

A Toolset for the Authoring, Simulation, and Rendering of Sensory Experiences

Markus Waltl, Benjamin Rainer, Christian Timmerer, Hermann Hellwagner

Alpen-Adria-Universität Klagenfurt, Multimedia Communication

Universitätsstraße 65-67

9020 Klagenfurt am Wörthersee, Austria

+43 (0)463 2700 3600

firstname.lastname@itec.aau.at

ABSTRACT

This paper describes a toolset for the authoring, simulating, and rendering of multimedia content annotated with Sensory Effect Metadata (SEM) descriptions as specified in Part 3 of the MPEG-V standard. This part of MPEG-V standardizes the description of sensory effects (e.g., light, wind) in order to be rendered on sensory devices (e.g., fans, vibration chairs) aiming at generating a sensory experience stimulating possibly all human senses. Our implementation comprises a toolset to author sensory effects associated with multimedia content and the simulation thereof. Furthermore, it includes a library, a standalone player, and a Web browser plug-in which enables the playback and rendering of sensory effects on off-the-shelf rendering devices and in various contexts. All software modules are available under the GNU General Public License (GPL) v3 and the GNU Lesser General Public License (LGPL) v3 respectively.

Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities

General Terms

Human Factors, Experimentation, Standardization.

Keywords

MPEG-V, Annotation Tool, Simulator, Media Player, Web Browser Plug-in, Sensory Effects, Sensory Experience.

1. INTRODUCTION

Multimedia content such as combinations of video, audio, text, and images is provided and consumed everywhere and during all times of the day. The focus of the currently available content is clearly on stimulating senses like audition and vision. Some cinemas, like in the Universal Studios theme parks [16], already provide an enhanced viewing experience by stimulating further senses like olfaction, mechanoreception, or thermoception. In recent years also devices for the home environments came into the market (e.g., [1][4][12]) and standards like MPEG-V were ratified including means for describing sensory effects (e.g., wind, vibration, light) to be attached to traditional multimedia [9]. A

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'12, October 29–November 2, 2012, Nara, Japan.

Copyright 2012 ACM 978-1-4503-1089-5/12/10...\$15.00.

brief introduction to the MPEG-V standard focusing on Part 3 is given in Section 2. In particular, there are a number of companies providing sensory devices such as perfumers [5][12][13], devices for rendering light effects [4], or a combination of different effects (i.e., light, wind, vibration) [1]. Furthermore, the research community started to contribute to this emerging area, e.g., with the evaluation of olfaction for video content [7] or the influence of light effects on 3DTV [15]. Hence, there is a need for tools supporting this area.

In this paper we describe our open source toolset – complementing our dataset [19] – for the authoring, simulation, and rendering of sensory experiences (cf. Section 3). For the authoring stage, we provide the Sensory Effect Video Annotation (SEVino) tool which is described in Section 3.1 whereas the Sensory Effect Simulator (SESim) is a tool for simulating sensory effects synchronized with the actual media data and described in Section 3.2). For the playback and rendering we provide a library plus Web browser plug-in and a standalone player described in Sections 3.3 and 3.4 respectively. To the best of our knowledge, this is the first complete toolset for authoring and rendering sensory experiences. There are other annotation tools (e.g., [3]) for sensory experiences but they are – as far as we know – not open source and do not provide tools for rendering the sensory experience. Section 4 describes a possible deployment of the introduced toolset for user studies conducted in this area including results thereof. Finally, we conclude the paper in Section 5 presenting also future work items.

2. MPEG-V: Media Context and Control

The MPEG-V standard, formally referred to as Media Context and Control [9], is an international standard that provides tools for describing virtual and real worlds. The main focus of this standard is the communication and representation from virtual world objects to real world objects and vice versa. Also the communication between virtual worlds is defined.

MPEG-V consists of seven parts: architecture (Part 1), control information (Part 2), sensory information (Part 3), virtual world object characteristics (Parts 4), data formats for interaction devices (Part 5), common types and tools (Part 6), and reference software and conformance (Part 7).

