SLOVENSKI STANDARD
SIST ETS 300 743 E1:2005

01-november-2005

Digitalna videoradiodifuzija (DVB) – Sistemi za podnaslove

Digital Video Broadcasting (DVB); Subtitling systems

Ta slovenski standard je istoveten z: ETS 300 743 Edition 1

ICS:

33.170	Televizijska in radijska difuzija	Television and radio broadcasting

SIST ETS 300 743 E1:2005 **en**

EUROPEAN TELECOMMUNICATION STANDARD

ETS 300 743

September 1997

Source: EBU/CENELEC/ETSI-JTC

Reference: DE/JTC-DVB-17

ICS: 33.020

Key words: DVB, digital, video, broadcasting, TV

European Broadcasting Union

Union Européenne de Radio-Télévision

Digital Video Broadcasting (DVB); Subtitling systems

ETSI

European Telecommunications Standards Institute

ETSI Secretariat

Postal address: F-06921 Sophia Antipolis CEDEX - FRANCE
Office address: 650 Route des Lucioles - Sophia Antipolis - Valbonne - FRANCE
X.400: c=fr, a=atlas, p=etsi, s=secretariat - **Internet:** secretariat@etsi.fr

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

Contents

Foreword 5

1 Scope 7

2 Normative references 7

3 Definitions and abbreviations 7
3.1 Definitions 7
3.2 Abbreviations 8

4 Introduction to DVB subtitling system 9
4.1 Overview 9
4.2 Data hierarchy and terminology 10
4.3 Temporal hierarchy and terminology 10

5 Subtitle decoder model 11
5.1 Decoder temporal model 11
5.1.1 Service acquisition 11
5.1.2 Presentation Time Stamps (PTS) 12
5.1.3 Page composition 12
5.1.4 Region composition 12
5.1.5 Points to note 13
5.2 Buffer memory model 13
5.2.1 Pixel display buffer memory 13
5.2.2 Region memory 14
5.2.3 Composition buffer memory 14
5.3 Cumulative display construction 14
5.4 Decoder rendering bandwidth model 14
5.4.1 Page erasure 14
5.4.2 Region move or change in visibility 14
5.4.3 Region fill 15
5.4.4 CLUT modification 15
5.4.5 Graphic object decoding 15
5.4.6 Character object decoding 15

6 PES packet format 16

7 The PES packet data for subtitling 16
7.1 Syntax and semantics of the PES data field for subtitling 16
7.2 Syntax and semantics of the subtitling segment 16
7.2.1 Page composition segment 17
7.2.2 Region composition segment 19
7.2.3 CLUT definition segment 21
7.2.4 Object data segment 22
7.2.4.1 Pixel-data sub-block 24
7.2.4.2 Syntax and semantics of the pixel code strings 25

8 Requirements for the subtitling data 27
8.1 Scope of Identifiers 27
8.2 Scope of dependencies 27
8.2.1 Composition page 27
8.2.2 Ancillary page 27
8.3 Order of delivery 28
8.3.1 PTS field 28
8.4 Positioning of regions and objects 28
8.4.1 Regions 28

8.4.2 Objects sharing a PTS 28
8.4.3 Objects added to a region 28
8.5 Avoiding excess pixel-data capacity 28

9 Translation to colour components 28
9.1 4- to 2-bit reduction 29
9.2 8- to 2-bit reduction 29
9.3 8- to 4-bit reduction 29

10 Default CLUTs and map-tables contents 30
10.1 256-entry CLUT default contents 30
10.2 16-entry CLUT default contents 31
10.3 4-entry CLUT default contents 31
10.4 2_to_4-bit_map-table default contents 32
10.5 2_to_8-bit_map-table default contents 32
10.6 4_to_8-bit_map-table default contents 32

