

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

**ISO/IEC JTC1/SC29/WG11 MPEG2013/M30301  
 July 2013, Vienna (AT)**

**Source**    **Telecom ParisTech**  
**Status**   **For consideration at the 105<sup>nd</sup> MPEG meeting**  
**Title**     **Report in WebVTT in ISOBMFF Implementation**  
**Author**   **Cyril Concolato**

## 1 Introduction

During the 104<sup>th</sup> MPEG meeting, study texts of ISO 14496-30 DIS and of 14496-12 DAM2 have been produced. They enable the carriage of WebVTT streams in MP4 files. This contribution reports on the implementation of these tools in the GPAC project.

## 2 WebVTT import into MP4

As reported during the previous meeting in m29230, GPAC's MP4Box tool supports typical operations on MP4 files. The summary of the possible operations are given here:

Operation	Command line example	Notes
WebVTT Import	<code>MP4Box -add file.vtt file.mp4</code>	Adds a track to the given file: <ul style="list-style-type: none"> <li>Overlapping cues are split into non-overlapping cues and stored in samples</li> <li>Comments and in-between cues text are ignored (support for <code>VTTAdditionalTextBox</code> to be added)</li> </ul>
	<code>MP4Box -add file.vtt:lang=en:layout=800x600x100x100 file.mp4</code>	Sets the language and track layout information
	<code>MP4Box -add file.vtt:delay=1000 file.mp4</code>	Adds an edit list
	<code>MP4Box -add file.vtt:dur=20 file.mp4</code>	Import only a certain duration of the input file
SRT Import as WebVTT track	<code>MP4Box -add file.srt:FMT=VTT file.mp4</code>	Forces the import of SRT files to generate a WebVTT track instead of a 3GPP Timed Text track

WebVTT Export	MP4Box -raw <trackId> file.mp4	Exports a WebVTT file where consecutive cues from different samples and with the same id, settings and payload are merged
	MP4Box -raw <trackId>:vttnomerge file.mp4	Dump samples as-is w/o merge, and with empty samples
	MP4Box -raws <trackId>:<sampleNumber> file.mp4	Exports only one VTT sample
File Editing	MP4Box -split-chunk start:end file.mp4	Split a track into 2 files
	MP4Box -cat file1.mp4 -cat file2.mp4 file3.mp4	Concatenation of tracks (second header is ignored)
DASH	MP4Box -frag 1000 file.mp4	
	MP4Box -dash 1000 file.mp4	Fragments and segments the file and produces an MPD containing: <Representation id="1" mimeType="video/mp4" codecs="wvtt" width="800" height="600" startWithSAP="1" bandwidth="1939">

### 3 WebVTT in MP4 playback

For this meeting, we have implemented support for WebVTT playback within GPAC's players. Currently, only WebVTT content stored in MP4 files is supported. Raw WebVTT files are not played. One of the reasons for that is that the storage of WebVTT cues in MP4 files ensures that each sample is a RAP, therefore at each sample previous cues can be safely discarded.

In details the rendering of WebVTT cues is performed as follows. Upon detection of a WebVTT stream in an MP4 file, a dedicated decoding pipeline is created (by the vtt\_in module) and a graphical layer is added on top of the video layer, if a video is also present in the MP4 file. The decoder is in charge of receiving the WebVTT samples, of parsing box-structured cues and of calling the graphical layer for rendering. The graphical layer is in charge of rendering WebVTT Cues. We have chosen to base the rendering upon the existing SVG rendering capabilities of GPAC. So, in real time, WebVTT cues are transformed into SVG elements, namely SVG Tiny 1.2 textArea elements. In fact, upon reception of a new WebVTT sample, for each cue, the decoder calls a JavaScript function to trigger the rendering of that cue. Due to the JavaScript approach, such rendering is highly configurable and can be modified without recompiling GPAC.

So far we have added support for:

- Top/Bottom-based line positioning
- Line wrapping and multiline cues
- Multiple simultaneous cues
- Basic rendering: italic, bold, underlined

Support for regions has not been added.

With this player we have been able to validate the playback of WebVTT cues in MP4 files in the following conditions:

- Stored in MP4 files and read from disk or HTTP;
- Stored in fragmented MP4 files and read as part of a DASH presentation;
- With a video or without a video

We have experimented seeking into the file, and have not encountered the need for the `WebVTTcueSourceIdBox`, as indicated in the FRNB comments on 14496-30.

Since GPAC's player is available on many platform, WebVTT playback is available on Windows, Mac, Linux, iOS, Android.

## **4 Conclusion**

This contribution reported the added support for playback of WebVTT files in GPAC, including in a DASH context. Comments and feedback from the MPEG community is appreciated.