In this paper we focus on Part 3 (sensory information). Sensory information describes sensory effects (e.g., wind, vibration, scent, light). Therefore, the standard defines the Sensory Effect Description Language (SEDL) which is an XML Schema-based language. SEDL describes building blocks and common attributes for all effects such as the effect type, intensity, duration, location, timestamp, fading, etc. The actual effects are defined in the

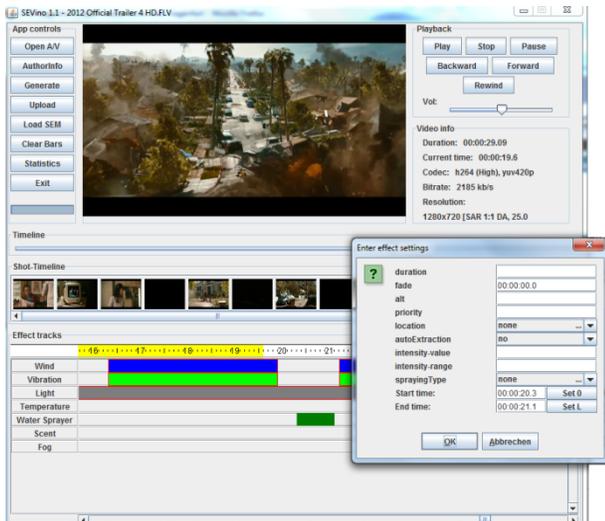


Figure 1. Sensory Effect Video Annotation.

Sensory Effect Vocabulary (SEV). SEV specifies each effect in detail with specific attributes. For example, light effects can have a color or olfactory effects can have a scent.

A description compliant to SEDL and SEV is referred to as Sensory Effect Metadata (SEM) and consists of a number of effects. Such SEM descriptions can accompany multimedia content for enhancing the viewing experience.

3. IMPLEMENTATION

This section provides information about the implementations of our annotation tool, simulator, Web browser plug-in, and media player. All tools are available under GNU General Public License (GPL) v3 and GNU Lesser General Public License (LGPL) v3 respectively. Downloads for each tool can be found under [14].

3.1 Sensory Effect Video Annotation

The Sensory Effect Video Annotation (SEVino) tool provides means for annotating multimedia content with various sensory effects (e.g., wind, vibration, light, scent). SEVino is written in Java and uses VLCJ [10] for the rendering of the multimedia content.

Figure 1 shows a screenshot of SEVino. In the upper part, one can see the video, playback control, and video information. In the center, navigation controls can be found. One can navigate through the video by using a slider-based or shot-based navigation. SEVino uses FFmpeg [6] for extracting frames from the video. FFmpeg provides extracted frames to a shot-detection module which extracts shots using a color-based detection algorithm. In the lower part, the sensory effect editor is located. Here one can enter an effect by dragging the mouse from a start time to an end time like in a Gantt diagram. Subsequently, a window is displayed where one can enter effect parameters such as intensity, duration, or effect-dependent parameters (e.g., scent, color). Each effect type is represented by its own timeline. The different effect types and its parameters are described within a configuration file which is evaluated when starting the application. Therefore, arbitrary sensory effects can be easily integrated by simply editing the configuration file.

From the entered sensory effects SEVino generates an MPEG-V-compliant SEM description. Note that sensory effects that are not defined within the SEV will be put into a separate namespace. Currently, SEVino supports only the namespaces specified in

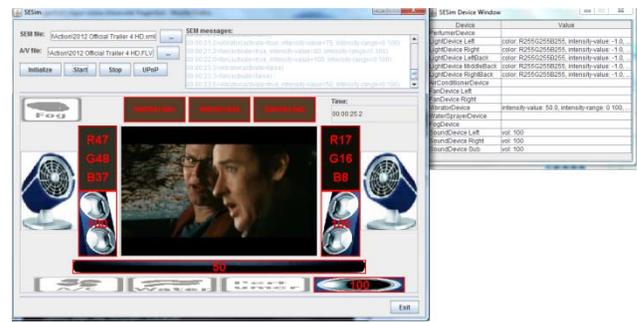


Figure 2. Sensory Effect Simulator.

SEDL/SEV but SEVino is still able to generate a compliant SEM description.

3.2 Sensory Effect Simulator

The Sensory Effect Simulator (SESim) is a tool for simulating SEM descriptions including various sensory effects, i.e., light, wind, vibration, fog, scent, temperature, and water-sprayer. The main functionality of the simulator is for testing generated SEM descriptions without the need of having the actual rendering devices on-site.