11 Structure of the pixel code strings (informative) 33

Annex A (informative): How the DVB subtitling system works 34

A.1 Data hierarchy and terminology 34

A.2 Temporal hierarchy and terminology 35

A.3 Decoder temporal model 35
A.3.1 Presentation Time Stamps (PTS) 35
A.3.2 Page composition 35
A.3.3 Region composition 36
A.3.4 Points to note 36

A.4 Decoder display technology model 36
A.4.1 Region based with indexed colours 36
A.4.2 Colour quantization 37

A.5 Decoder rendering bandwidth model 37
A.5.1 Page erasure 37
A.5.2 Region move or change in visibility 38
A.5.3 Region erasure 38
A.5.4 CLUT modification 38
A.5.5 Graphic object decoding 38
A.5.6 Character object decoding 38

A.6 Examples of the subtitling system in operation 39
A.6.1 Double buffering 39
A.6.1.1 Instant graphics 39
A.6.1.2 Stenographic subtitles 42

A.7 Glossary 44

History 45

Foreword

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

NOTE: The EBU/ETSI JTC was established in 1990 to co-ordinate the drafting of ETSs in the specific field of broadcasting and related fields. Since 1995 the JTC 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
Case Postale 67
CH-1218 GRAND SACONNEX (Geneva)
Switzerland

Tel: +41 22 717 21 11
Fax: +41 22 717 24 81

Digital Video Broadcasting (DVB) Project

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.

Transposition dates	
Date of adoption:	5 September 1997
Date of latest announcement of this ETS (doa):	31 December 1997
Date of latest publication of new National Standard or endorsement of this ETS (dop/e):	30 June 1998
Date of withdrawal of any conflicting National Standard (dow):	30 June 1998

Blank page

1 Scope

This European Telecommunication Standard (ETS) specifies the method by which subtitles, logos and other graphical elements may be coded and carried in DVB bitstreams. The system applies Colour Look-Up Tables (CLUTs) to define the colours of the graphical elements. The transport of the coded graphical elements is based on the MPEG-2 system described in ISO/IEC 13818-1 [1].

2 Normative references

This ETS incorporates by dated and undated reference, provisions from other publications. These normative references are cited at the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this ETS only when incorporated in it by amendment or revision. For undated references the latest edition of the publication referred to applies.

[1] ISO/IEC 13818-1: "Coding of moving pictures and associated audio".

[2] ETS 300 468: "Digital Video Broadcasting (DVB); Service Information (SI) in DVB systems".

[3] ISO/IEC 10646-1 (1993): "Information Technology - Universal Multiple Octet Coded Character Set (UCS) - Part 1: Architecture and Basic Multilingual Plane".

[4] ITU-R Recommendation 601-3 (1992): "Encoding parameters of digital television for studios".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of this ETS, the following definitions apply:

ancillary page: An optional page that can be used to carry CLUT definition and object data segments that can be shared by more than one subtitle stream. For example, the ancillary page can be used to carry logos or character glyphs.

Colour Look-Up Table (CLUT): A look-up table applied in each region for translating the objects' pseudo-colours into the correct colours on the screen. In most cases, one CLUT is sufficient to present correctly the colours of all objects in a region, but if it is not enough, then the objects can be split horizontally into smaller objects that, combined in separate regions, need not more than one CLUT per region.

CLUT-family: A family of CLUTs which consists of:

- one CLUT with 4 entries;
- one CLUT with 16 entries;
- one CLUT with 256 entries.

NOTE 1: Three CLUTs are defined to allow flexibility in the decoder design. Not all decoders may support a CLUT with 256 entries, some may provide sixteen or even only four entries. A palette of four colours would be enough for graphics that are basically monochrome, like subtitles, while a palette of sixteen colours allows for cartoon-like coloured objects. Having a CLUT of only four entries does not imply that only a rigid colour scheme can be used. The colours that correspond to the four entries can be redefined, for instance from a black-grey-white scheme to a blue-grey-yellow scheme. Furthermore, a graphical unit may be divided into several regions that are linked to different CLUTs, i.e. a different colour scheme may be applied in each of the regions.

composition page: The page which carries the page composition. This page may contain graphical elements as well. Those elements that may be shared by different screen layouts are carried in an "ancillary page".

