DVB webinar on subtitling standards

24th January, 2018

Introduction

- Subtitles (or captions) are an important feature of TV services; they are used to provide:
  - Dialogue transcripts in alternative languages
  - Service accessibility aids, e.g. Audio Description
  - Commentary or other ancillary information
- The first DVB subtitle specifications were for Teletext subtitles
  - EN 300 472 and EN 301 775, first published in 1994
- Graphical bitmap DVB Subtitling – EN 300 743
  - First published (V1.1.1) in 1997
  - Revised version V1.6.1 has been completed; publication by ETSI is under way
- Both of the existing specifications are still in wide use
- The new DVB TTML subtitling specification (Bluebook A174) specifies a mechanism to convey TTML subtitles in DVB systems – EN 303 560
  - Initial version V1.1.1 has been completed; publication by ETSI is under way

Webinar contents

- Part 1: Revised bitmap subtitling specification
  - prEN 300 743 V1.6.1 / DVB blue book A009
- Part 2: TTML introduction
- Part 3: New DVB TTML subtitling specification
  - prEN 303 560 V1.1.1 / DVB blue book A174
- Part 4: Deployment considerations

DVB (bitmap) subtitling specification revision prEN 300 743 V1.6.1

DVB webinar (part 1), 24th Jan. 2018

Paul Szucs, Sony Europe

Specification editor and ETSI work item rapporteur

paul.szucs@sony.com

Contents

- DVB subtitling specification – history
- Rationale for the latest revision
- Overview of the latest revision (V1.6.1)
  - General re-factoring
  - Bug-fixes and clarifications
  - New features

DVB Subtitling (EN 300 743) - history

- The original specification (V1.1.1) was published in 1997; it specified the bitmap subtitles format for use with DVB SDTV services
  - V1.2.1 (2002) was a general maintenance revision
- V1.3.1 (2006) added support for subtitles with HDTV services
- V1.4.1 (2011) added support for subtitles with 3DTV services
  - V1.5.1 (2014) was a general maintenance revision
- V1.6.1 (2018*) adds support for subtitles with UHDTV services, in conjunction with two new features

* ETSI publication is expected within 2018; in the meantime it is available as DVB blue book A009

DVB Subtitling - rationale for the latest revision

- The main reason to undertake the revision of EN 300 743 was to ensure compatibility with UHDTV services, i.e. DVB UHD-1 both Phases 1 and 2
- Despite the ongoing work on text-based next-generation subtitles, bitmap-based subtitles will continue to be a relevant feature in DVB broadcast systems for UHDTV, as well as for SDTV and HDTV services
- Bitmap-based subtitles offer a more pragmatic and efficient approach in DVB deployment regions where support is needed for subtitles in many languages with different fonts, some of which are extremely complex. This is especially the case in south-east Asia.
- General maintenance and minor bug-fixing was also needed

DVB Subtitling (EN 300 743) – what's new in V1.6.1?

- The specification has been extensively revised editorially in order to make it compliant with the latest ETSI specification drafting guidelines
- Clause 4 has undergone a major re-factoring; it is now a purely informative introduction; normative statements have been moved to the appropriate subsequent clauses
- Bug-fix applied - ambiguity in the Object Data Segment syntax has been resolved, about when *8_stuff_bits field* is present; the function *wordaligned()* had not been defined
- Added support for UHDTV services; subtitle bitmaps are limited to a maximum resolution of 1920x1080, but can use the existing DDS display window feature to position subtitles within the UHDTV (3840x2160) display raster
- New alternative CLUT segment, to enable subtitles to be provided in colour systems other than ITU-R BT.601
- New progressive bitmap encoding method based, aligning the compression method with the PNG format
- Clarifications added around backward compatibility with subtitle streams and IRDs

DVB Subtitling for UHDTV services

- The Display Definition Segment (DDS) shall be included in subtitle streams intended for UHDTV services, whereby subtitle graphics rendering is constrained to HDTV resolution (up to 1920 x 1080 pixels)
- Where the display window feature of the DDS is not used, the UHDTV IRD shall up-scales subtitles spatially before rendering them on a UHDTV resolution display

