5, rue de la Verrerie
38120 Le Fontanil (France)
Tel : + 33 476 26 1976 - Fax : + 33 476 26 2593
www.tcube.tv

Le Fontanil, Sept. 12th 2013

Subtitling from teletext, files or live

By Jean Luc Grimaldi, Tcube Marketing Manager

Introduction

Requirement for subtitling have grown very quickly and subtitling services are already or will become very soon mandatory in many countries.

This demand comes with request for better services, higher number of simultaneous languages, higher graphics quality and ease of deployment and maintenance.

At Tcube, we have seen the market evolving from old Teletext services in analogue to insertion of Teletext data and Subtitle file content in HD-SDI (using OP47 or SMPTE2031 containers), then direct creation of DVB Subtitle graphics for transmission within a broadcasting service.

Another request coming up is the subtitling of live events from an operator (typist). In this domain, use of speech recognition and automatic translation system open new perspectives for cheap and efficient subtitling systems.

The purpose of this document is to provide an explanatory survey of the different requirements and possible technical solutions.

Teletext legacy

The teletext system was invented in the late 70s and spread quickly from the early 80s.

At that time, the TV screen was divided into 24 rows of 40 characters each. Information was organized in pages and magazines. Originally based on a latin character set with additional pages for each language specificity (German, French, English, Nordic, …), system also accepted Cyrillic and Greek character set. Arabic and Hebrew were also supported.

In fact, the receiver (TV Set) was doing the job of creating the page from the transmitted data of 24 rows times 40 characters. A single font (non-proportional) was implemented.

Subtitling from a Teletext service was obtained just using using only one or two rows (typically 22 and 23) from a specific page and magazine (888 in France, 150 in Germany…).

Teletext was transmitted in the vertical blanking lines using analogue modulation carrying one row per line.

Despite new specifications supporting semi graphics characters and higher quality standard, nearly all Teletext receivers stopped at text level specification.

Terrestrial or Satellite Digital transmission drove Teletext gently to the end. This accelerated with HD transmission where quality is definitely found too poor.

Due to the long period of use and high volume of video processed, there exists a very large base of Teletext data stored on the vertical blanking of archived tapes.

STL files

In the early 90s, European Broadcasting Union normalized a file format for exchange of data subtitles. It is a binary format containing a descriptor and a succession of subtitle blocks, each containing In Timecode, Out Timecode, Row for display and Subtitle Text.

5, rue de la Verrerie
38120 Le Fontanil (France)
Tel : + 33 476 26 1976 - Fax : + 33 476 26 2593
www.tcube.tv

The descriptor field contains a set of characteristics, especially max number of characters per row and max number of row per field. Though this was a first step to get out of the Teletext strict format 40x24, the set was not complete and made assumptions about the font used (typically a non proportional). Font size is not defined in file; inappropriate font size versus max number of rows (e.g.) may result in either overlapping or too distant rows.

These defaults rose even more when HD came on the market.

STL has been a progress in terms of dematerialization, has proved to be efficient for text level insertion within transmission. However the unspecified characteristics make the rendering somehow unpredictable as it depends on local settings of receiver.

Other File formats

In the vein of STL, many formats appear to carry subtitle like SRT, used for DVD. Nearly all of them leave open font name and font size selection.

Specifying a modern subtitle description format

As subtitling becomes more used and as quality matters (especially when Subs are burned before transmission), a strong need for a fully described subtitle file has emerged.

Such a file should also have to:

- support all fonts (Latin, Greek, Cyrillic, Arabic, Hebrew, Chinese, Vietnamese, Thai, …)
- be resolution independent
- be editable in plain text
- be rendering system independent
- be upwards-compatible (that means new features not supported by a former system must not block the functioning)

One solution comes from the DLP Cinema projection, CineCanvas.

This is a XML file format (so guaranteeing by construction the text editable feature and the upwards compatibility – each feature comes within < - >, unknown feature are easy to ignore).

It includes the Load Font feature, making the system completely open to any font (especially Unicode), as long as you can reach font's location in the network.

It is resolution independent as all parameters for text location are given in percentages of width or height screen.

Font size is also given in percentage of screen height as explained hereunder.

The Cinecanvas document specifies that *'font is given in points. Fonts are rendered as if the screen height is 11 inches, so a 72pt would be 1/11 of the screen height.'*

To be clear, subtitles are always rendered as a percentage of screen height using the relation

Subtitle height = Screen Height * Font size /(72*11)

So it is resolution independent.

As a conclusion, Cinecanvas or precisely Subtitle specification (XML file format) for DLP Cinema Projection Technology, has proved to be a very robust, fully described and operational file description for subtitles.

5, rue de la Verrerie
38120 Le Fontanil (France)
Tel : + 33 476 26 1976 - Fax : + 33 476 26 2593
www.tcube.tv

Live subtitiling

Open subtitling of live events is a niche market that requires versatility and fast reaction.

Except NEWFOR protocol which is a private one, no public protocol exists.

Tcube decided to derive from Cinecanvas a subset of commands to easily send controls for font selection, font size, text position and text itself.

Labelled by Tcube as LIVE, the subset is transmitted over IP. It has been already successfully implemented by third parties worldwide.

Tcube offer

Tcube is a well known actor in the subtitling domain and ancillary data handling.

Capitalizing from successful development of data bridges and protocol translation (OP47, SMPTE2031) devices in the past years, Tcube moved to graphics subtitle rendering units either for open subtitles burned on picture or for DVB Subtitling generation (generation of complementary subtitles streams within TS Streaming).

Based on PEACH, Tcube proprietary hardware device supporting SDI and ASI I/O with Windows inside, subtitle applications can take either LIVE, Teletext or files as subtitle source.

VB1xxx product references address the open subtitle segment.

VB2xxx product references address the DVB graphics generation (generation of complementary subtitles streams –up to 8- within TS Streaming).

Easy to operate, easy to control, automation friendly PEACH make these small units (1/2 RU, 40W power consumption) the perfect fit for upgrading channels at lower cost without sacrificing on performances.

Deliberately open to existing standard and practices, PEACH currently supports the following file formats: STL, SRT, PAC, XML (cinecanvas).

For more information, please visit http://www.tcube.tv