept bit-streams which contain this field. It is recommended that contribution IRDs decode this field.

C.4.4.7 reserved flags

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

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

C.4.4.8 AC-3_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 [7], annex E, table E.3).

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

C.4.4.9 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: contribution IRDs shall be able to accept bit-streams which contain this field. contribution IRDs may ignore the data within this field.

C.4.4.10 mainid

Semantics: This 8-bit field is optional. It contains a number in the range 0-7 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 0-7.

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

C.4.4.11 asvc

Semantics: This 8-bit field is optional.

Encoding: Each bit (0-7) 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: contribution IRDs shall be able to accept bit-streams which contain this field. contribution IRDs may ignore the data within this field.

C.4.4.12 additional_info[l]

Semantics: These optional bytes are reserved for future use.

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

C.4.5 STD audio buffer size

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

C.5 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 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 $t0$, then audio stream 2 should also have frame n beginning with its sample 0 taken the identical time $t0$. 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.

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.)

C.5.3 Byte-alignment

The 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 frame shall reside in a single byte which is carried by the MPEG-2 data stream.

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

This clause is based on ISO/IEC 13818-3 [3]. The ancillary data format is also compatible with the DVD Video ancillary data format.

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.

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.

D.4 Detailed Specification

D.4.1 DVD-Video Ancillary Data

The transmission of "dynamic_range_control" in MPEG audio is optional. If applied, 16 bit of ancillary data [b15..b0] (situated at the end of each MPEG audio base frame) shall be used.

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.

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: $G = 24{,}082 - 6{,}0206\ X - 0{,}2007\ Y$

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

If the 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: This field shall be read by the decoder, and the decoder shall interpret the value G as a gain value applied to all subband 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 (REPLACE WITH 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 contribution 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 bit stream.

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 1 152 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 subband 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 subband 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 attentuated 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 roomtype 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: This field may be ignored by the decoder.

D.4.2.15 downmixing_levels_MPEG2

The detailed syntax is described on table D.13. The downmixing levels describe the downmix 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 downmix 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 downmix 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 downmix 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 2mc_header is not filled in. If MPEG2_copyright_ident_present is set to '1' the copyright identification in the MPEG 2mc_header is used.

Decoding: This field may be ignored by the decoder. It may be parsed be multiplexers and bit stream monitors to simplify extraction of these parameters from a bit stream.

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 downmix 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 msbs and a value Y in the four lsbs. The actual value is 48,164 - 6,0206 X - 0,4014 Y dB.

Decoding: These values shall be applied to the subband 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 timecode 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 on table D.17.

Table D.17: fine grain timecode syntax

Syntax	No. of bits	Mnemonic

fine_grain_timecode ( ) {
fine_grain_timecode_on
fine_grain_timecode_value
}
	
	2
	14
		
		bslbf
		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 EN 300 401 [26].

Encoding: It recommended to include this data 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.

Annex E (informative): Coding of Data Fields in the Private Data Bytes of the Adaptation Field

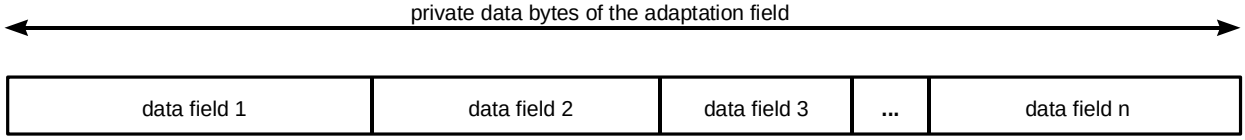
E.1 Introduction

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 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 of the PMT.

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 to 0x9F	Reserved for future use
0xA0 to 0xFF	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.

Annex F (informative): Basic Interoperable Scrambling System (BISS)

F.1 Scope

This annex contains the guidelines for scrambling and descrambling the MPEG-2 transport stream in DVB bit-streams in accordance with the Basic Interoperable Scrambling System (BISS) as specified in EBU Tech 3290 [27]. The application of scrambling to the MPEG-2 transport stream is optional.

F.2 Introduction

The Basic Interoperable Scrambling System [27] is based on the DVB common scrambling algorithm as defined in ETR 289 [6]. The scrambling system supports the following four modes of operation:

Mode 0: No scrambling.

Mode 1: All components are scrambled by a fixed Control Word (CW).

Mode 2: All components are scrambled by a single CW sequence. The Scrambler fixes a CW from the sequence for the duration of the crypto-period.

Mode 3: Each component may be scrambled by a different CW sequence as in Mode 2.

The Scrambler implements the Super Scrambling operations as defined in the DVB Common Scrambling Specification (part 2) in ETR 289 [6]. The scrambling mechanism is applied at Transport level only.

F.3 Specification

Encoding: To support the various modes of operation, the Scrambler inserts Entitlement Control Message (ECM) streams into the multiplex, identified by means of the **CA_descriptor** within the PMT. The value of the **CA_system_ID**, used to indicate the type of CA system applicable for the associated ECM streams, is 0x2600 for mode 1 and 0x2601 for modes 2 and 3. For mode 1, no ECM information is required, so the **CA_PID** contains the value 0x1FFF. A single **CA_system_ID** identifies both modes 2 and 3, however the modes are distinguished by the location of the **CA_descriptor**(s) within the PMT.

The use of Entitlement Management Message (EMM) streams has no application within the modes of operation described within BISS, however compatible equipment may utilize such streams as appropriate.

A Conditional Access Table (CAT) is present in the multiplex for modes 1, 2 and 3, although the table is empty, as no EMM stream will be present. Again, compatible equipment utilizing EMM streams should identify them correctly within the CAT.

A Scrambler that supports only a subset of the defined modes of operation should do so according the following hierarchy:

- a scrambler providing support for mode 2 should also support modes 0 and 1;
- a scrambler providing support for mode 3, should also support modes 0, 1 and 2.

Decoding: A contribution IRD that supports only a subset of the defined modes of operation should do so according to the same hierarchy as for the scrambler:

- a contribution IRD providing support for mode 2 should also support modes 0 and 1;
- a contribution IRD providing support for mode 3, should also support modes 0, 1 and 2.

History

Document history		
V1.1.1	April 2001	Publication