Enhancing audiovisual experience with haptic feedback: a survey on HAV

Fabien Danieau, Anatole Lécuyer, Philippe Guillotel, Julien Fleureau, Nicolas Mollet, Marc Christie

To cite this version:

Fabien Danieau, Anatole Lécuyer, Philippe Guillotel, Julien Fleureau, Nicolas Mollet, et al.. Enhancing audiovisual experience with haptic feedback: a survey on HAV. IEEE Transactions on Haptics (ToH), IEEE, 2013, 6 (2), pp.193-205. <hal-00766259>

Enhancing audiovisual experience with haptic feedback: a survey on HAV

Fabien Danieau, Anatole Lécuyer, Philippe Guillotel, Julien Fleureau, Nicolas Mollet, and Marc Christie.

Abstract—Haptic technology has been widely employed in applications ranging from teleoperation and medical simulation to art and design, including entertainment, flight simulation and virtual reality. Today there is a growing interest among researchers in integrating haptic feedback into audiovisual systems. A new medium emerges from this effort: haptic-audiovisual (HAV) content. This paper presents the techniques, formalisms and key results pertinent to this medium. We first review the three main stages of the HAV workflow: the production, distribution and rendering of haptic effects. We then highlight the pressing necessity for evaluation techniques in this context and discuss the key challenges in the field. By building on existing technologies and tackling the specific challenges of the enhancement of audiovisual experience with haptics, we believe the field presents exciting research perspectives whose financial and societal stakes are significant.

Index Terms—haptic interfaces, multimedia, audiovisual, user experience

◆

1 INTRODUCTION

In 1962, Heilig introduced *Sensorama*, a system where one could watch a 3D movie, sense vibrations, feel the wind and smell odors [1]. This pioneering work opened the path for research in virtual reality, providing high-end interfaces that involve real-time simulations and interactions through multiple sensorial channels [2]. The importance of the sense of touch (haptics) to the immersion of the user in the virtual reality environment has been particularly studied. But the use of haptic technologies is much wider than the field of virtual reality: numerous applications have been found for them in medical, robotic and artistic settings. However, it is virtual reality that has triggered the development and evaluation of numerous haptic interfaces that enable the study of complex physical interaction with virtual objects [3], [4].

In contrast, research and technology for audiovisual entertainment remains essentially focused on improving image and sound. Although the potential industrial impact appears to be significant, haptic feedback in a multimedia context, in which haptic feedback is combined with one or more media such as audio, video and text, remains underused. Only a few systems, known as "4D-cinemas", currently exploit this technology. However, the number of articles reporting the potential of haptic feedback for multimedia is increasing. In parallel, contributors working with virtual reality, such as Reiner [5], have showed haptic feedback to be a key factor in user immersion, and thus of great interest to entertainment applications.

Recent works defend this view. O'Modhrain et al. have demonstrated that the benefits of haptic feedback observed in virtual reality are applicable to multimedia applications [6]. Haptic feedback may open new ways to experience audiovisual content: the relation between users and audiovisual content is not limited to a passive context where the user just listens and watches but could enable physical involvement in a more immersive experience [7]. As well as physical sensations in parallel with the audiovisual content, the user could expect to receive a complementary piece of information or to intensify an emotion through haptic interaction, moving beyond simple immersion. The combination of haptics and audiovisual content becomes the complete medium of haptic-audiovisual (HAV [8]) content, worthy of study distinct from virtual reality, with its own specific requirements and scientific challenges.

This fresh field of study introduces many new questions. How to deliver haptic technology to the user? To what extent can haptics affect the user's perception and understanding of the audiovisual content, and how can haptics be employed efficiently in conjunction with image and sound? What about the acceptability of complex haptic interfaces for users? How will the quality of the user experience be evaluated? Moreover, to what extent can the same haptic effect be experienced in different viewing scenarios (mobile tv, cinema or user living space, potentially shared) with possibly different devices?

The aim of this review is to gather and classify the results obtained in this young field of research by identifying its key challenges. We then propose future paths for research.

The paper is organized as follows. We first describe a general workflow for adding haptic effects to audio-

---

- *F. Danieau is with Technicolor R&I, France and VR4I, INRIA France. e-mail: fabien.danieau@technicolor.com*
- *A. Lécuyer is with VR4I, INRIA, France.*
- *P. Guillotel, J. Fleureau and N. Mollet are with Technicolor R&I, France.*
- *M. Christie is with MimeTIC, IRISA, France.*

visual content and build on this workflow to detail its three main stages: (i) production of haptic effects, (ii) distribution of haptic effects and (iii) rendering of haptic effects. We then discuss techniques for evaluating the quality of experience of users (QoE) in such systems. We conclude by discussing the developing prospects in the field.

2 A WORKFLOW FOR ADDING HAPTIC FEEDBACK TO AUDIOVISUAL CONTENT

This review is organized in a manner analogous to the typical workflow for video-streaming. This comprises three stages: (i) production, (ii) distribution and (iii) rendering (see Figure 1). We use the term "haptic effect" to designate the use of a haptic feedback in audiovisual content (a generalization of the term employed in the specific context of video viewing [6], [9], [10]).

The first stage in the workflow deals with the **production** of the content, i.e. how haptic effects can be created or generated in synchronization with the audiovisual content. Three techniques emerge from the literature: the capture and processing of data acquired from sensors, automatic extraction from a component of the audiovisual content (image, audio or annotations) and manual authoring of haptic effects. These production techniques and tools will be reviewed in Section 3.

The second stage in the workflow deals with the **distribution** of haptic effects. Current technologies allow mass distribution of media over networks, so there is a strong requirement for haptic effects also to be distributable in this way. This raises questions on formalizing haptic effects. The synchronized transmission of haptic effects over networks is termed haptic broadcasting [11]. Various models, formalizations and techniques for this are reviewed in Section 4.

Finally an encoded haptic effect is **rendered** on a specific haptic device and experienced by the user. Section 5 offers an overview of the wide range of published techniques classified by the type of device (wearable, handheld, desktop or seat).

The evaluation of the user experience is a key aspect that cuts across production, distribution and rendering. Most interest to date has focused on the technical aspects of these three stages, but there is also a clear necessity to measure the quality of haptic-enhanced audiovisual experiences and provide common tools and metrics for such evaluations. This **quality of experience** (QoE see [12]) is reviewed Section 6.

3 PRODUCTION

Production is the task of creating haptic effects in order to enhance audiovisual content. Three methods to create them have been reported in the literature: (i) capturing haptic effects from the real world using physical sensors, (ii) generating haptic effects by an automated analysis of audio and/or visual contents, and (iii) manually synthesizing haptic effects from scratch or by editing effects obtained with the previous methods. Haptic effects will be classified according to their perceptual characteristics (tactile, kinesthetic, and proprioception).

Haptic Perception	**Haptic Effect**	**Ref.**
Tactile	Temperature	[13] [14]
	Vibration	[15] [16] [17] [18] [19] [20] [14]
	Pressure	[21] [14] [22] [23]
Kinesthetic	Movement	[24] [25]
	Force	[26] [9] [10] [25] [27]
Proprioception	Body Motion	D-Box[9] Mediamation[10]

Table 1
List of potential haptic effects for audiovisual content. Individual effects can be combined to create complex effects.

3.1 Haptic effects for audiovisual content

Waltl's classification of haptic effects is the most exhaustive yet published [18]. He details several sensory effects such as taste, smell and haptic. Haptic effects reported were temperature, wind, whole body vibration, water sprayer, passive kinesthetic motion and force (the user simply holds a force-feedback device), active kinesthetic (the user can explore actively the content thanks to a force-feedback device), tactile and rigid body motion (the whole body of the user is moved as in motion simulators). This classification linked each effect to a specific device.

