Are Sensory Effects Ready for the World Wide Web?

Christian Timmerer, Markus Waltl, and Hermann Hellwagner

Klagenfurt University, Multimedia Communication Group, Universitätsstrasse 65-67, 9020 Klagenfurt, Austria
{*firstname.lastname*}@itec.uni-klu.ac.at

Abstract. The World Wide Web (WWW) is one of the main entry points to access and consume Internet content in various forms. In particular, the Web browser is used to access different types of media (i.e., text, image, audio, and video) and on some platforms is the only way to access the vast amount of information on the Web. Recently, it has been proposed to stimulate also other senses than vision or audition while consuming multimedia content through so-called sensory effects, with the aim to increase the user's Quality of Experience (QoE). The effects are represented as Sensory Effects Metadata (SEM) which is associated to traditional multimedia content and is rendered (synchronized with the media) on sensory devices like fans, vibration chairs, lamps, etc. In this paper we provide a principal investigation of whether the sensory effects are ready for the WWW and, in anticipation of the result, we propose how to embed sensory effect metadata within Web content and the synchronized rendering thereof.

Keywords: MPEG-V, Sensory Effects, Quality of Experience, World Wide Web.

1 Introduction

In recent publications, it was proposed [1][2] to annotate multimedia content (i.e., videos) with a so-called Sensory Effect Metadata (SEM) description that can be used to stimulate also other senses than vision or audition, e.g., olfaction, mechanoreception, or thermoception which shall lead to an enhanced, unique user experience. That is, the actual video content is rendered on traditional devices (e.g., TV screens) and the sensory effects (i.e., described via SEM and synchronized with the video) are rendered on sensory devices like fans, vibration chairs, lamps, perfumer, etc. The concept of this approach is depicted in Figure 1 with the aim to provide an enhanced, unique, and worthwhile user experience. Note that the SEM is currently being standardized within Part 3 of MPEG-V Media Context and Control [3].

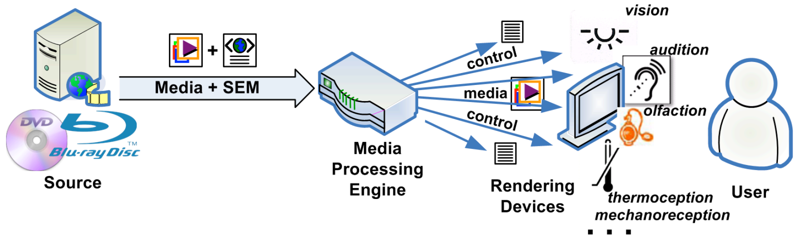

Figure 1. The Concept of Sensory Effects [1].

In [4] we present promising results regarding the usage of sensory effects in conjunction with video content. However, current implementations primarily focus on video content. In this paper we argue that this could be easily extended to Web content, of course including videos also. Therefore, in Section 2 we propose an approach on how to embed SEM descriptions into Web content (i.e., (X)HTML) and the processing thereof as a browser plug-in based on an MPEG Extensible Middleware (MXM) [5] engine capable of parsing the SEM description. To the best of our knowledge, this is the first attempt trying to bring sensory effects to the World Wide Web. The conclusions and future work are outlined in Section 3.

2 Proposed Approach for Integrating Sensory Effects within Web Content

The preferred way to link resources such as a SEM description to a Web document – herein after referred to as document – is the `link` element. An example usage of the `link` element to reference a SEM description is shown in Listing 1.

Listing 1. Example usage of the `link` element to reference a SEM description.

```
<link rel="alternate" type="text/xml" title="SEM description"
      href="http://server.org/sem.xml" media="light, wind, vibration"/>
```

The attributes `type`, `title`, and `href` are used as specified by HTML. The `rel` attribute indicates that the SEM description provides an *alternate representation of the current document* which, in our view, conforms with the concept of sensory effects as depicted above. In particular, if used in conjunction with the `media` attribute, the `ref` attribute indicates that the referenced SEM description is intended for use with the media specified. Please note that the media attribute must support some version of the media query language [6]. Currently, the following media types are defined: ‘aural’, ‘braille’, ‘handheld’, ‘print’, ‘projection’, ‘screen’, ‘tty’, ‘tv’. As none of them match with the concept of sensory effects we propose adding new media types corresponding to the sensory effects as defined in [3], e.g., ‘light, wind, vibration’ as shown in Listing 1.