NOTE 2: Thus, alternative screen layouts, defined as different page compositions, may use the same CLUTs and objects. There is no need to convey the common information for each screen layout separately. This sharing is particularly useful when subtitles are provided in several languages, all combined with the same logo. To retain flexibility, the position at which a region is shown on the screen is not a property of that region itself, but defined in the page composition, so that a shared region may be shown in different locations on different screen layouts.

decoder state: Pixel and composition buffer memory allocations and values.

display: A completed set of graphics.

display set: The set of segments that operate on the decoder state between page composition segments to produce a new display.

display sequence: A sequence of one or more displays.

epoch: The period between resets to the decoder state caused by page composition segments with page state = "mode change".

object: Anything that can be presented on a TV screen, e.g. a subtitle, a logo, a map, etc. An object can be regarded as a graphical unit. Each has its own unique ID-number.

packet identifier: See ISO/IEC 13818-1 [1].

page composition: The top-level definition of a screen layout. Several regions may be shown simultaneously on the screen; those regions are listed in the page composition. At any one time, only one page composition can be active for displaying, but many may be carried simultaneously in the bitstream.

PES packet: See ISO/IEC 13818-1 [1].

pixel-data: A string of data bytes that contains, in coded form, the representation of a graphical object.

region: A rectangular area on the screen in which objects are shown. Objects that share one or more horizontal scan lines on the screen are included in the same region.

NOTE 3: A region therefore monopolizes the scan lines of which it occupies any part; no two regions can be presented horizontally next to each other.

transport packet: See ISO/IEC 13818-1 [1].

transport packet stream: A sub-set of the transport packets in a transport stream sharing a common Packet Identifier (PID).

transport stream: See ISO/IEC 13818-1 [1]. A data stream carrying one or more MPEG programs.

subtitle stream: A stream of subtitling segments that when decoded will provide a sequence of subtitling graphics meeting a single communication requirement (e.g. the graphics to provide subtitles in one language for a one program). A subtitling stream may contain data from a single page (the composition page) or from two pages (the composition page and the ancillary page).

3.2 Abbreviations

For the purposes of this ETS, the following abbreviations apply:

bslbf	bit string, left bit first
Cb	as defined in ITU-R Recommendation 601-3 [4] (see subclause 7.2.3)
CLUT	Colour Look-Up Table
Cr	as defined in ITU-R Recommendation 601-3 [4] (see subclause 7.2.3)
DVB	Digital Video Broadcasting
IRD	Integrated Receiver Decoder
MPEG	Moving Pictures Experts Group

PCR	Programme Clock Reference
PCS	Page Composition Segments
PES	Packetized Elementary Stream
PID	Packet IDentifier
PMT	Program Map Table
PTS	Presentation Time Stamp
RCS	Region Composition Segments
ROM	Read-Only Memory
TS	Transport Stream
uimsbf	unsigned integer, most significant bit first
Y	as defined in ITU-R Recommendation 601-3 [4] (see subclause 7.2.3)

4 Introduction to DVB subtitling system

This ETS specifies the transport and coding of graphical elements in the DVB subtitling system.

4.1 Overview

To provide efficient use of the display memory in the decoder this subtitling system uses region based graphics with indexed pixel colours. Each display is composed of a number of regions with specified position. A region is a rectangular area with a horizontal and vertical size, pixel depth. A region can have a defined background colour and graphical objects can be positioned within the region.

Pixel depths of 2, 4 and 8-bits are supported allowing up to 4, 16 or 256 different pixel codes to be used in each region. Each region is associated with a CLUT which defines the colour and transparency for each of the pixel codes.