In contrast, the classification we propose is based on haptic perceptual capabilities. Haptic feedback is often separated into two categories: tactile and kinesthetic feedback. There are three types of tactile stimuli: perception of vibration, of pressure [28] and of temperature [29]. Two types of kinesthetic stimuli may be defined [30]: perception of movement and limb position and the perception of forces. Finally, haptic perception may result from the motion of the user's own body [31]. Both the vestibular system and proprioception contribute to the perception.

We then propose a table summarizing haptic effects in HAV systems in which each category is mapped to contributions from the literature (see table 1). The reader may also refer to the guidelines for the design of vibrotactile effects [32] or haptic feedback in multimodal environments [33]. These individual effects can be combined to create more complex effects. For example, the haptic effect associated with an explosion might be defined with a combination of temperature and vibration.

Haptic effects are mostly used to represent physical events which occur in the scene (see references in table 1). The user perceives stimuli which are directly related to the audiovisual content (e.g. bumps when driving off-road), augmenting the physical event and the sense of "being physically present". However other

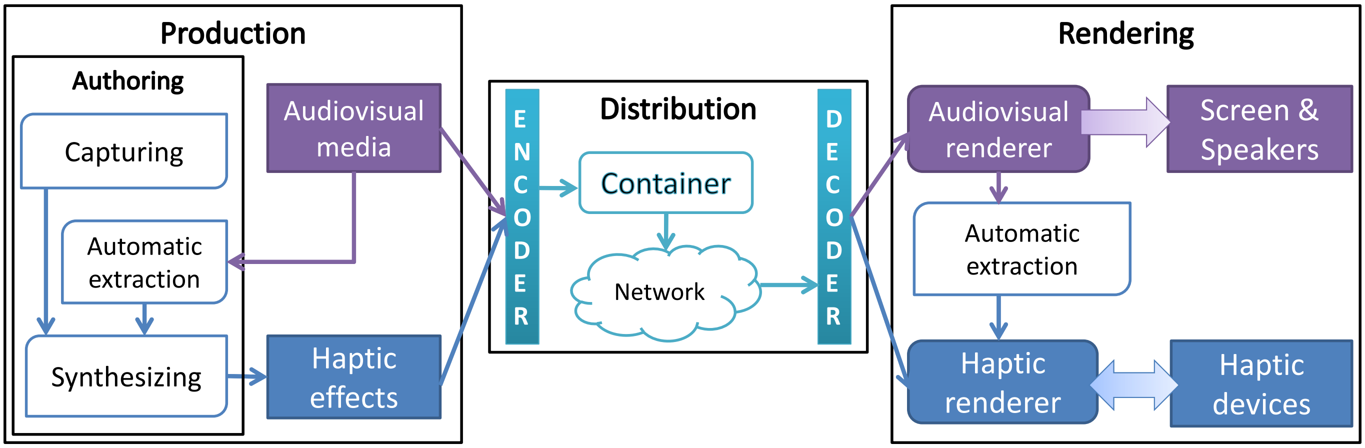


Figure 1. Workflow for adding haptic effects to audiovisual content. In this review, we consider haptic effects as a component of a multimedia content. Effects are typically produced, distributed and rendered in the user living space in parallel to the audiovisual content.

aspects of an audiovisual content, such as ambiance, can be enhanced [20]. The role of haptic effects in audiovisual content is analogous to that of audio in movies: audio is used for increasing the realism (sound effects) and to create ambiance (music). In movies, a clear separation is drawn between diegetic sounds (a sound for which the source belongs to the diegesis, the recounted story) and non-diegetic sounds (a sound for which the source is neither visible nor implied in the action, typically such as a narrator's comment or mood music). Non-diegetic haptic effects have similar potential. Non-visual content could be augmented by providing additional information that is perceived by the user.

The use of haptic effects to enhance ambiance or emotion is not straightforward. The haptic effect designer may explore results from research on affective haptics: recent works attempt to communicate affect with haptic feedback [34] or trigger users' emotions with the help of haptic patterns [17], [35].

3.2 Capturing haptic effects from the real world

One approach to creating haptic effects is to capture haptic effects related to an object or actor in a scene. Piezo-electric sensors can also be used to capture forces [6] or vibrations but, most of the time, accelerometers are used to record accelerations and deduce forces applied to the targeted object. Brady et al. equipped a radio-controlled car to capture accelerations on X, Y and Z axes [36]. These recorded data were then directly transmitted and rendered to the user's control device. Recorded accelerations on X and Y axes control an embedded 2DoF force-feedback device and acceleration on the Z-axis drives a vibration device. Similarly, Danieau et al. placed a camera together with an accelerometer on an actor's chest to capture a first-person point-of-view video and the associated motion [27]. Different scenarios were used to capture different kinds of movements: riding a bike, riding a horse and being in a car as it braked or turned. The videos were then replayed with haptic effects of force generated from the recorded accelerations. Kuchenbecker et al. recorded haptic events in a database to enable replay later [37]. The authors recorded accelerations resulting from the impact of a stylus on different materials (wood, foam). These accelerations were transduced into forces and replayed by a force-feedback device when the user touched virtual materials.

A second approach consists of capturing haptic effects related to a whole scene. Depth (or 2.5D) cameras have been used to build touchable images [10]. A more precise result could be obtained with 3D trackers [38] but these devices are more expensive and the analysis of the scene would take longer. The problem of capturing haptic effects remains strongly constrained by the available hardware. In contrast to video and sound recording, only a limited number of devices exist, mainly accelerometers and 3D cameras with considerable variations in precision and cost.

3.3 Automatic extraction of haptic effects from audiovisual content

Haptic effects can also be created automatically by extraction. The key idea is to generate haptic effects which are consistent with media content in order to highlight specific aspects. For example a scene showing an explosion could be enhanced by haptic feedback such as vibrations and heat. Video and sound analysis might be used to detect explosions and then automatically add haptic effects.

Automatic extraction can occur in the production stage or in the rendering stage (see figure 1). In the production stage, haptic effects are automatically generated and can be modified by the creator. In the rendering stage, haptic effects are automatically generated on the client side.

3.3.1 Generation from visual content

A classical way to extract content from an audiovisual media consists in using video analysis techniques. Typical algorithms rely on feature detectors to extract points of interest inside an image to build derived information (e.g. object identification) [39]. There are significant variations in the features they offer such as robustness to light variations, motion and computational cost. Some specific algorithms are dedicated to the detection of specific features such as faces [40] or motion [41]. Detecting events is also possible. Video abstraction [42] and video data mining [43] have both been used for event detection but are restricted to specific subjects such as sports, where the potential range of events is limited. Once a targeted event is detected in the audiovisual content, a haptic effect could be generated. For instance, Réhman et al. have shown how to automatically extract events from a soccer game video and to display them with a vibrotactile device [16]. Five vibration patterns were designed to represent the position of the ball on the field, to the team leading the game or to the goals. However the main focus was on how to render the effects rather than the video analysis. An earlier study was conducted in the context of sensory substitution, but the aim was to use haptic feedback to replace visual information rather than using haptics to enhance these data [44].

The difficulty of direct extraction of haptic information from video was pointed out by Mc Daniel et al. [38]. To simplify the problem, the authors built a database which maps visual information (a picture of an object) to haptic information (the 3D shape of the object). The database is used to generate appropriate haptic feedback for each object identified from visual information.

Even if computer vision provides a broad range of tools, most techniques to analyze and generate haptic feedback have not been yet explored in detail. The robustness and adaptability of the detection algorithms remain typical issues in the field [39].

3.3.2 Generation from audio content

Haptic effects can also be created from the audio content within audiovisual media. The main approach is to transduce an audible signal into a signal suitable for vibration motors. Chang and O'Sullivan used a band-pass filter to isolate frequencies compatible with a targeted vibration motor and then amplify and render the output signal on this device [45]. This system was developed for mobile phones which then vibrate according to ringtones. The MOTIV[1] development platform from Immersion is a similar commercially available system. The "Reverb" module allows the automatical addition of haptic effects to any application using the output audio stream.