Based on the information in the `link` element, the SEM description could be parsed either directly via JavaScript or within a browser plug-in utilizing an MXM engine as proposed in [5]. The SEM MXM Engine – part of the MPEG-V reference software – is used to access and parse the data within the SEM description and used to control the actual rendering devices (e.g., fans, vibration chairs, lamps). Therefore, the corresponding application programming interfaces (APIs) of the rendering devices must be accessed from within the Web application/browser. In case of the JavaScript solution, this might be an issue but for the browser plug-in the rendering devices could be accessed directly from within the plug-in.

However, the main issue is the synchronization with possible multimedia resources included within the document such as video or audio content. Therefore, it is required to access the current playback position of the actual multimedia resources which is provided thanks to HTML5's `video` and `audio` elements and corresponding DOM attributes [7]. In particular, the `currentTime` DOM attribute provides the current playback position in seconds that can be used for synchronization with the sensory effects. Additionally, the `timeupdate` event indicates that the current playback position changed as part of normal playback which also serves the synchronization purpose. Listing 2 shows an excerpt of how to use the `timeupdate` event and `currentTime` attribute for the synchronization of the video content with the sensory effects.

Listing 2. Usage of `timeupdate` and `currentTime` for synchronization of video content with sensory effects.

```
<script>
...
function time_update() {
  var vid = document.querySelector('video');
  var now = vid.currentTime;
  // note: sem holds sensory effects information previously parsed
  for (var i=0; i<sem.length; i++ {
    if (now >= sem[i].start && now <= sem[i].end) {
       update_sensory_devices(sem[i]);
       break;
} } } // end function time_update()
...
</script>
...
<video src="http://server.org/video.mp4" ontimeupdate="time_update()"
       autobuffer controls>
```

Finally, in [1] we presented means for the automatic extraction of color information from the video frames in order to control additional light sources, e.g., placed behind or next to the TV screen. In particular, we evaluated a choice of algorithms within different color spaces (average color: RGB; dominant color: RGB, HSV, HMMD). However, in all cases we need to access the content of the frame which is not possible within HTML5 as something like `currentFrame` does not exist and is not foreseen. On the other hand, the `canvas` element allows for manipulating the display's content (examples can be found in [8]) and can be used for this purpose. In particular, the `canvas` element is used to extract the pixel information from the display and, consequently, this information is used for extracting the color information. For the actual color extraction and its usage the interested reader is referred to [1].

3 Conclusions and Future Work

In this paper we proposed an approach that brings sensory effects to the World Wide Web. In particular, it provides a conceptual answer to the question raised in the title of this paper. That is, yes, sensory effects are ready for the World Wide Web but this approach requires some implementation work to be done as part of our future work. Furthermore, currently it is not clear how these sensory effects are perceived by the users of the Web content which calls for formal subjective quality assessments. Finally, we would like to study how to apply sensory effects for the whole Web site/content instead of only focusing on audio and video parts.

Acknowledgments. This work was supported in part by the European Commission in the context of the NoE INTERMEDIA (NoE 038419), the P2P-Next project (FP7-ICT-216217), and the ALICANTE project (FP7-ICT-248652).

References

1. Waltl, M., Timmerer, C., Hellwagner, H.: A Test-Bed for Quality of Multimedia Experience Evaluation of Sensory Effects, In: 1st Int'l. Workshop on Quality of Multimedia Experience (QoMEX'09), San Diego, USA (2009)
2. Choi, B. S., Joo, S., Lee, H.: Sensory Effect Metadata for SMMD Media Service, In: 4th Int'l Conf. on Internet and Web Applications and Services (ICIW'09), pp. 649-654, Venice/Mestre, Italy (2009)
3. ISO/IEC 25005-3 FCD: Information technology – Media context and control – Sensory Information, ISO/IEC JTC 1/SC 29/WG 11/N10987, Xi'an, China (2009)
4. Waltl, M., Timmerer, C., Hellwagner, H.: Increasing the User Experience of Multimedia Presentations with Sensory Effects, In: 11th Int'l. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'10), Desenzano del Garda, Italy (2010)
5. Timmerer, C., Chiariglione, F., Preda, M., Rodriguez-Doncel, V.: Accelerating Media Business Developments with the MPEG Extensible Middleware, In: FIA Valencia Book 2010, IOS Press (2010)
6. Lie, H., Çelik, T., Glazman, D., van Kesteren, A. (eds): Media Queries, Candidate Recommendation, W3C (2009), http://www.w3.org/TR/2009/CR-css3-mediaqueries-20090915/
7. Hickson, I., Hyatt, D.: HTML5, Working Draft, W3C (2009), http://www.w3.org/TR/2009/WD-html5-20090825/
8. Manipulating video using canvas, https://developer.mozilla.org/En/Manipulating_video_using_canvas