

The Chroma+ Approach to Enrich Video Content using HTML5

Philipp Oehme, Michael Krug, Fabian Wiedemann, Martin Gaedke
Department of Computer Science
Chemnitz University of Technology
Chemnitz, Germany

{philipp.oehme,michael.krug,fabian.wiedemann,martin.gaedke}@informatik.tu-chemnitz.de

ABSTRACT

The Internet has become an important source for media content. Content types are not limited to text and pictures but also include video and audio. Currently audiovisual media is presented as it is. However, these media do not integrate the huge amount of related information, which is available on the Web. In this paper we present the Chroma+ approach to improve the user experience of media consumption by enriching media content with additional information from various sources in the Web. Our approach focuses on the aggregation and combination of this related information with audiovisual media. This approach involves using new HTML5 technologies and with WebVTT a new annotation format to display relevant information at definite times. Some of the advantages of this approach are the usage of a rich annotation format and extensibility to include heterogeneous information sources.

Categories and Subject Descriptors

D.2.11 [Software Engineering]: Software Architectures – *Domain-specific architectures*; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems; H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia;

Keywords

Media enrichment, media aggregation, streaming, HTML5, mashup, video streaming, video mashup, widgets

1. INTRODUCTION

Recently, the market of media is changing dramatically. The Web becomes a more important way of distributing audiovisual content or text-based information. In addition, the increasing spread of Internet connections with high bandwidth facilitates the delivery of media content, such as video or audio, to end users. We define end users as people consuming media in front of any device capable of displaying media, such as TVs, computers and mobile devices. Simultaneously, the Internet offers a rich source of information to enrich this media content.

More than half the users who are online are dealing with video platforms like YouTube [1, 3]. While these platforms are limited to show only video content, the Chroma+ approach describes

how to enrich media content with widgets containing related information such as text, pictures, videos or maps. We consider video content as the visual part of a video whereas the metadata of the video is the other part. The purpose of this paper is to demonstrate our approach by using Web technologies such as HTML5 which offers the opportunity to stream video easily via HTTP [2].

This paper is structured as follows: We provide an overview of our demonstration. This section will also examine the technical aspects, including the HTML5 features used, the research on WebVTT [5] and our current prototype, which we will present during demonstration sessions. Finally, we offer an outlook to further developments.

2. DEMONSTRATION

The mission of the Chroma+ project is to build a platform that facilitates the time-oriented composition of media streams with HTML components or widgets. One of the current approaches to enrich media with additional information is popcorn.js¹. In contrast to popcorn.js, where all annotations are described in custom JavaScript files, we are using a separate WebVTT file. Therefore, it is highly adjustable and provides support for extensions. Currently, the annotations have to be written manually, but our project partners are working on an automatic annotation framework.

2.1 VIDEO WITH HTML5

The new video element in HTML5 offers the opportunity to display enriched media. While many browsers support the video element itself by now, support for video codecs varies heavily. Therefore, the opportunity of including several <source> elements is important for the accessibility of content. This demo exhibits media enrichment, by means of the <track> element by including explicitly timed annotations. The support of the <track> element is under development. Currently this feature is supported by Chrome 23+ [4], IE 10, Opera 12.10 and Safari 6 [6]. Additionally, Mozilla is working on the support in Firefox. Apart from this, there is no support in mobile browsers.

The HTML code example listed below shows the exemplary use of the <video> and <track> element in our demo. The track element with the attribute “kind” set to “metadata” is not displayed by browsers, but accessible for JavaScript components through the text track API. Thus, it does not collide with regular subtitles. We intend to display the annotations at explicit timestamps. Therefore, we use a JavaScript event provided by the <track> element. The “oncuechange” event is fired when a

Copyright is held by the author/owner(s).
WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