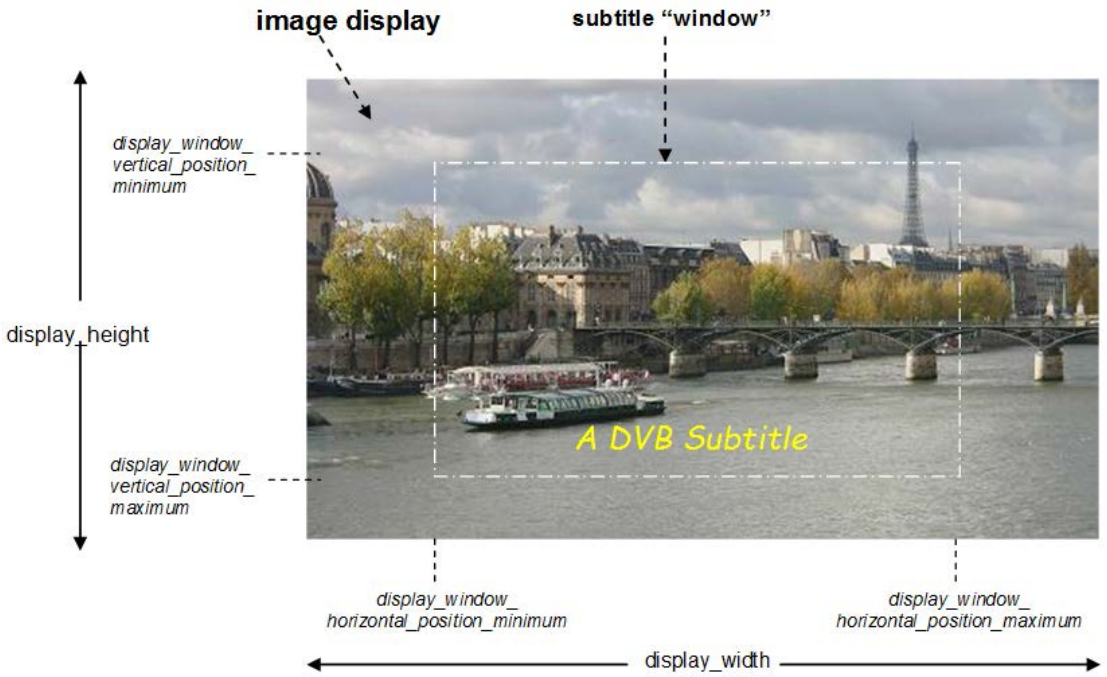
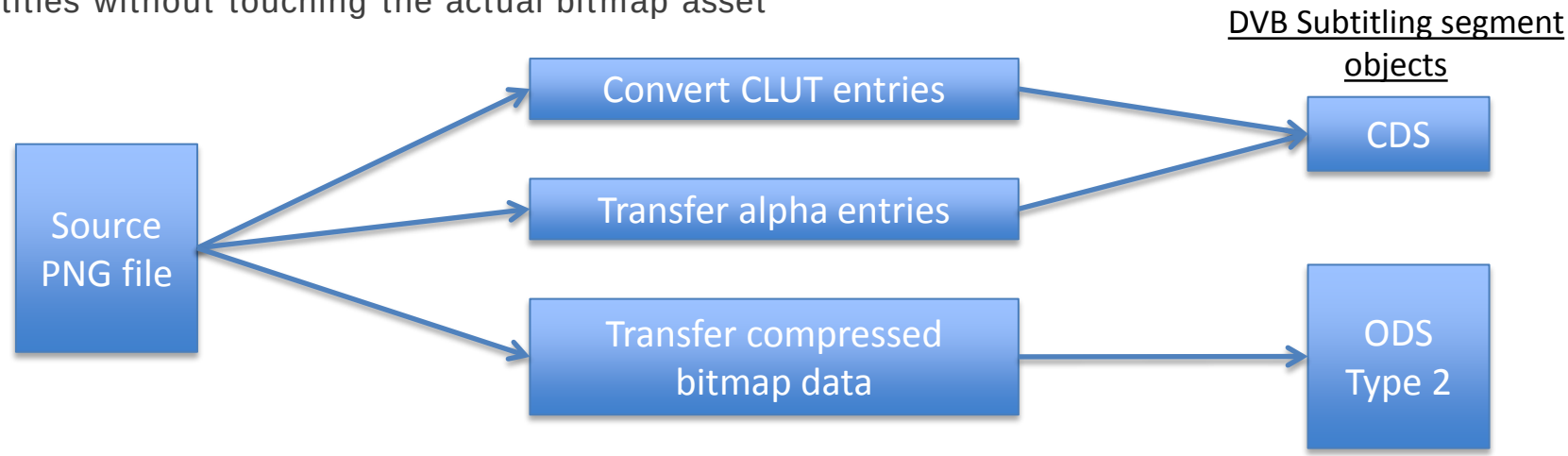


Figure 4: Use of Display definition segment parameters

Progressively coded subtitle bitmaps

- Prior to V1.6.1, the working assumption was that subtitles were coded in interlaced format, i.e. consisting of two fields of bitmap data
- V1.6.1 adds a new Object Data Segment (ODS) *coding_type* (=2) for progressively coded subtitle bitmaps, which consist of a single block of bitmap data
- The progressively coded bitmap subtitles option makes use of the same compression technique as used in the Portable Network Graphics (PNG) format, rather than the legacy run-length coding algorithm that is unique to EN 300 743
- This enables more convenient re-use of subtitling assets that might be deployed in multiple target delivery systems; suitably encoded PNG files can be converted into DVB progressively coded bitmap subtitles without touching the actual bitmap asset