At the discretion of the encoder, objects designed for displays supporting 16 or 256 colours can be decoded into displays supporting fewer colours. A quantization algorithm is defined to ensure that this process is predictable by the originator. This feature allows a single data stream to be decoded by a population of decoders with mixed, and possibly evolving, capabilities.

This subtitling system provides a number of techniques that allow efficient transmission of the graphic data:

- pixel structures that occur more than once within a bitmap can be transmitted only once, and then positioned multiple times within the bitmap;

- pixel structures used in more than one subtitle stream shall only be transmitted once;

- pixel data is compressed using run-length coding;

- where the gamut of colours required for part of a graphical object is suitably limited, that part can be coded using a smaller number of bits per pixel and a map table. For example, an 8-bit per pixel graphical object may contain areas coded as 4 or 2-bits per pixel each preceded by a map table to map the 16 or 4 colours used onto the 256 colour set of the region. Similarly, a 4-bit per pixel object may contain areas coded as 2-bits per pixel;

- colour definitions can be coded using either 16 or 32-bits per CLUT entry. This provides a trade off between colour accuracy and transmission bandwidth.

The above features require only compliance with this ETS. Additional features are provided that allow more efficient operation where there are additional agreements between the data provider and the manufacturer of the decoder:

- graphic objects resident in ROM in the decoder can be referenced;
- character codes, or strings of character codes, can be used in place of graphic object references. This requires the decoder to be able to generates glyphs for these codes.

This ETS is not concerned with the private agreements required to make these features operate.

4.2 Data hierarchy and terminology

The "building block" of the subtitling information is the subtitling_segment. These segments are carried in PES packets which are in-turn carried by Transport Packets.

All the broadcast data required for a subtitle stream will be carried by a single transport packet stream (i.e. on a single PID). A single transport packet stream can carry several different streams of subtitles. The different subtitle streams can be subtitles in different languages for a common program. Alternatively, they can be for different programs (provided that the programs share a common PCR).

Different subtitle streams can also be supplied to address different display characteristics or to address special needs. For instance:

- different subtitle streams can be provided for 4:3 and 16:9 aspect ratio displays;
- subtitle streams can be provided for viewers with impaired hearing. These may include graphical representations of sounds.

Within a transport packet stream the segments for different subtitling streams are identified by their page identifiers. One or more subtitling_descriptors ETS 300 468 [2] in the PMT for a program describe the available subtitling streams and specify the PID and page ids that shall be decoded for each subtitling stream.

A subtitling stream may contain data from a single page (the composition page) or from two pages (the composition page and the ancillary page). The ancillary page can be used to carry objects that are common to 2 or more subtitle streams. For example, the ancillary page can carry a logo that is common to subtitle streams for several different languages.

The PTS in the PES packet provides presentation timing information for the subtitling data. The number of segments carried by each PES packet is only limited by the maximum length of a PES packet defined by MPEG.

In summary the data hierarchy is:

- Transport Stream (TS);
- transport packet stream (common PID);
- PES (provides timing);
- subtitle stream (composition or composition and ancillary pages);
- page;
- segment.

4.3 Temporal hierarchy and terminology

At the segment level in the data hierarchy there is temporal hierarchy. The highest level is the epoch. This is analogous to the MPEG video sequence. No decoder state is preserved from one epoch to the next.

An epoch is a sequence of one or more displays. Each display is a completed screen of graphics. Consecutive displays may differ little (e.g. by a single word when stenographic subtitling is being used) or may be completely different. The set of segments that form each display is called a display set.

Within a display set the sequence of segments (when present) is:

- page composition;
- region composition;
- CLUT definition;
- object data.

All segments associated with composition page shall be delivered before any segments from the optional ancillary page. The ancillary page may only carry CLUT definition or object data segments.

5 Subtitle decoder model

The subtitle decoder model is an abstraction of the processing required for the interpretation of subtitling streams. The main purpose of this model is to define a number of constraints which can be used to verify the validity of subtitling streams. The following figure shows a typical implementation of a subtitling decoding process in a receiver.