The approach selected by Nanayakkara et al. is even more direct and does not require any processing of the audio stream [46]. The authors developed a chair for deaf people which renders music and vibration. The sound is played by speakers attached to the seat, which are specially designed to propagate vibrations to the surface they are attached to.

Most research follows this straightforward technique of the transduction of audio into vibrations. The approach could be extended by attempting to represent the information conveyed by the audio stream. Audio analysis techniques to extract specific features would then be useful. For example the system described by Zhang and Kuo permits the identification of music, speech and environmental sound in an audio signal [47].

3.3.3 Generation from metadata

Metadata can contain information about movements or physical properties of objects within the media. Yamaguchi et al. extracted data from a Flash[2] animation to compute force feedback as the user explores the content [9]. Since this format allows access to the geometry and position of elements within a 2D animation, it is possible to compute a force-feedback related to one of the objects in the scene. The authors defined a virtual mass for the targeted object and then computed a force-feedback relative to the acceleration and mass of this object. This technique can be applied to computer animations where a 3D model of the scene is available. But the system remains specific to animations and is not suitable for standard video. However some data formats allow for the description of audiovisual content. The MPEG-7 standard focuses on the description of multimedia content and can contain a description of movement within a scene [48], opening many possibilities for the generation of haptic effects.

3.4 Graphical creation tools for synthesizing haptic effects

Although haptic effects can be created automatically, the need to create them before their integration with audiovisual content remains. Original effects may need to be edited. Neither of these functions can be automated.

Two main categories of graphical creation tools have been designed. The first allows users to specify the behavior of one or several actuators. In this case the designer has to use the same device as the end-user. In the second category the designer edits haptic cues that the user will perceive without referring to specific hardware. Various data formats and graphical tools are summarized in table 2.

3.4.1 Device-oriented effects

The behavior of an actuator is typically controlled by specifying a curve representing the amplitude of the stimulation (vibration or the force in time). The Haptcons editor [49] was created to edit trajectory patterns called "haptic icons" on a 1DOF force feedback device

---

[1] http://www.immersion.com/products/motiv/

[2] http://www.adobe.com/products/flash.html

(a knob). This kind of tool is already used in the industry. The aforementioned MOTIV[1] development platform provides a curve editor for designing vibrotactile patterns for various devices (mobile phones, gamepads, etc.).

Quite different graphical interfaces are used to edit the behavior of an array of motors. The user must specify the behavior of each motor in time. Representative examples have been developed by Rahman et al. [19] and Kim et al. [20].

3.4.2 User-oriented effects

The second type of graphical tool focuses on describing what the user should feel instead of defining how actuators should behave. This implies that the haptic rendering is handled by dedicated software.

Ryu et al. have created the posVib Editor to edit vibrotactile patterns [50]. The intensity of the vibration felt by the user is represented by a curve.

The MPEG Rose Annotation tool was designed to associate sensory effects to multimedia content [18] (see Section 3.1). It allows the designer to tune sensory effects all along a movie. One or several effects can be added on a timeline which determines when they start and when they finish.

A different approach consists in describing material properties of objects within a scene. It implicitly determines what users feel when they touch objects. This type of tool resembles a 3D editor in which the author directly visualizes the 3D object being manipulated, but haptic (friction, stiffness) rather than visual properties are edited. We refer the readers to the presentation of the K-Haptic Modeler [51] as well as the HAMLAT tool [52] which is a graphical editor for HAML (see Section 4.1.1).

4 DISTRIBUTION

The second stage consists in formalizing haptic effects into data to be synchronized, stored and transmitted with the audiovisual media. Even though the range and nature of haptic effects is not yet well defined, there have been several attempts at providing formalizations. These formats are summarized in table 2 which displays, when available, the associated authoring tools (see Section 3.4), and solutions to transmit haptic effects over the network (see Video Container column of table 2).

4.1 Data formats for haptic effects

Though there are several contributions which use dedicated formats to encode haptic feedback for audiovisual content, most approaches rely on generic formats. We consider two ways to formalize haptic effects: "device-oriented" that defines the actuators' precise behavior, and "user-oriented" that describes effects from the user's point of view. The formats presented in this section are however suitable for both usages. Choosing between them only influences the way in which the rendering stage has to be handled: device-oriented data are used to control haptic devices directly, but user-oriented data must be interpreted. Since there is no obvious way to classify the encoding of haptic effects, we will use a per-format classification. We will detail contributions based on XML, a versatile description language, CSV a simple format to store data and VRML, a language dedicated to descriptions of 3D worlds. These formats are summarized in Table 2.

The issue of formalizing haptic effects has been solved by companies such as D-Box[9] or Immersion[1] who have developed commercial solutions for rendering haptic effects along with audiovisual content. D-Box have created a proprietary language to add haptic effects to a movie, called D-Box Motion Code™. However, details of these formats are not currently available and the effects cannot be edited by the end-user.

4.1.1 XML-based

The first method of formalizing haptic feedback relies on XML language. The Haptic Application Meta-Language (HAML [54]) is a generic format for describing haptic feedback which contains information about the haptic device, haptic rendering and visual rendering (see Listing 1). The purpose of this format is to be able to use any haptic interface with any virtual world, the system adapting the haptic rendering to the capabilities of the haptic interface used. This language is dedicated to virtual reality applications but it could be used to describe scenes in audiovisual content: objects and their location, geometry, haptic properties (stiffness, damping, friction), etc. This format respects the MPEG-7 standard which yields standardized tools to structure and organize descriptions of multimedia content [48].

```
<HAML>
...
    <SceneDS>
        <Object>
            <Type>Mesh</Type>
            <Name>Cube</Cube>
            ...
            <Tactile>
                <Stiffness>0.8</Stiffness>
                <Damping>0.9</Damping>
                <SFriction>0.5</SFriction>
                <DFriction>0.3</DFriction>
            </Tactile>
        </Object>
    </SceneDS>
</HAML>
```

Listing 1. Example of an xml-based file (HAML [8]). Here, the haptic properties (stiffness, friction and damping) of a 3D cube are defined.

Closely related to video viewing, the Sensory Effect Description Language described by Waltl also relies on XML [18]. This language is designed to add sensory effects to any multimedia content: movies, video games, web content, etc. Users can create groups of effects and synchronize them with other media (see Section 3.1 for the list of effects). For each effect the designer can specify at least its intensity and duration. However devices

Type of Effect	Format	Data Content	GUI	Video Container	Ref.
User-oriented	MPEG-V (XML)	Description and organization of sensory effects in a multimedia content	Yes (MPEG RoSE Annotation Tool)	MPEG-2 TS	[18] [53]
	MPEG-7 (XML)	Description of a 3D scene, haptic device and haptic rendering	Yes (HAMLAT)	n/a	[54] [52]
	XML	Haptic properties of a 3D scene: friction, stiffness, etc. of objects	Yes (K-HapticModeler)	n/a	[51]
		Vibration patterns	Yes (PosVib Editor)	n/a	[50]
	VRML	Description of 3D objects and associated haptic rendering methods	No	n/a	[55]
	MPEG-4 BIFS (VRML)	Information about depth, stiffness, friction of a scene	No	MPEG-4	[10]
	CSV	Information about motion into a scene	No	n/a	[27]
Device-oriented	CSV	Trajectory patterns	Yes (Hapticon Editor)	n/a	[49]
	XML	Description of haptic device properties and description of how they are activated	Yes (TouchCon)	n/a	[56]
		Vibration patterns of a tactile array	Yes	n/a	[19]
	MPEG-4 BIFS (VRML)	Vibration patterns of a tactile array	Yes	MPEG-4	[20] [25]