Colour and HDR

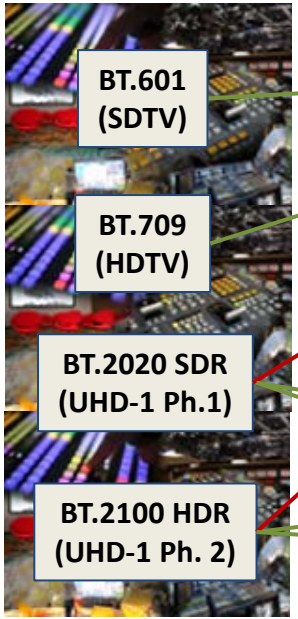
- All previous versions of the specification allowed subtitles to be delivered **only in the ITU-R BT.601** (SDTV) colour space, by use of a CLUT (Colour Look-Up Table), carried in the CLUT Definition Segment (CDS) (if the default CLUTs are not used)
  - Even when HDTV services use ITU-R BT.709 (HDTV) colour space
  - Default CLUTs were specified for use with 4, 16, and 256-colour graphics systems; these provide an adequate pre-defined wide range of colours tailored for each system capability
- In V1.6.1 the new feature of the **Alternative CLUT Segment (ACS)** allows subtitles to be delivered in **other colour and dynamic range systems**, e.g. ITU-R BT.709, ITU-R BT.2020 and ITU-R BT.2100 - PQ10 or HLG
  - So that subtitles that might be authored in these other systems have their colours maintained deterministically through to the decoder
  - Schematic subtitle workflows are shown in the next slide
- Use/support of the CDS and/or default CLUTs remains to be mandatory
- Use of the ACS is optional for subtitle content and its support is optional for the subtitle decoder

Schematic diagram of subtitle workflows

Production
(Video, including subtitles)

Subtitle service
CLUT segments

IRD



CDS

Subtitles are rendered using CDS and/or default CLUTs

SDTV, HDTV, UHD-1; EN 300 743 V1.5.1 & earlier

HDR- and WCG-capable IRDs might apply some arbitrary conversion to subtitle properties

The CDS shall always be provided, so IRDs that do not support ACS will use CDS

Subtitles graded in BT.2100 / BT.2020 are converted to BT.601

CDS

ACS

UHD-1 Ph.2, UHD-1 Ph.1; EN 300 743 V1.6.1

Subtitles graded in BT.2100 preserve their intended characteristics

Using the ACS allows rendering subtitles as close as possible to the service's intended appearance

Clarifications around backward compatibility

- V1.6.1 includes explicit compliance recommendations to align service signalling and the decoder
- The *subtitling_type* signalled in the PMT in effect indicates the version of the specification with which the subtitle stream complies, and hence the minimum version of the subtitle decoder needed to decode the stream

Table 5: Subtitling type usage

Subtitling type in the subtitling_descriptor (see ETSI EN 300 468[2])	EN 300 743 version compliance	Indicative service compatibility	Features that are not recommended for the subtitle service
0x10-0x13, 0x20-0x23	V1.1.1, V1.2.1	SDTV	DDS, DSS, ACS, ODS object coding type = '2'
0x14, 0x24	V1.3.1	HDTV, UHDTV [1]	DSS, ACS, ODS object coding type = '2'
0x15, 0x25	V1.4.1, V1.5.1	3DTV	ACS, ODS object coding type = '2'
0x16, 0x26	V1.6.1	HDTV [2], UHDTV	None

NOTE 1: The subtitle service may use only the CLUT definition segment (CDS) to define the available subtitle colours within the ITU-R BT.601 [3] colour system.

NOTE 2: The subtitle service may use ODS object coding type = '2' but in that case decoders compliant with V1.5.1 or earlier of the present document will not be able to decode the subtitles.

In summary…

- *EN 300 743 V1.6.1 (DVB Subtitling) maintains the relevance of the DVB bitmap subtitles specification for the Ultra-HDTV age, as one of the three approaches to subtitling available in DVB systems*


DVB

Webinar on DVB Subtitling Systems

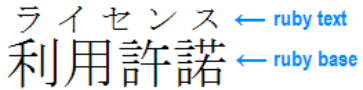
Part 2 – Stefan Pöschel (IRT)

TTML introduction

TTML

- Timed Text Markup Language 1 (**TTML1**), W3C 2013
  - was: Distribution Format Exchange Profile (DFXP)
- based on XML → extensible by new features/extensions, metadata
- powerful/complex → has to be restricted!

