Synchronization and Customization of Subtitles in Web-based Main-Screen and Multi-Screen Scenarios

Zuzanna Parcheta, Mario Montagud, Jordi Belda, Fernando Boronat
Universitat Politècnica de València
Grau de Gandia (SPAIN)
{zupar@epsg., mamontor@, jorbelva@epsg., fboronat@dcom.}upv.es

ABSTRACT
This paper presents a web-based platform that enables the customization and synchronization of subtitles when using both a shared main screen and multiple secondary screens. This platform enables the dynamic adaptation of subtitles’ format (font family, size, color...) and position according to the users’ preferences and/or needs. Likewise, the users can dynamically select the subtitles’ language and adjust the number of lines to be displayed. Moreover, a delay offset to the subtitles can be applied. This platform can also be used in multi-screen scenarios to enable personalized media experiences. By using a secondary device, each user can enable the consumption of subtitles in two different languages and navigate between them, being able to set the position of the video by clicking on a specific subtitle line. Synchronization between subtitles and other media is achieved thanks to standardized HTML5 elements and Javascript. Synchronization across devices is achieved by using Javascript components (Node.js and Socket.IO) and a designed virtual clock synchronization mechanism.

Author Keywords
IDES, Multi-Screen, Subtitles, Synchronization, Web.

ACM Classification Keywords
C.2.4 [Communication Networks]: Distributed Systems.

INTRODUCTION
Subtitles and other types of text-based media (e.g., captions, descriptions...) are key components to provide enriched and personalized media services. Additionally, for many people, such as hearing impaired or non-natives, text-based media information is the only way to access and interpret audiovisual content. This paper presents a web-based platform that enables the synchronization (sync, hereafter) and customization of subtitles when using both a shared main screen and multiple secondary screens (see Fig.1). A key feature of this platform is the exclusive reliance on standard web-based technologies, such as HTML5 and Javascript. This guarantees cross-platform and cross-device support, as well as ubiquitous deployment. Our goal was to constitute a flexible and accurate platform to enable the dynamic adaptation and customization of subtitles according to the users’ needs (e.g., language), preferences (e.g., format), capabilities (e.g., hearing or visual impairments), the available devices (e.g., smartphone, tablet, laptop, TV...), application dynamics and contextual conditions, while guaranteeing synchronized playout.

Other solutions for synchronizing subtitles with audiovisual content in web browsers have been proposed in the last years (e.g., [1] and [2]). The system presented in [1] enables the customization, adaptation and sync of subtitles across devices (which is a use case of Inter-Device Sync or IDES). It takes advantage of the native support for Scalable Vector Graphics (SVG) for presenting the subtitles, using the obsolete SubRip (.srt) format, and of Synchronized Multimedia Integration Language (SMIL) to achieve sync. Our platform differs from that one because: i) it only uses standardized HTML5 elements, which enables higher flexibility and performance; ii) it uses the WebVTT (Web Video Text Tracks) subtitles’ format (.vtt), which is the current W3C standard for displaying timed text-based media; iii) it provides accurate sync levels, by including timestamps and a virtual clock sync mechanism; iv) it allows the concurrent selection of different subtitles’ languages; and v) it enables the navigation between subtitles. In [2], WebVTT subtitles are synchronized with DASH (Dynamic Adaptive Streaming over HTTP) content in a live scenario, but no customization, navigation and IDES functionalities are provided.

Our platform can be very useful in many use cases. On the one hand, it can be beneficial for language therapy, literacy and learning purposes. On the other hand, it can be a relevant tool for social integration. First, it can be used by people with audiovisual impairments, who can adapt the subtitles’ format according to their needs. Second, it can be used in public (e.g., transport systems, airports...) and private spaces (e.g., hotels, theaters...) to allow people from different countries the selection of their native language for subtitles.

Although the design and development of our platform are still in progress, the technological components to achieve the targeted functionalities are available, and presented in this demo paper. The audience will be able to experiment with
our platform, from whom we expect valuable feedback about
design aspects, applicability and future functionalities.

SUBTITLE SYNC PLATFORM
Four main technological components have been used to
develop our platform. The first one is the HTML5 video
element, which allows embedding full-fledged media players
into websites. Besides, the HTML5 track element (child of
the video element) is used to automatically handle subtitles
and other timed text-based data (e.g., metadata, descriptions,
captions, chapters…). The second one is Node.js, which is an
open-source, cross-platform runtime environment, written in
Javascript, for server-side and networking web-based
applications. The third one is Socket.IO, which is a
lightweight Javascript library that enables real-time
communication between web clients and a (Node.js) web
server. The fourth one is a virtual clock sync mechanism to
achieve a global coherent notion of time in the session.

Subtitles Presentation and Customization
A WebVTT subtitle file is composed of various blocks,
called cues. A cue basically contains an ID (optional), time
settings (start and end time), the text-based information
and a blank line (end of the cue). The HTML5 track element is
responsible of presenting each cue during the proper period,
according to the media components being played out.
Moreover, our platform enables the dynamic adjustment of
subtitles’ format (e.g., color, font, size, layout…) and
position (e.g., drag and drop, replacement beyond the video
window or the main screen…) to optimize their readability
or the application’s esthetics.

Language Selection
Each user can select the subtitle language through a drop
down list (which sets the src attribute of the track element
to the proper value). In the secondary screens, each user can
also enable the consumption of the original subtitles, which
will be received from the main device (see Fig.1).

Subtitles Synchronization
The track element is able to automatically handle the sync of
text-based media with the audio and video content.
Moreover, our platform requires IDES. This is achieved by
periodically exchanging messages between the involved
entities, including their current playout times. To achieve a
global and coherent notion of time, we could make use of
Network Time Protocol (NTP), at either the system or
application-level (e.g., via Javascript), but we decided on
developing a virtual clock sync mechanism. It is based on
periodically exchanging timestamped messages between the
involved entities to estimate and compensate the clock skews
between them, by also taking into account the measured the
(two-way) latency, in a similar way as NTP performs.

The users can additionally set a customized delay offset to
the subtitles. This can be useful in certain situations. For
example, users could prefer to consume the subtitles a bit
earlier or later than the associated audio content to check how
words are pronounced or if they are able to identify the
spoken words (or sentences), respectively.

Moreover, our platform allows sharing the execution of the
navigation control or VCR-like commands (e.g., play, pause,
seek…) in a synchronized way.

Subtitles Navigation
When using secondary devices, each user can set the number
of subtitles lines to be displayed. In this case, the user can
also seek to a specific video position by clicking into the
targeted subtitle line.

A demo video showing the capabilities of the presented
platform can be watched at goo.gl/xS9HVX.

FUTURE WORK
Several extensions to the presented platform are planned.
First, we want to include artificial intelligence
functionalities, such as voice synthesis and recognition.
This will enable to use it for live subtitling as well as to provide
the automatic generation of both audio and subtitles for other
languages than the ones available in the media files. Second,
we want to improve the User Interface Design (UID). Third,
we want to explore the suitability of alternatives methods for
sharing a common timeline across the involved entities.

REFERENCES
1. A. Rodriguez, G. Talavera, P. Orero, J. Carrabina,
Subtitle Synchronization across Multiple Screens and Devices, Sensors, 12(7),
8710-8731, June 2012
2. C. Concolato and J. Le Feuvre, Live HTTP streaming
of video and subtitles within a browser, MMSys ’13,
146-150, Oslo (Norway), February 2013.

Figure 1. Subtitles Customization and Sync across Devices.