Table 2
Overview of existing formats to edit and store haptic effects. Two types of haptic effect can be described: effects focused on what the user will perceive (user-oriented), and effects focused on how the actuators will behave (device-oriented). Most of the time a graphical user interface is designed to easily edit data. Some formats are to be embedded with a container enabling both audiovisual and haptic contents to be distributed via streaming platforms.

and techniques to render effects are not specified. If converting an intensity into vibrations is simple, the rendering of a forward movement over 2 meters with an acceleration of $30cm.s^{-2}$ is less straightforward (see Listing 2). At the time of writing this paper, this language is close to being standardized by the MPEG working group as the MPEG-V format [57].

```
<sedl:SEM>
    <sedl:Effect xsi:type=”sev:
         RigidBodyMotionType” activate=”true” si:
         pts=”1593000”>
        <sev:MoveToward distance=”200”
             acceleration=”30”/>
    </sedl:Effect>
    <sedl:GroupOfEffects si:pts=”1647000”>
        <sedl:Effect xsi:type=”sev:VibrationType
             ” activate=”true” intensity−range=”0
              100” intensity−value=”10”/>
        <sedl:Effect xsi:type=”sev:WindType”
             activate=”true” intensity−range=”0
             100” intensity−value=”5”/>
    </sedl:GroupOfEffects>
</sedl:SEM>
```

Listing 2. Example of an xml-based file (MPEG-V [18]). Here a “Move Toward” effect is defined followed by a group of effects combining “Wind” effect and a “Vibration” effect.

In an approach dedicated to instant messaging applications, Kim et al. [56] developed an XML-based format to exchange haptic feedback called “TouchCons”. This allows users to send haptic messages such as vibration patterns or thermal effects. Two main files are used in this system. First, the Library XML describes a list of haptic messages and how they should be rendered (device used, intensity, duration). Second, the Device XML describes the available devices and associated capabilities. To send a message, the user chooses one from the Library XML file. When he receives a message, it is rendered according to the capabilities of the devices listed in the user’s Device XML file. This framework could be used, instead of TouchCons, to describe haptic effects and then to send them to the end-user. The effects would be then rendered according to the user’s devices configuration.

Finally XML representation can be used to determine the behavior of actuators directly. For example, Rahman et al. [19] described vibration patterns of a vibrotactile array: the vibration intensity of each motor is specified in an XML file. This approach is simple but the effects described can be rendered only by a specific device.

4.1.2 CSV-based

Comma Separated Values (CSV) is a file format where data are stored in a simple text file separated by commas. Enriquez et al. relied on this format to store knob positions [49]. This direct approach is simple but device specific. Danieau et al. [27] also used this type of format but the authors stored information about the motion embedded in a video (acceleration in $m.s^{-2}$ on 3 axes for each instant t). The motion effect is then rendered by the user's haptic device.

4.1.3 VRML-based

A third method used to describe haptic content uses VRML/X3D. This language serves to represent 3D worlds and contains information needed by visual rendering systems. Sourin and Wei [55] proposed an extension of this language by adding haptic rendering techniques. One purpose of this language is to transmit virtual objects and their associated haptic rendering algorithms over the internet. In a similar way to HAML, this solution allows an audiovisual scene and the associated rendering techniques to be described.

The two techniques presented hereafter are based on the MPEG-4 BIFS format, also known as MPEG-4 Part 11 [58]. BIFS, which stands for Binary Format for Scenes, is a scene description protocol based on VRML. Cha et al. extended this format to add haptic properties to a video [10]. The authors built a ”touchable” movie, i.e. a movie in which spectators can feel the depth of the images using a force-feedback device. For each frame of the video the authors associated texture properties (stiffness, static friction and dynamic friction; see Listing 3).

```
Shape{
  appearance Appearance {
    texture ImageTexture {
      url ”color_image.jpg”
      }
    hapticSurface HapticTextureSurface {
      stiffnessRange 0.1 10
      staticFrictionRange 0.2 0.9
      dynamicFrictionRange 0.3 0.9
      maxHeight 1.0
      hapticTexture ImageTexture{
        url ”haptic_image.jpg”
        }
      }
    }
  geometry Depth {
    focalLength 6.983
    pixelWidth 0.00123
    nearPlace 10
    farPlane 200
    texture ImageTexture {
      url ”depth_image.png”
      }
    }
}
```

Listing 3. A VRML-based file (Extended MPEG-4 BIFS [10]). This file describes haptic properties of a visual scene (color_image.jpg). The depth map and associated friction are specified.

This modified BIFS format can also be used to store vibrotactile patterns used to drive an array of vibration motors. Kim et al.'s encoded a pattern in a grey-scale image where each pixel represents an actuator and the intensity of the pixel corresponds to actuator activation intensity: from black (0) for idle to white (255) for maximal vibration [20]. In a similar way, vibrotactile patterns can be associated with video frames (see Listing 3: instead of ”haptic_image.jpg” a ”tactile_pattern.jpg” would be associated with the visual scene). Thus the MPEG-4 BIFS format extended by Cha et al. can both describe a 3D scene and/or contain data to drive vibrotactile arrays. These two possibilities have been implemented by Kim et al. for adding haptic textures effects or vibration effects to educational videos [25].

4.2 Haptic-video containers

A container is a meta-file format that can hold several files in a single file which makes distribution easier. In the HAV context, a container is a single file that regroups haptic, visual and audio content. This stage is depicted in Figure 1. All components are compressed and synchronized into a single container for network transmission [59]. These containers are mainly used in multimedia applications to store both audio and visual content into a single file.

Several containers embedding audio and video exist (ogv, avi, mp4, etc.), but those combining haptic content are less common. A simple solution would consist of directly embedding the file containing the haptic data into a container that allows the attachment, such as the mkv container. O'Modhrain and Oakley used on the Flash standard to distribute videos enhanced with haptic effects [26]. They integrated haptic feedback in their home-made animation and the media was played by a web browser embedding the Immersion Web plug-in. This alternative is suitable for distribution purposes, although limited to the rendering capability of the plug-in and to a specific type of audiovisual content (animation).

To take advantage of streaming platforms, one solution is to develop formats for haptic effects compatible with video containers that permit playback as they are downloaded. Some formats (see Section 4.1) were designed to support this streaming feature. Modified MPEG-4 BIFS [10] can be embedded into a classical MPEG-4 container. In a similar way MPEG-V is compatible with the MPEG-2 TS container [53]. This streaming challenge has been identified as **haptic broadcasting** by Cha et al. [11]. This is a specific challenge different from the classical transmission of data for teleoperation [60]. The purpose is not to control a device remotely but to send multimedia containing audio, video and haptic content. The two formats presented are at an early stage of development but demonstrate the possibility of haptic broadcasting.

5 RENDERING

Once the haptic content has been transmitted to the user, the haptic device needs to decode and render the content to provide the appropriate effect (in the same way that video is displayed on the screen or audio is rendered on the speakers, see Figure 1). Here we review a list of haptic interfaces proposed for “enhanced” video viewing.