- **TTML2**: Working Draft; e.g. ruby annotations
  - TTML1 content backward compatible



TTML styling features (incomplete)

text/background color; opacity

bold/*italic*/underline

font size/line height

font family

bi-directional text

alignment horizontal/vertical

line wrapping/padding

line padding (EBU-TT)

multi row align
(EBU-TT)

fill line gap (IMSC1 Text)

TTML example

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xmlns="http://www.w3.org/ns/ttml" xmlns:tts="http://www.w3.org/ns/ttml#styling" xml:lang="en">
      <head>
            <styling>
                  <style xml:id="pStyle" tts:textAlign="center"/>
                  <style xml:id="sStyle" tts:color="#ffffff" tts:backgroundColor="#000000" tts:fontSize="160%"
                        tts:fontFamily="monospace" tts:fontStyle="normal"/>
            </styling>
            <layout>
                  <region xml:id="bottom" tts:origin="10% 10%" tts:extent="80% 80%" tts:displayAlign="after"/>
            </layout>
      </head>
      <body>
            <div>
                  <p xml:id="sub1" region="bottom" begin="00:00:05.000" end="00:00:10.300" style="pStyle">
                        <span style="sStyle">One Subtitle.</span>
                  </p>
                  <p xml:id="sub2" region="bottom" begin="00:00:22.000" end="00:00:30.000" style="pStyle">
                        <span style="sStyle">A second Subtitle.</span>
                        <br/>
                        <span style="sStyle">With a second line</span>
                  </p>
            </div>
      </body>
</tt>
```

TTML ISD

- Intermediate Synchronic Document (ISD), virtual TTML document
  - represents particular displayed state between two “changes”
- rendering an TTML document ↔ consecutive rendering of resulting ISDs

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xmlns="http://www.w3.org/ns/ttml" xmlns:tts="http://www.w3.org/ns/ttml#styling" xml:lang="en">
  <head>
    <styling>
      <style xml:id="pStyle" tts:textAlign="center"/>
      <style xml:id="sStyle" tts:color="#ffffff" tts:backgroundColor="#000000" tts:fontSize="160%"
        tts:fontFamily="monospace" tts:fontStyle="normal"/>
    </styling>
    <layout>
      <region xml:id="bottom" tts:origin="10% 10%" tts:extent="80% 80%" tts:displayAlign="after"/>
    </layout>
  </head>
  <body>
    <div>
      <p xml:id="sub1" region="bottom" begin="00:00:05.000" end="00:00:10.300" style="pStyle">
        <span style="sStyle">One Subtitle.</span>
      </p>
      <p xml:id="sub2" region="bottom" begin="00:00:22.000" end="00:00:30.000" style="pStyle">
        <span style="sStyle">A second Subtitle.</span>
        <br/>
        <span style="sStyle">With a second line</span>
      </p>
    </div>
  </body>
</tt>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xmlns="http://www.w3.org/ns/ttml" xmlns:tts="http://www.w3.org/ns/ttml#styling" xml:lang="en">
  <head>
    <styling>
      <style xml:id="pStyle" tts:textAlign="center"/>
      <style xml:id="sStyle" tts:color="#ffffff" tts:backgroundColor="#000000" tts:fontSize="160%"
        tts:fontFamily="monospace" tts:fontStyle="normal"/>
    </styling>
    <layout>
      <region xml:id="bottom" tts:origin="10% 10%" tts:extent="80% 80%" tts:displayAlign="after">
        <body>
          <div>
            <p xml:id="sub2" style="pStyle">
              <span style="sStyle">A second Subtitle.</span>
              <br/>
              <span style="sStyle">With a second line</span>
            </p>
          </div>
        </body>
      </region>
    </layout>
  </head>
</tt>
```

EBU-TT-D

- **EBU-TT-D 1.0** (EBU Tech 3380), EBU 2015
  - based on contribution format EBU-TT (TTML1 based)
  - considers subtitling requirements of EBU's broadcasters
  - used by HbbTV 2.0 (further constraints/implications) and DVB DASH
- restrictions, e.g. nesting, timing locations

IMSC1

- TTML Profiles for Internet Media Subtitles and Captions 1.0 (**IMSC1**), W3C 2016
  - use features from EBU-TT-D/SMPTE-TT/CFF-TT/SDP-US, for usable subset
  - Text Profile (+ Image Profile; both profiles used by ATSC 3.0)
    - superset of EBU-TT-D (but minor exceptions)

Upcoming EBU-TT-D/IMSC1 updates

- **EBU-TT-D 1.0.1d**, draft
  - **refered to by DVB TTML spec**
  - adopting two new attributes of IMSC 1.0.1
  - currently waiting for IMSC 1.0.1 Recommendation

- **IMSC 1.0.1**: Candidate Recommendation, 2017
  - **Text Profile refered to by DVB TTML spec**
  - adds activeArea + fillLineGap (on DVB TM-SUB's request)
  - currently waiting for second implementation per feature to progress
- **IMSC 1.1**: Working Draft; refers to TTML2

DVB default conformance point ...