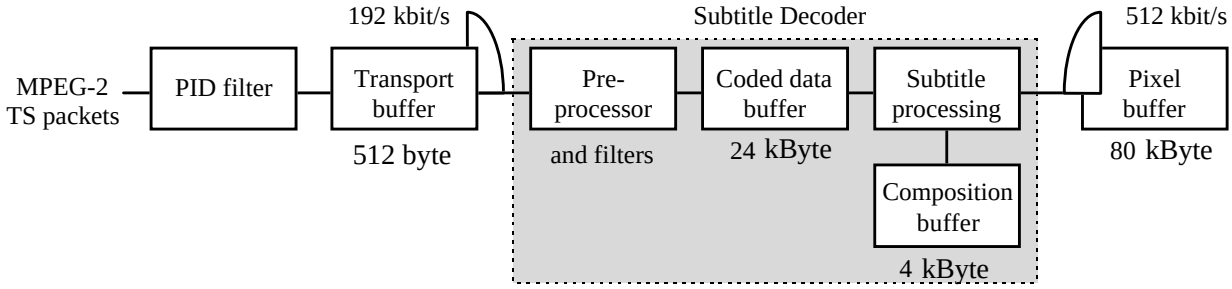

Figure 1: Subtitle decoder model

The input to the subtitling decoding process is an MPEG-2 Transport Stream (TS). After a selection process based on PID value, complete MPEG-2 Transport Stream packets enter into a transport buffer with a size of 512 byte. When there is data in the transport buffer, data is removed from this buffer with a rate of 192 kbit/s. When no data is present, the data rate equals zero.

The MPEG-2 transport stream packets from the transport buffer are processed by stripping off the packet headers of TS packets and of Packetized Elementary Stream (PES) packets with the proper data_identifier value. The Presentation Time Stamp (PTS) fields shall be passed on to the next stages of the subtitling processing. The output of the pre-processor is a stream of subtitling segments which are filtered based on their page_id values.

The selected segments enter into a coded data buffer which has a size of 24 kbyte. Only complete segments are removed from this buffer by the subtitle decoder. The removal and decoding of the segments is instantaneous (i.e. it takes zero time). If a segment produces pixel data, the subtitle decoder stops removing segments from the coded data buffer until all pixels have been transmitted to the pixel buffer. The rate for the transport of pixel data into the pixel buffer is 512 kbit/s.

5.1 Decoder temporal model

A complete description of the memory use of the decoder shall be delivered at the start of each epoch. Hence, epoch boundaries provide a guaranteed service acquisition point. Epoch boundaries are signalled by page composition segments with a page state of "mode change".

The pixel buffer and the composition buffer hold the state of the subtitling decoder. The epoch for which this state is defined is between Page Composition Segments (PCSs) with page state of "mode change". When a PCS with state of "mode change" is received by a decoder all memory allocations implied by previous segments are discarded i.e. the decoder state is reset.

All the regions to be used in an epoch shall be introduced by the Region Composition Segments (RCSs) in the display set that accompanies the PCS with page state of "mode change" (i.e. the first display set of the epoch). This requirement allows a decoder to plan all of its pixel buffer allocations before any object data is written to the buffers. Similarly, all of the CLUT entries to be used during the epoch shall be introduced in this first display set. Subsequent segments can modify the values held in the pixel buffer and composition buffer but may not alter the quantity of memory required.

5.1.1 Service acquisition

The other allowed values of page state are "acquisition point" and "normal case". The "acquisition point" state (like the "mode change" state) indicates that a complete description of the memory use of the decoder is being broadcast. However, the memory use is guaranteed to be the same as that previously in operation. Decoders that have already acquired the service shall only look for development of the existing display (e.g. new graphical objects to be decoded). Decoders trying to acquire the service can treat a page state of "acquisition point" as if it is "mode change".