Type of interface	Device	Actuator	Haptic Effect	Ref.
Wearable	Vibrotactile armband	7x10 vibration motors	Vibrations (related to position of a ball during a soccer game)	[15]
	Vibrotactile glove	20 vibration motors (4 per finger)	Vibrations	[20]
	Vibrotactile armband or jacket	Array of vibration motors (variable size)	Vibrations	[19]
	Vibrotactile jacket	16x4 vibration motors	Vibrations (related to user's emotions)	[17]
	Vibrotactile vest	Vibration motors + solenoids + peltier elements	Pressure (gunshot), temperature (blood flow), vibrations (slashing)	[14]
	Vibrotactile vest	8 air cells	Vibrations and pressure (gunshots, acceleration, explosion)	TNGames[3]
Handheld	Mobile phone	Vibration motor	Vibrations (related to status of soccer game)	[16]
	Mobile phone	Vibration motor	Vibrations	Immersion[1]
	Remote control	2DOF joystick	Force	[26]
	Computer mouse	2DOF joystick	Force	[9]
	Portable TV	10x10 array of ultrasound transducers	Pressure	[23]
Desktop	Force-feedback device	3DOF motorized arm	Movement	[24]
	Phantom[5]	6DOF motorized arm	Movement	[25]
	Novint Falcon[6]	3DOF motorized arm	Force (texture of an image)	[10]
	Novint Falcon[6]	3DOF motorized arm	Force (motion in the video)	[27]
	n/a	Array of 324 ultrasound transducers	Pressure	[22]
	Air receiver	Array of air-jets	Pressure	[21]
	Philips[7] AmBX	Vibration motor + 2 fans (+ 2 LED spotlights)	Vibration (+ wind & light)	[18]
Haptic Seat	Vibrotactile blanket	176 vibration motors	Vibrations (related to user's emotions)	[61]
	Vibrotactile chair	3x4 vibration motors	Vibrations	[62]
	Couch	Vibration motor	Vibrations (of the whole seat)	Guitammer[8]
	Moving chair	4 compressors under chair legs	3DOF body motion (pitch, roll, heave)	D-Box[9]

Table 3
Overview of existing haptic devices used for enhancing audiovisual content.

We classified these devices into four categories: wearable devices, handheld devices, desktop devices and haptic seats. The results are presented in Table 3.

5.1 Wearable devices

Wearable devices are designed to be worn by as the user experiences audiovisual content. Typically they are composed of several vibrotactile actuators embedded into clothes, as detailed in Rahman et al. [19]. This topic has been intensively studied for virtual reality purposes [63] and many devices have been designed.

In the HAV context, exploring the idea of enhancing live sports experience, Lee et al. [15] proposed a device with vibrotactile sensation through an assembly of 7x10 vibrotactors attached to the user's forearm. This prototype was used to render movements of the ball on the field during a soccer game. The tactile array was mapped to the field and vibrations were triggered at ball locations. According to the authors this device allows the user to better understand ambiguous game situations.

Kim et al. designed a tactile glove for immersive multimedia [20], [25]. It contains 20 tactile actuators per glove (4 per finger). The gloves are wireless-controlled and produce vibrotactile patterns as the user watches a movie. These patterns were first created, then synchronized with the video.

A tactile jacket has also been developed by Lemmens et al. [17]. They explored the influence of tactile devices on spectators' emotional responses, and designed a tactile jacket with 16 segments of 4 vibration motors covering the torso and the arms. Motors are activated following patterns related to specific emotions. For example, the feeling of love is enhanced by activating motors overlying the abdomen in a circular manner.

Palan et al. [14] presented a vest with embedded vibration motors, solenoids and Peltier elements. The vest was designed to display three haptic effects as realistically as possible: gunshots, slashing and blood flow, with the motivation of improving video games experience. Similarly, a commercially available jacket manufactured by TNGames[3] produces effects such as explosions, gunshots or accelerations using 8 air cells.

While the embedded devices do not yield a significant change in weight or wearability of clothes, being composed of simple vibrotactile actuators, the range of possible haptic effects is rather limited.

5.2 Handheld devices

Users can experience haptic feedback through portable devices held in the hand. Vibrotactile technology appears well-suited for portable devices. For years, the gaming industry has used vibrating joypads to enhance immersion video games. Mobile devices (phones and tablets) are now equipped with vibration motors which may be used to enhance multimedia contents[4]. Using this technology, Réhman et al. relied on a mobile phone equipped with a vibration motor to display haptic cues related to a soccer game [16]. Alexander et al. developed a prototype of a mobile TV providing tactile feedback using ultrasound [23]. The device is a screen with a 10x10 array of ultrasonic transmitters set on the reverse side. The user holds the device to observe the audiovisual content and experiences haptic feedback through the fingers.

The remote control developed by O'Modhrain and Oakley is a different sort of handheld device that provides force-feedback [26]. A gaming joystick was rehoused in a device resembling a remote control. Similarly Yamaguchi et al. used a computer mouse with a 2DOF force-feedback joystick [9].

As with clothes-based devices, handheld devices cannot embed heavy actuators and so only a restricted range of haptic effects can be rendered. However, the use of a common device in the user living space (remote control, mobile phone) seems well on the way to popular acceptance.

5.3 Desktop devices

In virtual reality settings, force-feedback devices are mainly used to interact with virtual objects. The user can feel and often modify the displayed content. With video viewing the user cannot modify the content. The user receives haptic cues, sometimes while actively exploring the content, but the audiovisual content does not change. For example in the solution devised by Gaw et al. [24], the user holds a force-feedback device and is guided along a prerecorded path while viewing a movie. The same technique was used by Kim et al. to enhance educational videos with a Phantom[5] device [25]. In a similar way, Danieau et al. used a force-feedback device to enable the user to feel the captured acceleration associated with a video [27] .

These devices have also been adapted to the task of "touching" images in a video [10]. In this study the user could actively explore the video content and received haptic feedback through a Novint Falcon device[6].

Other desktop devices have been designed to convey haptic feedback to the user without direct contact. An example is a fan which generates air streams, simulating the haptic effect of wind. Associated with a thermal device, a fan may be used to create temperature variations [13]. Fans providing wind effects are commercially available. The Philips amBX system[7] generates not only wind effects but also lighting effects and enables keyboard vibration. This kind of device is simple to use, which results in more ecological interaction.

Contact with virtual objects is possible without directly handling a device. Hoshi et al. [22] used ultrasound to exert pressure remotely on a user's skin. Their prototype was composed of an array of 324 airborne ultrasound transducers, able to exert a force of 16mN at a 20mm focal point diameter over a 180x180mm surface. This invisible surface is created at 200mm above the device. Combined with a 3D display system, the author succeeded in creating touchable floating images. A similar system has been previously developed by Suzuki and Kobayashi [21], based on air jets.

5.4 Haptic seats

Our fourth device category is the haptic seat. The user sits on a modified chair and passively senses haptic effects.

Vibrotactile actuators have once again been used in a number of ways. The tactile blanket [61], a variant for the theme Lemmens' Jacket [17], is equipped with 176 actuators and displays vibration patterns designed to enhance the user's emotion.

More recently Israr and Poupyrev [62] embedded an array of 12 vibrotactile actuators in the back of a chair, with an original controller. The user experienced the tactile illusion of a continuous stimulus though the actuators were at discrete locations.

Several commercial products in this category are already available. One example is the "couch shaker" from The Guitammer Company[8]. This device uses actuators to shake the couch or sofa, operating like a subwoofer which propagates low-frequency vibrations to the couch instead of playing sounds. Some seating devices attempt to provide more complex effects such as motion. Typically such seats are fixed on actuators or motion

---

[3] http://tngames.com/
[4] http://www.immersion.com/markets/mobile/products/
[5] http://www.sensable.com
[6] http://www.novint.com
[7] http://www.ambx.philips.com
[8] http://www.thebuttkicker.com

platforms. For example, the D-Box[9] seat features 3 DOF: pitch, roll and heave.

Haptic seats are commonly encountered in theme parks or amusement arcades where they are typically used as motion simulators. Some of them even embed several devices to provide a wide range of effects (water spray, air blast, leg ticklers, etc. See the MediaMation[10] company.) These devices are not, however, adapted to the end-user living space and their cost is prohibitive for the mass market. In contrast, the D-Box[9] seat is a consumer product designed for living room use though it remains expensive. Devices based on vibrotactile arrays are also available but the range of tactile effects which can be rendered is quite limited.