- re-use existing EBU-TT-D implementations (HbbTV 2.0, DVB DASH)
- but also support popular IMSC1 (Text) standard
- decoders of either profile able to process it!
- support mandatory (→ DVB TTML spec slides)

- additional requirements for EBU-TT-D files
  - UTF-8
  - Hypothetical Render Model (HRM) can be applied
  - not more than four regions active at the same time

… and additional decoder requirements

- generic font family name mappings, e.g. “proportionalSansSerif”:
  - IMSC1 Text: glyph dimensions like at Arial/Helvetica/Liberation Sans
  - HbbTV: Tiresias
- support itts:fillLineGap / ebutts:linePadding / ebutts:multiRowAlign
- initial text color: white (implementation-dependent in EBU-TT-D)


DVB

Webinar on DVB Subtitling Systems

Part 3 – Peter Cherriman (BBC)

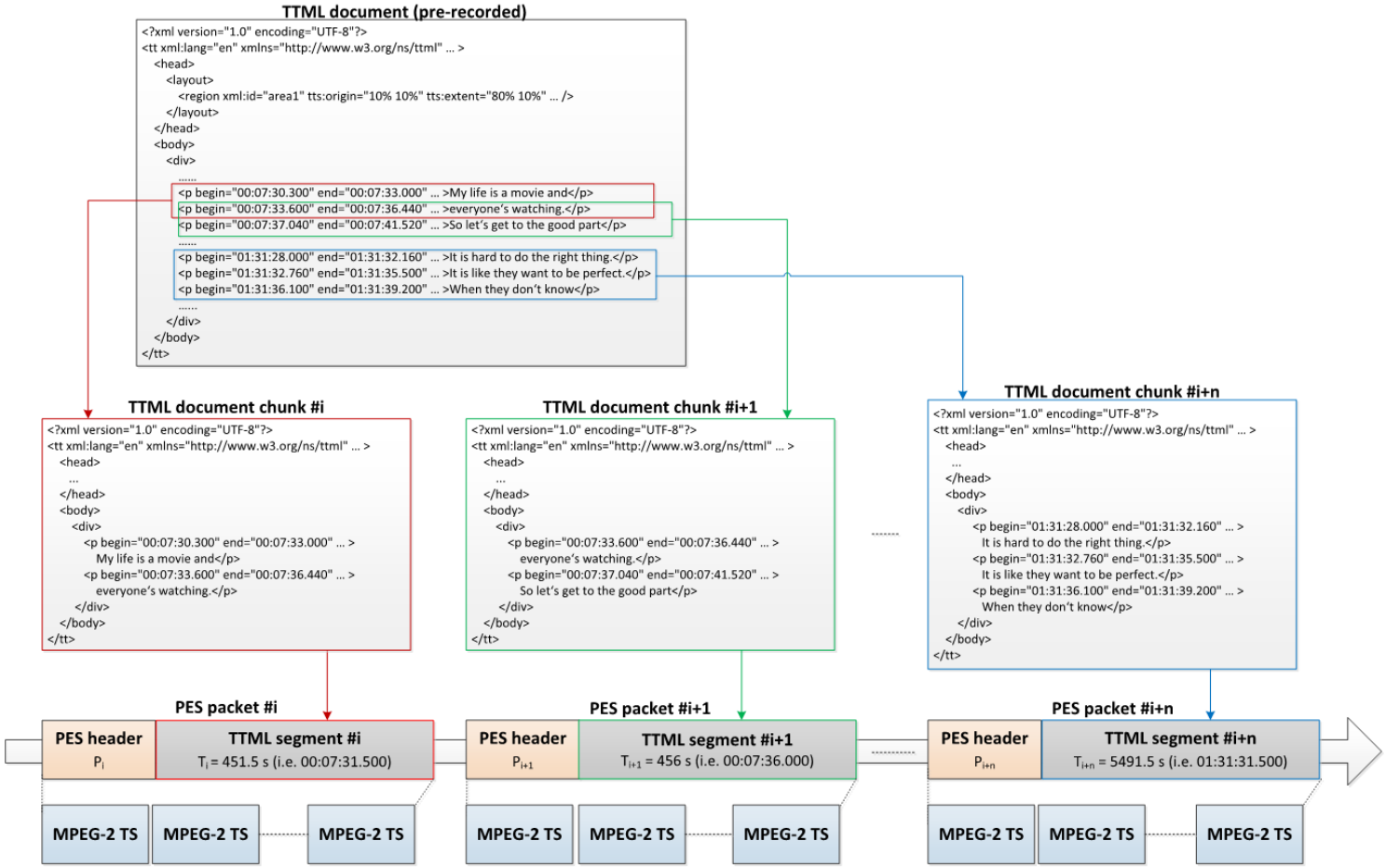
TTML subtitle delivery specification

DVB TTML Subtitle specification

- DVB Bluebook A174 is available at https://www.dvb.org/standards
- Submitted to ETSI as prospective EN 303 560
- It defines a mechanism for delivering TTML documents over transport stream
- Currently three TTML processor profiles are supported: EBU-TT-D v1.0.1d, IMSC1.0.1, DVB default conformance point
- IRD requirements and recommendations are also defined. These include required fonts and font mappings needed for compatibility reasons.

TTML Transport mechanism

- The DVB bluebook A174 defines a PES payload format for conveying TTML documents (can be gzip compressed)
- It provides a mechanism for synchronising the TTML document’s timeline values with the PTS in the PES packet conveying the document.
- It defines a mechanism for segmenting TTML documents into short duration standalone TTML documents that can be conveyed in a PES packet.