¹ <http://popcornjs.org/>

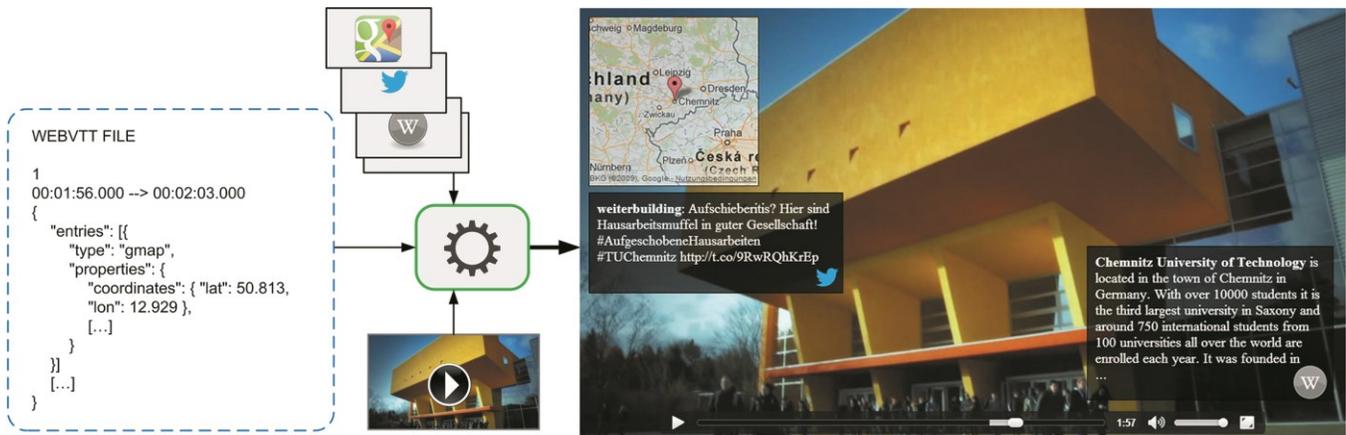


Figure 1. Architecture overview and resulting video mashup

specified cue in the metadata file is entered or left. This enables us to show metadata at the certain specified times.

```
<video id="demo" controls >
  <source src="demovideo.webm" type="video/webm">
  <source src="demovideo.mp4" type="video/mp4">
  <track src="metadata.vtt" label="Metadata"
    kind="metadata" default>
</video>
```

2.2 ANNOTATIONS WITH WEBVTT

The approach to use the HTML5 <video> and <track> element was chosen because as yet there is no support in browsers for a video format which can contain additional metadata in one container. We use an external annotation format combined with supported standard video formats. This offers the possibility to adapt our approach to many scenarios. Furthermore, the use of WebVTT provides the option to use JSON to describe annotations. A sample data structure extracted from the demonstration annotation is illustrated in Figure 1. The flexibility of JSON enables the creator of annotations to properly specify the information and provides the capability to use plugins.

The WebVTT file contains several cues consisting of an optional identifier, an explicit timestamp and the cue payload, which can contain any text content. The timestamp is given in “hh:mm:ss.ttt” format, describing the hours, minutes, seconds and thousandths of a second in prescribed order.

2.3 PRESENTATION OF OUR DEMO

A first implementation of our approach is the core of our demonstration². We are presenting a promotion video from the Chemnitz University of Technology, which is annotated and then combined with several widgets showing related information. Based on jQuery we implemented a script, which uses the “oncuechange” event to synchronize the video with the annotations from the metadata file. The specified information is processed and displayed as a widget in a predefined way. The position of the widgets must not be on top of the video. They can also be placed next to it. In further developments end users shall be able to move the widgets. The demonstration uses several Web service APIs to enrich media. One example service is Google Maps, which is used to search for a specified location and display this location on a

map. Another service used is Twitter. Twitter enables us to search for entries (Tweets) containing a hashtag defined by the video annotations. Additionally, we included information from Wikipedia to provide the users with an advanced possibility to inform themselves about the topic. Furthermore, a connector for Google+ was implemented, which facilitates the opportunity to create a relation between individuals seen in the video and their social profiles. Figure 1 shows the architecture of our implementation and the result of media enrichment.

3. LESSONS LEARNED AND OUTLOOK

Our first experiments show that the composition of video and related Web content using HTML5 techniques works. First user interviews for gathering requirements for our new platform show high interest in the topic. The combination of WebVTT and widgets is challenging, as the loading behavior of some widgets has to be taken into account. We are currently working on a video mashup platform that allows users to specify the video streams and corresponding Web content. This results in an enriched media stream improving the user experience in media consumption.

4. ACKNOWLEDGMENTS

This work was supported by the Sächsische Aufbaubank within the European Social Fund in the Free State of Saxony, Germany (Project Chrooma+).

5. REFERENCES

- [1] Google Inc. 2012. Make the Web Work for You. (2012).
- [2] HTML - Living Standard: 2013. <http://www.whatwg.org/specs/web-apps/current-work/>. Accessed: 2013-02-22.
- [3] Statista Inc. 2012. Online video market in the U.S. - Statista Dossier 2012.
- [4] Web Platform Status (chromestatus.com) - The Chromium Projects: 2013. <http://www.chromium.org/developers/web-platform-status>. Accessed: 2013-02-22.
- [5] WebVTT Standard: 2013. <http://dev.w3.org/html5/webvtt/>. Accessed: 2013-02-22.
- [6] WebVTT support in browsers | Web Media Text Tracks Community Group: 2012. <http://www.w3.org/community/texttracks/2012/08/23/webvtt-support-in-browsers/>. Accessed: 2013-02-22.

² <http://chroomaplus.eu/www2013-demonstrator>