6 QUALITY OF EXPERIENCE

Haptic effects aim at enhancing the audiovisual experience. This means that the quality of experience (QoE) of a video viewing session with haptic feedback would be higher than when haptic feedback is not present. But how should this hypothesis be assessed? Jain discusses the necessity of capturing the QoE for system evaluation [64]. He underlines the difficulty of identifying and measuring the factors that characterize this metric due to its subjective nature.

Nevertheless Hamam et al. [8], [65] have proposed an initial model for the evaluation of QoE in multimedia haptics which identifies four factors: rendering quality, and the user-centered measures of physiology, psychology and perception. The rendering quality is dependent on the quality of the visual, audio and haptic feedback. Perception measures describe the way the user perceives the system depending on the user's experience, fatigue and other factors which may alter the user's perception. Physiological measures identify how the system modifies the user's biological state, and psychological measures highlight changes in mental state. The authors detail an exhaustive list of parameters related to each factor (e.g. respiration rate, body temperature or blood pressure for physiological measures). While this provides a taxonomy of the different factors influencing the quality of experience, techniques to evaluate them were not presented.

In this section we detail classical techniques to measure the QoE of HAV systems. The typical approach found in the literature is a subjective measure based on questionnaires. Other approaches capture biosignals which provide an objective measurement of the user's physiological state from which emotional state is inferred.

6.1 Subjective measures: questionnaires

Most contributions in HAV rely on simple questionnaires to evaluate the impact of haptic feedback on the quality of experience. Participants are usually asked to respond to questions on a Likert-scale. For example, Kim et al. [20] studied the benefits of vibrotactile feedback for enhancing movies by using 4 general questions (Is this more interesting than movies? Is the tactile content easy to understand? Is the tactile content related to the scene? and Does the tactile content support immersion?). Ur Rhéman et al. covered the same aspects using a more detailed questionnaire [16], while other authors have limited their analysis only to user satisfaction (see [66]).

A more elaborate approach characterizes the quality of experience using multiple factors. Hamam et al. [67] evaluated the five factors (extracted from their model described above) of realism, usefulness, intuitivism, fatigue and QoE. Danieau et al. [27] identified 4 factors: sensory, comfort, realism and satisfaction. "Sensory" characterized how the haptic feedback contributed to the immersion. "Realism" described the realism of the simulation and how it was consistent with the user's representation of the real world. "Comfort" measured the overall comfort of the system (a proxy for acceptance). "Satisfaction" measured how much the user enjoyed using the system. These 4 factors were combined into one QoE measure. This variation highlights the need for a standardized questionnaire to better evaluate and compare different systems. Identifying the factors to be measured is not an easy task, but several have already been evaluated in a systematic way: comfort, interest, acceptance and satisfaction. They can serve as a basis on which to build a subjective measure of the QoE.

6.2 Objective measures: physiological data

Another approach to the evaluation of the quality of experience consists of measuring changes in the user's physiological state. The QoE cannot be directly determined from this measure, but it can be used to infer the user's emotional state, which contributes to the QoE. To the best of our knowledge, no work has been done using these techniques in the context of HAV systems. Nonetheless, inspiring results can be found in the context of virtual reality applications and video viewing.

In the context of virtual reality, Meehan et al. gathered heart rate, skin conductance and skin temperature data from subjects in a stressful virtual environment [68]. These measures helped to determine the user's feeling of "presence" and were compared to subjective users' self-reports (see [69] for a survey on "presence"). These authors suggest that heart rate has the strongest correlation with a sensation of presence. Skin conductance correlated less strongly and skin temperature not at all. Haptic feedback significantly improved presence.

Mandryk et al. observed biosignals in video game players to determine their user experience [70]. Skin conductance, heart rate, facial muscle activity and respiration rate were captured. The authors concluded that, for most participants, playing against a friend is more enjoyable than playing against the computer. The physi-

---

[9] http://www.d-box.com
[10] http://www.mediamation.com

ological measures were significantly consistent with the self-reported measures.

In a video viewing context, Fleureau et al. studied the potential of physiological signals for detecting emotional events [71]. Participants simply watched several videos while their heart rate, skin conductance and facial muscle activity were recorded. A detector based on machine learning techniques was designed. Given the user's biosignals, the system was robustly able to determine whether user were experiencing an emotional event and if this event was positive or negative.

The physiological chosen signals in these studies were mostly similar: heart rate, galvanic skin response, and facial muscle activity. All yielded significant results despite the various settings of virtual reality, video games and video viewing. The implications for the evaluation of HAV experiences is clear. Furthermore, closed-loop systems, in which physiological signals are used to control the nature and intensity of haptic events offer interesting possibilities for adapting the haptic effects to the individual user.

7 DISCUSSION

We have presented an overview of how haptic effects can enhance audiovisual content. Studies relevant to each stage of haptic production, distribution and rendering have been presented. Some of these studies present solutions that address all stages and may be seen as implementations of the generic workflow displayed in Figure 1. These general approaches are summarized in Table 4.

While the existing solutions clearly demonstrate how haptic effects can be used with audiovisual content using tactile or kinesthetic feedback, the studies reported do not explore combinations of effects (e.g. kinesthetic and tactile). This is mostly because the devices studied have generally had only one type of actuator. As a consequence, the range of effects that can be generated is narrow and the conjunction of effects is rarely explored, despite the significant potential benefits. Furthermore, there appears to be a gap between the use of portable haptic interfaces (wearable or handheld), conveying weak effects, and complex devices (motion simulators) which are not adapted to the user living space. There is a clear opportunity to design new haptic devices dedicated to audiovisual enhancement. This implies in turn a better understanding of the requirements for HAV systems, which seem to differ significantly from those in virtual reality systems.

Further research on user perception should be conducted to determine relevant haptic stimuli for effective audiovisual entertainment. The link between haptic stimuli and user experience is not thus far well established. Haptic effects are mainly used in a similar way to the use of haptic feedback in virtual reality: to immerse the user physically in the audiovisual scene. The use of haptic effects to enhance non-diegetic aspects of a video such as the ambiance or emotions has been little studied. This appears as a key challenge and opportunity in this nascent field.

The distribution stage also requires research effort. Each solution currently uses a different technique to formalize haptic effects in the absence of a common definition for haptic effects. Only half of the studies have proposed methods for the transmission of the media to a remote display device. But several techniques allowing haptic broadcasting are emerging. Multimedia containers embedding audiovisual and haptic effects are currently being developed and standardized (MPEG-V, MPEG-4 BIFS). The MPEG-V format is a promising standard for distribution currently under development by the MPEG group. The draft standard presents a list of haptic effects along with an XML-based method to describe them. This format is also designed to be compatible with streaming technologies. However the new standard will have to follow the evolution of this emerging field of study. New haptic effects and devices will almost certainly be developed.

In most solutions haptic effects are synthesized: authors manually create and synchronize haptic effects to the audiovisual content. Each solution currently offers a different technique for editing haptic effects, though general editing tools may arrive with the advent of new standards. The automatic extraction of haptic cues from visual content has also been reported. Such cues are currently limited to specific audiovisual content: soccer games following pre-defined rules, and animations where the position and geometry of objects is already known. The automatic extraction of haptic effects for any audiovisual content remains a complex task, and more work will be necessary to adapt current algorithms to this new purpose. Extraction can be facilitated by metadata that describe the content of the media, but extracting haptic effects from videos is a new challenge for which new specific techniques will have to be designed.

One final aspect to be discussed in this review is the quantification of the benefits lent to audiovisual content by haptic effects. Some of the studies presented here have conducted user evaluations, mostly based on questionnaires. Most show that haptic effects enhance the user experience but the various studies are heterogeneous and hardly comparable. There is pressing need for shared validated tools to evaluate this quality of experience.