TTML document (pre-recorded)
<?xml version="1.0" encoding="UTF-8"?>
<tt xml:lang="en" xmlns="http://www.w3.org/ns/ttml" ... >
<head>
<layout>
<region xml:id="area1" tts:origin="10% 10%" tts:extent="80% 10%" ... />
</layout>
</head>
<body>

.....
<p begin="00:07:30.300" end="00:07:33.000" ... >My life is a movie and</p>
<p begin="00:07:33.600" end="00:07:36.440" ... >everyone's watching.</p>
<p begin="00:07:37.040" end="00:07:41.520" ... >So let's get to the good part</p>
.....
<p begin="01:31:28.000" end="01:31:32.160" ... >It is hard to do the right thing.</p>
<p begin="01:31:32.760" end="01:31:35.500" ... >It is like they want to be perfect.</p>
<p begin="01:31:36.100" end="01:31:39.200" ... >When they don't know</p>
.....

</body>
</tt>
TTML document chunk #i
<?xml version="1.0" encoding="UTF-8"?>
<tt xml:lang="en" xmlns="http://www.w3.org/ns/ttml" ... >
<head>
...
</head>
<body>

<p begin="00:07:30.300" end="00:07:33.000" ... >
My life is a movie and</p>
<p begin="00:07:33.600" end="00:07:36.440" ... >
everyone's watching.</p>

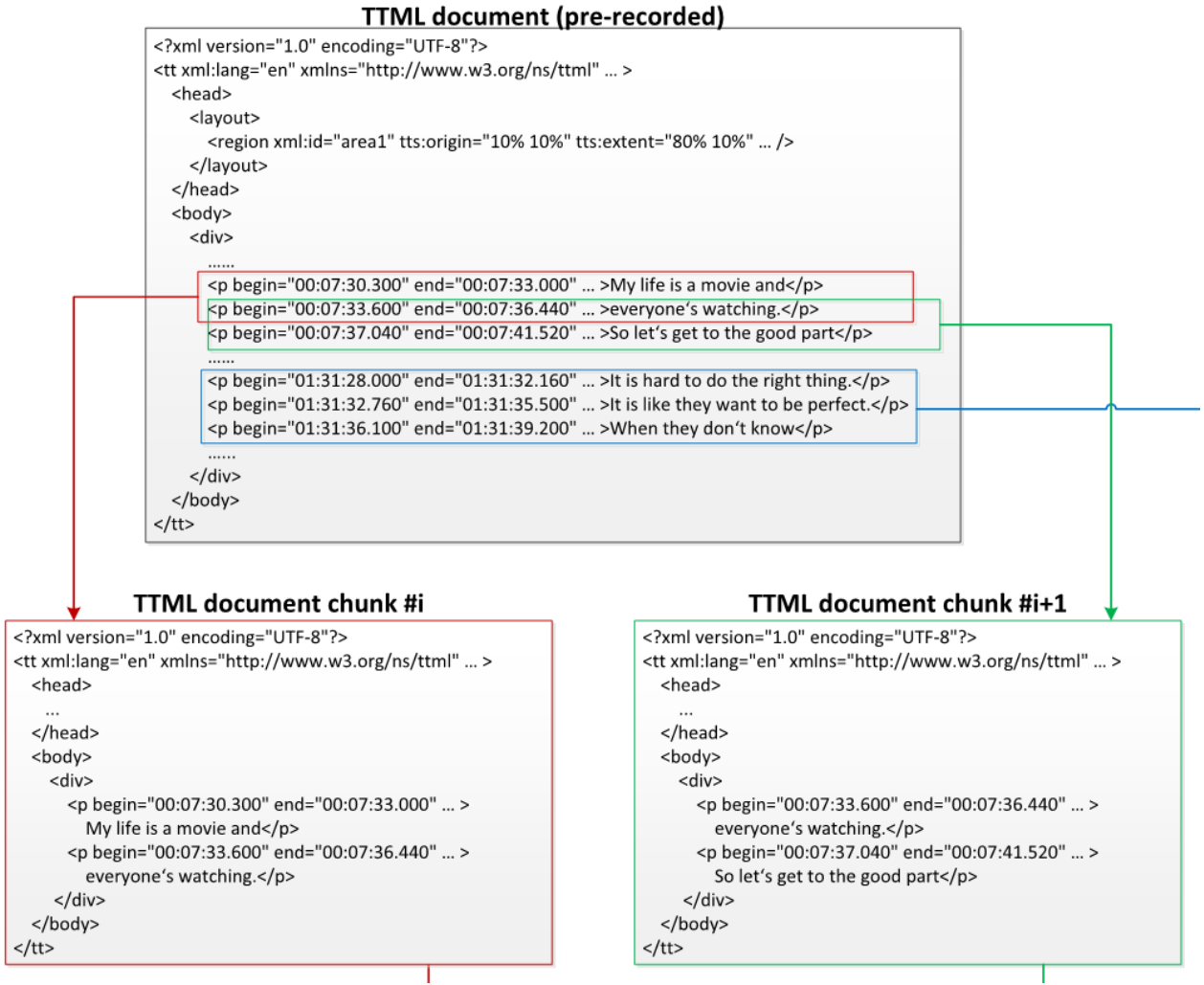
</body>
</tt>
TTML document chunk #i+1
<?xml version="1.0" encoding="UTF-8"?>
<tt xml:lang="en" xmlns="http://www.w3.org/ns/ttml" ... >
<head>
...
</head>
<body>