SESim is written in Java and uses VLCJ [10] for rendering the multimedia content. Figure 2 illustrates a screenshot of SESim. One can see the video in the center and around it the devices for simulating sensory effects. The setup of the lights, fans, and vibration panel is based on the ambX system [1]. In the top-right corner of SESim the SEM messages provide debugging information about the currently simulated sensory effects. The SESim Device Window as shown in the top-right corner of Figure 2 provides also information about the currently simulated sensory effects but in a separate window.

In addition to parsing and interpreting SEM descriptions, SESim supports an automatic average color extraction. If the SEM description indicates that automatic color extraction is allowed, SESim extracts every 0.1 second a frame and splits it vertically into three equal parts. In particular, for each part the average color is calculated and rendered on the corresponding light. That is, the left part of the frame is rendered on the left light, the center part on the lights above the video, and the right part on the right light.

3.3 Ambient Library and Browser Plug-in

The Ambient Library (AmbientLib) provides a library that is capable of parsing SEM descriptions and rendering the defined sensory effects onto real devices. Therefore, AmbientLib provides the possibility to extend the range of the supported devices through so-called device drivers. The AmbientLib is responsible for parsing SEM descriptions and for calculating the average color of a video frame in real time. Therefore, the video frames are partitioned into 9 pieces, reflecting the cardinal points plus an area in the center. The parsed metadata and the average color of the video frames are provided to the device drivers. The device drivers are responsible for rendering sensory effects on real devices.

For providing the possibility to use sensory effects with videos embedded in Web sites we developed the Web browser plug-in. The Web browser plug-in supports the most commonly used Web browsers, i.e., Internet Explorer, Mozilla Firefox, Google Chrome, Safari, and Opera. The Web browser plug-in was created using the Firebreath Open Source Project [2]. Furthermore, the plug-in uses the AmbientLib to parse SEM and render sensory effects. The

Web browser plug-in supports frame grabbing from the Adobe Flash Player and video frames extracted from the HTML5 *video* element can be handed over to the plug-in [8]. Figure 3 depicts the Web browser plug-in using the AmbientLib to render sensory effects for a video sequence on the amBX hardware [1]. The AmbientLib and the Web browser plug-in form a so-called media processing engine. Both the Ambient Library and the Web browser plug-in are developed in C/C++. Web sites making use of the Web browser plug-in only have to include a JavaScript script which initializes the plug-in and conveys the location of the SEM description. In Listing 1 the JavaScript for loading the plug-in and handling the SEM descriptions is presented.

```
var obj;
function computeFrame() {
  if(obj.calcFrame()){
    obj.grabFrame(10,this.video.getWidth(),
      this.video.getHeight(),
      this.video.getWidth(),
      this.video.getHeight());
  }
  obj.playEffectAtTimeAbs(
    this.video.getPosition());
  return;
}
function timerCallback() {
  if (this.video.getState() == "PLAYING") {
    computeFrame();
  }
  setTimeout(function() {timerCallback();}, 40);
}
function doLoad(SEMName) {
  this.video = jwplayer();
  obj = document.getElementById("plugin");
  obj.initialize(0, this.video.getWidth(),
    this.video.getHeight());
  obj.setCalculationRate(0,0);
  obj.loadSEM("http://www.yourdomain.com/SEMs/",
    SEMName);
  setTimeout(function() {timerCallback();}, 40);
}
```

Listing 1. JavaScript for Loading and Handling SEMs.

For using the JavaScript from Listing 1 one has to insert the code presented in Listing 2 into the Web page. Note in Listing 2 we used the JWPlayer for playback but one is free to use any other player but has to adjust Listing 1 and Listing 2 accordingly.

```
<script>
function gup(name) {
  name = name.replace(/[\/]/, "\\");
  .replace(/[\/]/, "\\");
  var regexS = "[\\?&]" + name + "=([^\&#]*)";
  var regex = new RegExp(regexS);
  var results = regex.exec(window.location.href);
  if(results == null)
    return "";
  else
    return results[1];
}
function initialize() {
  var video = gup("video");
  var sem = video + ".xml";
  var videoP = "http://www.yourdomain.com/videos/"
    + video + ".FLV";
  jwplayer('video').setup({
    flashplayer: 'player.swf', id: 'video',
    width: '1280', height: '720', wmode: 'window',
    file: videoP, events: {
      onReady: function(event){
        doLoad(sem); jwplayer().play();
      },

```



Figure 3. Ambient Library and Web Browser Plug-in.

```
onComplete: function(event){
  window.location = 'index.html';
}
});
</script>
<body>...
<div id="video"></div>
<object id="plugin" type="application/x-ambient"
  width="0" height="0">
  <param name="onload" value="initialize"/>
</object>...
</body>
```

Listing 2. Code for Activating Sensory Effects in a Web Page.