8 CONCLUSION

In this review we have explored the possibilities provided by haptic feedback for enhancing audiovisual content. Several trends can be identified within this emerging field. The studies presented have been arranged against a generic workflow and the key challenges that pertain to this new way of experiencing videos identified.

The first stage, related to production of haptic effects, is the identification and generation of haptic effects

Audiovisual Content		Haptic Effect	Production	Distribution	Rendering	Ref.
Category	Details					
Sport	Soccer game (3D simulation)	Vibrations (ball position)	[Automatic extraction] The system traces the ball during soccer game	n/a	Vibrotactile array embedded into an arm band	[15]
	Soccer game (simulation)	Vibrations (ball position, goals, team leading)	[Automatic extraction] Video analysis of events from a soccer game (not implemented, events are received from the simulation)	n/a	Mobile phone equipped with vibration motor	[16]
Animation	Animation (home-made with Flash)	Force (related to an object of the animation)	[Automatic creation] Force-feedback is computed from the position and geometry of the object	Flash	Mouse with a joystick (2DOF force feedback)	[9]
	Cartoon (home-made with Flash)	Force (related to onscreen character)	[Synthesis] Force-feedback is defined during edition of the cartoon	Flash	Remote control with a joystick (2DOF force feedback)	[26]
	Cartoon / Movie	Movement (user's hand is guided according to a trajectory)	[Capturing] Trajectories recorded from force feedback device	n/a	Force-feedback device	[24]
Movie	Movie	Force (user touches the image)	[Synthesis / Capturing] Material properties for each frame (depth, stiffness, etc.) stored into MPEG-4 BIFS	MPEG-4	Novint Falcon (3 DOF force-feedback)	[10]
	Movie (from Youtube)	Vibrations	[Synthesis] Vibration patterns stored into XML file	XML file on a web server	Vibrotactile array embedded into arm band or jacket	[19]
	Movie	Vibrations	[Synthesis] Vibration patterns stored into MPEG-4 BIFS	MPEG-4	Vibrotactile array embedded into gloves	[20]
	Movie	Vibrations and wind	[Synthesis] Sensory effects stored into MPEG-V file	MPEG-2 TS	Philips amBX system	[18]
	Educational video	Vibrations or force (user touches the image)	[Synthesis] Haptic effects (vibrations or haptic properties) stored into MPEG-4 BIFS	MPEG-4	Vibrotactile gloves or Phantom device (6DOF force-feedback)	[25]
	Movie	Force (related to the motion in the video)	[Capturing] The motion is captured by accelerometers	n/a	Novint Falcon (3DOF force-feedback)	[27]

Table 4

Summary of existing schemes for adding haptic effects to audiovisual content. Each system offers a solution for synchronizing and rendering haptic feedback within an audiovisual content. Some schemes also specify ways to distribute the media over the network.

which must be delivered to the user during the display of the media. We detailed different formats to store and synchronize haptic effects to audiovisual media, from a simple text-based representation to standardized XML formats. The key issue is the creation of haptic feedback. While a number of authoring tools are available, these effects may also be captured from physical sensors or generated from an other part of the media (video, audio or metadata).

Once the media has been enriched with haptic effects, it must be sent to the user. Media streaming platforms to distant users is now a common method of distribution. This stage is dependent on the way haptic data are stored. Though these issues are largely solved for audiovisual media, there are few standards for media with haptic effects. However some pioneering contributions have demonstrated the feasibility of this approach.

In the last stage the user perceives the media through a haptic device. These haptic interfaces are generally designed and dedicated to the purpose of displaying haptic cues during video viewing.

The results of our survey suggest that research effort is needed in the design of data formats and technology for distributing HAV content. The development of haptic media creation tools is also necessary. This may lead to a new type of professional activity in the cinema industry. Just as 3D movies now need “stereographers”, so will new HAV content require “haptographers”. Moreover the development of tools to evaluate the quality of experience and the acceptance of such systems is mandatory. Tackling the challenges of this young but promising field of study will yield new tools and methods for adding haptic content to multimedia, leading to a more compelling user experience in combination with audiovisual content.

REFERENCES

[1] M. Heilig, “Sensorama simulator,” Aug. 1962. U.S. Patent 3050870.

[2] G. C. Burdea and P. Coiffet, *Virtual reality technology*. Wiley-IEEE Press, 2003.

[3] S. J. Biggs and M. Srinivasan, “Haptic interfaces,” *Handbook of virtual Environments*, pp. 93 – 116, 2002.

[4] V. Hayward, O. R. Astley, M. Cruz-Hernandez, D. Grant, and G. Robles-De-La-Torre, “Haptic interfaces and devices,” *Sensor Review*, vol. 24, no. 1, pp. 16–29, 2004.

[5] M. Reiner, “The role of haptics in immersive telecommunication environments,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 392–401, Mar. 2004.

[6] S. O’Modhrain and I. Oakley, “Touch TV: adding feeling to broadcast media,” in *European Conference on Interactive Television: from Viewers to Actors*, 2003.

[7] N. Magnenat-Thalmann and U. Bonanni, “Haptics in virtual reality and multimedia,” *IEEE Multimedia*, vol. 13, no. 3, pp. 6–11, 2006.

[8] A. El Saddik, M. Orozco, M. Eid, and J. Cha, *Haptic technologies - bringing Touch to multimedia*. Springer, 2011.

[9] T. Yamaguchi, A. Akabane, J. Murayama, and M. Sato, “Automatic generation of haptic effect into published 2D graphics,” in *Eurohaptics*, 2006.

[10] J. Cha, M. Eid, and A. E. Saddik, “Touchable 3D video system,” *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 5, pp. 1–25, Oct. 2009.

[11] J. Cha, Y.-S. Ho, Y. Kim, J. Ryu, and I. Oakley, “A framework for haptic broadcasting,” *IEEE Multimedia*, vol. 16, pp. 16–27, July 2009.

[12] K. Kilkki, “Quality of experience in communications ecosystem,” *Journal of universal computer science*, vol. 14, no. 5, pp. 615–624, 2008.

[13] J. Dionisio, “Projects in VR virtual hell: a trip through the flames,” *IEEE Computer Graphics And Applications*, pp. 11–14, 1997.

[14] S. Palan, R. Wang, N. Naukam, L. Edward, and K. J. Kuchenbecker, “Tactile Gaming Vest (TGV),” 2010. http://iroboticist.com/2010/03/26/tgv/, accessed on September 2012.

[15] B. Lee, J. Lee, J. Cha, C. Seo, and J. Ryu, “Immersive live sports experience with vibrotactile sensation,” *Human-Computer Interaction*, pp. 1042–1045, 2005.

[16] S. ur Rehman, “Turn your mobile into the ball: rendering live football game using vibration,” *IEEE Transactions on Multimedia*, vol. 10, pp. 1022–1033, Oct. 2008.

[17] P. Lemmens, F. Crompvoets, D. Brokken, J. V. D. Eerenbeemd, and G.-j. D. Vries, “A body – conforming tactile jacket to enrich movie viewing,” *Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 7–12, 2009.

[18] M. Waltl, *Enriching multimedia with sensory effects: annotation and simulation tools for the representation of sensory effects*. VDM Verlag Saarbrücken, Germany, 2010.

[19] M. A. Rahman, A. Alkhaldi, and J. Cha, “Adding haptic feature to YouTube,” *International conference on Multimedia*, pp. 1643–1646, 2010.

[20] Y. Kim, J. Cha, J. Ryu, and I. Oakley, “A tactile glove design and authoring system for immersive multimedia,” *IEEE Multimedia*, vol. 17, no. 3, pp. 34–45, 2010.

[21] Y. Suzuki, M. Kobayashi, and A. J. Interface, “Air jet driven force feedback in virtual reality,” *IEEE computer graphics and applications*, vol. 25, no. 1, pp. 44–7, 2005.