<p begin="00:07:33.600" end="00:07:36.440" ... >
everyone's watching.</p>
<p begin="00:07:37.040" end="00:07:41.520" ... >
So let's get to the good part</p>

</body>
</tt>
TTML document chunk #i+n
<?xml version="1.0" encoding="UTF-8"?>
<tt xml:lang="en" xmlns="http://www.w3.org/ns/ttml" ... >
<head>
...
</head>
<body>

<p begin="01:31:28.000" end="01:31:32.160" ... >
It is hard to do the right thing.</p>
<p begin="01:31:32.760" end="01:31:35.500" ... >
It is like they want to be perfect.</p>
<p begin="01:31:36.100" end="01:31:39.200" ... >
When they don't know</p>

</body>
</tt>
PES packet #i
PES header
P_i
TTML segment #i
T_i = 451.5 s (i.e. 00:07:31.500)
PES packet #i+1
PES header
P_{i+1}
TTML segment #i+1
T_{i+1} = 456 s (i.e. 00:07:36.000)
PES packet #i+n
PES header
P_{i+n}
TTML segment #i+n
T_{i+n} = 5491.5 s (i.e. 01:31:31.500)
MPEG-2 TS
MPEG-2 TS
MPEG-2 TS
MPEG-2 TS
MPEG-2 TS
MPEG-2 TS
MPEG-2 TS
MPEG-2 TS
MPEG-2 TS


TTML document (pre-recorded)
<?xml version="1.0" encoding="UTF-8"?>
<tt xml:lang="en" xmlns="http://www.w3.org/ns/ttml" ... >
<head>
<layout>
<region xml:id="area1" tts:origin="10% 10%" tts:extent="80% 10%" ... />
</layout>
</head>
<body>

.....
<p begin="00:07:30.300" end="00:07:33.000" ... >My life is a movie and</p>
<p begin="00:07:33.600" end="00:07:36.440" ... >everyone's watching.</p>
<p begin="00:07:37.040" end="00:07:41.520" ... >So let's get to the good part</p>
.....
<p begin="01:31:28.000" end="01:31:32.160" ... >It is hard to do the right thing.</p>
<p begin="01:31:32.760" end="01:31:35.500" ... >It is like they want to be perfect.</p>
<p begin="01:31:36.100" end="01:31:39.200" ... >When they don't know</p>
.....

</body>
</tt>
TTML document chunk #i
<?xml version="1.0" encoding="UTF-8"?>
<tt xml:lang="en" xmlns="http://www.w3.org/ns/ttml" ... >
<head>
...
</head>
<body>

<p begin="00:07:30.300" end="00:07:33.000" ... >
My life is a movie and</p>
<p begin="00:07:33.600" end="00:07:36.440" ... >
everyone's watching.</p>

</body>
</tt>
TTML document chunk #i+1
<?xml version="1.0" encoding="UTF-8"?>
<tt xml:lang="en" xmlns="http://www.w3.org/ns/ttml" ... >
<head>
...
</head>
<body>