3.4 Sensory Effect Media Player

The Sensory Effect Media Player (SEMP) is a multimedia player written in C/C++ which uses DirectShow for video playback. In addition to the multimedia playback, SEM is able to parse SEM descriptions and render sensory effects on real-world devices such as the amBX system [1] (light, wind, vibration), Cyborg Gaming Lights [4] (light), and Vortex Activ [5] (scent). Currently, SEM supports light, wind, vibration, and scent effects through their respective devices as specified in Part 3 of MPEG-V.

Furthermore, SEM offers the functionality for taking user ratings between 1 and 5 which are stored in a log file together with a timestamp which can be used for subjective quality assessments.

Additionally, SEM supports an automatic color calculation. It uses a SampleGrabber for retrieving a frame every 0.1 second and splits the frame into three parts similar to SESim (cf. Section 3.2) and renders the color on the corresponding device.

4. DEPLOYMENTS

During the development of this toolset a series of subjective quality assessments were conducted [11][17][18][20]. Therefore, we used SEVino to create SEM descriptions for various video sequences and received user ratings either via SEM (cf. Section 3.4) or the Web browser plug-in (cf. Section 3.3), depending on the used scenario. In the following we shortly discuss the conducted user studies and their results to show the benefits of our toolset and MPEG-V.

In our first user study [18] we showed the benefits of using additional sensory effects (i.e., light, wind, and vibration) with multimedia content. Thus, we used different video sequences from different genres (i.e., action, sports, commercial, news, and documentary) – generated with SEVino – and presented each video sequence once with sensory effects and once without sensory effects – using SEM. The participants rated their viewing experience while watching the video sequences with

sensory effects compared to the same sequence without sensory effects. The results show that sensory effects are enriching the viewing experience for the action, documentary, and sports genre. On the other side, the genres news and commercial only benefit slightly from sensory effects or sometimes are even annoying.

Our second user study [17] evaluated the influence of sensory effects on the perceived video quality. We used the documentary and action sequences from our previous study [18] and generated for each sequence four different bit-rate versions (2-6 Mbit/s). SEMP was used again and the participants had to rate the video quality for each video sequence independent from whether or not they come along with sensory effects. The results show that sensory effects can conceal visible artifacts and, thus, receive a higher perceived video quality. The average increase in the participants' rating is around 0.5 Mean Opinion Score (MOS) points for both video sequences.

In our third user study [20] we shifted the context to the World Wide Web and performed two evaluations using AmbientLib and the Web browser plug-in. Additional Web video sequences were annotated using SEVino. First, we evaluated – similar to our first study [18] – the impact of sensory effects on the QoE but this time only with light effects. Second, we measured the performance of the automatic color extraction. The purpose of this evaluation was to determine if it is better to skip entire frames, pixel rows, or pixel columns when there are performance issues (e.g., a too slow computer). The results of these studies are the following: the first study more or less confirmed our results from the first study but in the Web context. The second study showed that it is better to skip entire frames than rows and/or columns if there are performance issues. Skipping of rows and/or columns can lead to mismatching colors and, thus, lead to a worse viewing experience.

Finally, in our fourth user study [11] we investigated the impact of sensory effects on the users' emotional response in Web videos. Similar to our third user study we performed the evaluations in the context of the Web using the AmbientLib/Web browser plug-in. In this study we investigated again the enhancement of the QoE using sensory effects. Furthermore, we evaluated the impact of sensory effects on emotions. This study was performed jointly at Alpen-Adria-Universität Klagenfurt (Austria), RMIT University (Australia), and University of Wollongong (Australia). The results of this study more or less confirmed our previous studies regarding QoE. For emotions, the study indicated that emotions such as fun, worry, interest etc. are increased and, thus, provide the viewer an immersive viewing experience.