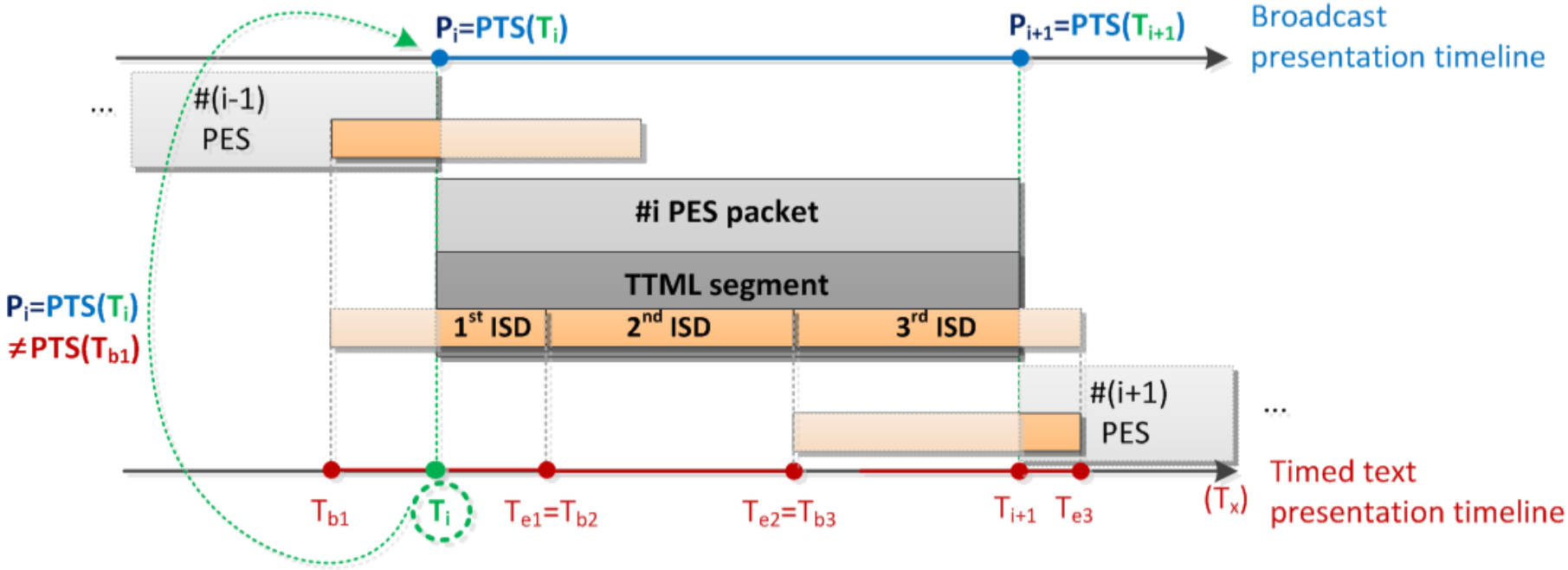
<p begin="00:07:33.600" end="00:07:36.440" ... >
everyone's watching.</p>
<p begin="00:07:37.040" end="00:07:41.520" ... >
So let's get to the good part</p>

</body>
</tt>

TTML segmentation

- To meet the requirements for random access and live subtitling a TTML document needs to be split into standalone short duration documents
- A PES packet becomes active at it’s presentation time (defined by PTS) and is active for a maximum of 5 seconds, or until next PES packet becomes active.
- A subtitle that is “bisected” by the segmentation process appears in multiple document chunks.

Example: PES packet



TTML synchronisation

- Each PES packet carries a single standalone TTML document to aid rapid acquisition.
- PES header has PTS to synchronise with video/audio in the transport stream.
- TTML document has is own timeline, which doesn’t need to be modified.
- The PES also contains a TTML mediatime that corresponds to the PTS value in the PES header. This can be used to convert between times in TTML document and MPEG system clock.

Downloadable font functionality

- To better support different languages an optional mechanism is specified for IRDs to acquire and download required fonts from the internet and/or a broadcast DSM-CC carousel.
- A broadcast table provides information on the fonts available and their location. The table can be shared by several services.
- Signalling can indicate if a downloadable font is required to render the subtitle stream, called an “essential font”.

Colour and HDR

- TTML uses the sRGB colour-space
- sRGB similar to the ITU-R Rec.709 colour space used for HD services.
- Compositing subtitles with HDR video needs some care.
- DVB specification provides guidance for compositing.
- TTML2 provide a informative annex on how to map sRGB values for compositing with HDR video for display on a HDR display.

Webinar on DVB Subtitling Systems

Part 4

Deployment considerations and comparison of new specifications

Recap

- Two new specifications for UHD
  - DVB bitmap subtitles (Bluebook A009)
  - DVB TTML subtitles (Bluebook A174)
- Alternative
  - DVB Teletext
    - Limited character set
    - Limited colours
    - Limited presentation options

Which new spec should I use for UHD subtitles?

- It depends!
- Depends on current use of subtitles (broadcast & IP)
- Depends on future subtitles needs (broadcast & IP)
- Depends on receivers in target market
- New or updated receivers are required to use either of these new specifications.

Why use DVB Bitmap subtitles (EN 300 743)

- Reasons to use new version:
  - Similar to existing SD/HD bitmap subtitles which may already be use.
  - Gives broadcasters more control over look of subtitles, however as with all content, the resultant quality is somewhat dependent of the quality of receiver scaling and compositing.
- Downsides:
  - Higher bitrate required, and it increases with resolution.
  - However the newly added PNG compression should help.
  - Even though UHD subtitles are limited to HD resolution bitmaps, the bitrate will likely be higher than HD subtitles if you use the recommended 8-bit CLUT and encoding.

DVB TTML subtitles

- Reasons to consider bluebook A174 (EN 303 560):
  - Similarity with DVB DASH subtitles, as supported by HbbTV.
  - TTML used for broadcast can be reused later for DASH/CMAF for VoD.
  - More future proof
  - Ideal for new greenfield deployments
  - Bitrate required is independent of resolution.
- Downsides:
  - If already using bitmapped subtitles, it requires a change of process.
    - However if a broadcaster needs TTML for DASH/CMAF delivery, they could migrate to producing subtitles in TTML and converting them for legacy platforms; this could improve legacy bitmap subtitles, if currently being generated from Teletext, with its limitations.


DVB