5. CONCLUSION AND FUTURE WORK

In this paper we presented an overview of the MPEG-V standard and an open-source toolset for generating and consuming MPEG-V-compliant descriptions. Our toolset can be used for enriching multimedia content in an easy way and, thus, providing a more immersive viewing experience. The presented deployments and evaluation results are promising and encouraging for further investigations in this area.

In the future we want to support further sensory devices and the streaming of sensory effects. Furthermore, we want to investigate the (semi-)automatic extraction of sensory effects other than light.

6. ACKNOWLEDGMENTS

This work was supported in part by the EC in the context of the ALICANTE (FP7-ICT-248652), SocialSensor (FP7-ICT-287975), and QUALINET (COST IC 1003) projects.

7. REFERENCES

- [1] amBX UK Ltd. <http://www.ambx.com>. (Last access: April 2012).
- [2] Bateman, R. Firebreath Project. <http://ww.firebreath.org/>. (Last access: May 2012).
- [3] Choi, B., Lee, E.-S., and Yoon, K. 2011. Streaming Media with Sensory Effect. *Int'l Conference on Information Science and Applications (ICISA)*, Jeju Island, Korea.
- [4] Cyborg Gaming Lights. <http://www.cyborggaming.com>. (Last access: April 2012).
- [5] Dale Air. <http://www.daleair.com>. (Last access: April 2012).
- [6] FFmpeg. <http://www.ffmpeg.org>. (Last access: April 2012).
- [7] Ghinea, G. and Ademoye, O. A. 2010. Olfaction-enhanced Multimedia: Perspectives and Challenges. *Multimedia Tools and Applications*. Springer, Netherlands, 601-626.
- [8] Hickson, I. W3C Working Draft. <http://www.w3c.org/TR/HTML5>. (Last access: May 2012).
- [9] ISO/IEC 23005. 2011. Information technology - Media context and control.
- [10] Java Framework for the vlc Media Player (VLCJ). <http://code.google.com/p/vlcj/>. (Last access: April 2012).
- [11] Rainer, B., Walzl, M., Cheng, E., Shujau, M., Timmerer, C., Davis, S., Burnett, I., Ritz, C., and Hellwagner, H. 2012. Investigating the Impact of Sensory Effects on User Quality of Experience and Emotional Response in Web Videos. *Proceedings of the 4th Int'l Workshop on Quality of Multimedia Experience (QoMEX)*, Yarra Valley, Australia.
- [12] Scent Sciences. <http://www.scentsscience.com>. (Last access: April 2012).
- [13] Sensory Acumen. <http://www.sensoryacumen.com>. (Last access: April 2012).
- [14] Sensory Experience Lab (SELab). <http://selab.itec.aau.at>. (Last access: April 2012).
- [15] Seuntjens, P., Vogels, I., and van Keersop, A. 2007. Visual Experience of 3D-TV with Pixelated Ambient Light. *Proceeding of the 10th Annual Int'l Workshop on Presence*. 339-344.
- [16] Universal Studios. <http://www.universalstudios.com>. (Last access: April 2012).
- [17] Walzl, M., Timmerer, C., and Hellwagner, H. 2010. Improving the Quality of Multimedia Experience through Sensory Effects. *Proceedings of the 2nd Int'l Workshop on Quality of Multimedia Experience (QoMEX)*, Trondheim, Norway.
- [18] Walzl, M., Timmerer, C., and Hellwagner, H. 2010. Increasing the User Experience of Multimedia Presentations with Sensory Effects. *Proceedings of the 11th Int'l Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, Desenzano del Garda, Italy.
- [19] Walzl, M., Timmerer, C., Rainer, B., and Hellwagner, H. 2012. Sensory Effect Dataset and Test Setups. *Proceedings of the 4th Int'l Workshop on Quality of Multimedia Experience (QoMEX)*, Yarra Valley, Australia.
- [20] Walzl, M., Timmerer, C., Rainer, B., and Hellwagner, H. 2012. Sensory Effects for Ambient Experiences in the World Wide Web. *Multimedia Tools and Applications*, Springer, Netherlands, 1-20.