[22] T. Hoshi, M. Takahashi, T. Iwamoto, and H. Shinoda, “Non-contact tactile display based on radiation pressure of airborne ultrasound,” *IEEE Transactions on Haptics*, vol. 3, no. 3, pp. 155–165, 2010.

[23] J. Alexander, M. T. Marshall, and S. Subramanian, “Adding haptic feedback to mobile tv,” in *ACM Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1975–1980, 2011.

[24] D. Gaw, D. Morris, and K. Salisbury, “Haptically annotated movies: reaching out and touching the silver screen,” *Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 287–288, 2006.

[25] Y. Kim, S. Park, H. Kim, H. Jeong, and J. Ryu, “Effects of different haptic modalities on students’ understanding of physical phenomena,” in *IEEE World Haptics Conference*, pp. 379–384, 2011.

[26] S. O’Modhrain and I. Oakley, “Adding interactivity: active touch in broadcast media,” in *IEEE Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 293–294, 2004.

[27] F. Danieau, J. Fleureau, A. Cabec, P. Kerbiriou, P. Guillotel, N. Mollet, M. Christie, and A. Lécuyer, “A framework for enhancing video viewing experience with haptic effects of motion,” in *IEEE Haptics Symposium*, pp. 541 – 546, 2012.

[28] K. Shimoga, “A survey of perceptual feedback issues in dexterous telemanipulation. II. Finger touch feedback,” in *IEEE Virtual Reality Annual International Symposium*, pp. 271–279, IEEE, 1993.

[29] L. A. Jones and M. Berris, “The psychophysics of temperature perception and thermal-interface design,” in *IEEE Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 137–142, 2002.

[30] L. A. Jones, “Kinesthetic sensing,” *Human and Machine Haptics*, pp. 1–10, Mar. 2000.

[31] A. Berthoz, *The brain’s sense of movement*. Harward University Press, 2000.

[32] J. B. van Erp, “Guidelines for the use of vibro-tactile displays in human computer interaction,” in *Eurohaptics*, vol. 2002, pp. 18–22, 2002.

[33] K. Hale and K. Stanney, “Deriving haptic design guidelines from human physiological, psychophysical, and neurological foundations,” *IEEE Computer Graphics and Applications*, vol. 24, no. 2, pp. 33–39, 2004.

[34] J. Smith and K. MacLean, "Communicating emotion through a haptic link: design space and methodology," *International Journal of Human-Computer Studies*, vol. 65, no. 4, pp. 376–387, 2007.

[35] D. Tsetserukou, A. Neviarouskaya, H. Prendinger, N. Kawakami, and S. Tachi, "Affective haptics in emotional communication," in *IEEE International Conference on Affective Computing and Intelligent Interaction and Workshops*, no. 1, pp. 1–6, 2009.

[36] A. Brady, B. MacDonald, I. Oakley, S. Hughes, and S. O'Modhrain, "Relay: a futuristic interface for remote driving," in *Eurohaptics*, 2002.

[37] K. Kuchenbecker, J. Fiene, and G. Niemeyer, "Event-based haptics and acceleration matching: portraying and assessing the realism of contact," *Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 381–387, 2005.

[38] T. L. McDaniel, K. Kahol, P. Tripathi, D. P. Smith, L. Bratton, R. Atreya, and S. Panchanathan, "A methodology to establish ground truth for computer vision algorithms to estimate haptic features from visual images," *IEEE International Workshop on Haptic Audio Visual Environments and their Applications*, pp. 94–99, 2005.

[39] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: a survey," *Foundations and Trends in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177–280, 2007.

[40] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003.

[41] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 34, no. 3, pp. 334–352, 2004.

[42] B. T. Truong and S. Venkatesh, "Video abstraction: a systematic review and classification," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 3, no. 1, 2007.

[43] X. Zhu, X. Wu, A. K. Elmagarmid, Z. Feng, and L. Wu, "Video data mining: semantic indexing and event detection from the association perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 5, pp. 665–677, 2005.

[44] P. Bach-y Rita and S. W. Kercel, "Sensory substitution and the human-machine interface," *Trends in Cognitive Sciences*, vol. 7, no. 12, pp. 541–546, 2003.

[45] A. Chang and C. O'Sullivan, "Audio-haptic feedback in mobile phones," *Extended Abstracts of the Computer Human Interaction Conference on Human factors in Computing Systems*, pp. 1264–1267, 2005.

[46] S. Nanayakkara, E. Taylor, L. Wyse, and S. Ong, "An enhanced musical experience for the deaf: design and evaluation of a music display and a haptic chair," in *ACM International conference on Human factors in computing systems*, pp. 337–346, 2009.

[47] T. Zhang and C.-C. Kuo, "Audio content analysis for online audiovisual data segmentation and classification," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 4, pp. 441–457, 2001.

[48] S. Chang, T. Sikora, and A. Purl, "Overview of the MPEG-7 standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 688–695, 2002.

[49] M. Enriquez and K. MacLean, "The hapticon editor: a tool in support of haptic communication research," *Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 356–362, 2003.

[50] J. Ryu and S. Choi, "posVibEditor: graphical authoring tool of vibrotactile patterns," *IEEE International Workshop on Haptic Audio visual Environments and Games*, pp. 120–125, Oct. 2008.

[51] Y. Seo, B.-c. Lee, Y. Kim, J.-p. Kim, and J. Ryu, "K-HapticModeler: a haptic modeling scope and basic framework," *International Workshop on Haptic Audio Visual Environments and their Applications*, pp. 136 – 141, 2007.

[52] M. Eid, S. Andrews, A. Alamri, and A. El Saddik, "HAMLAT: a HAML-based authoring tool for haptic application development," *Haptics: Perception, Devices and Scenarios*, pp. 857–866, 2008.

[53] K. Yoon, B. Choi, E. Lee, and T. Lim, "4-D broadcasting with MPEG-V," in *IEEE Multimedia Signal Processing*, no. 2, pp. 257–262, 2010.

[54] F. El-far, M. Eid, M. Orozco, and A. Saddik, "Haptic Applications Meta-Language," *IEEE International Symposium on Distributed Simulation and Real-Time Applications*, pp. 261–264, 2006.

[55] A. Sourin and L. Wei, "Visual immersive haptic rendering on the web," *ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, p. 1, 2008.

[56] Y. Kim, H. Shin, and M. Hahn, "A bidirectional haptic communication framework and an authoring tool for an instant messenger," in *IEEE Advanced Communication Technology*, vol. 3, pp. 2050–2053, 2009.

[57] MPEG Group, "ISO/IEC FDIS 23005-3 Information technology – media context and control – Part 3: sensory information." 2011.

[58] ISO/IEC 14496-11, "Information technology – coding of audio-visual objects – Part 11: scene description and application engine," 2005.

[59] D. Wu, Y. T. Hou, W. Zhu, Y.-Q. Zhang, and J. M. Peha, "Streaming video over the Internet: approaches and directions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 282–300, 2001.

[60] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: an historical survey," *Automatica*, vol. 42, no. 12, pp. 2035–2057, 2006.

[61] E. Dijk, A. Weffers-Albu, and T. De Zeeuw, "A tactile actuation blanket to intensify movie experiences with personalised tactile effects," in *International Conference on Intelligent Technologies for Interactive Entertainment*, pp. 2–3, 2010.

[62] A. Israr and I. Poupyrev, "Tactile brush: drawing on skin with a tactile grid display," *Computer Human Interaction*, pp. 2019–2028, 2011.

[63] R. Lindeman, Y. Yanagida, H. Noma, and K. Hosaka, "Wearable vibrotactile systems for virtual contact and information display," *Virtual Reality*, vol. 9, no. 2, pp. 203–213, 2006.

[64] R. Jain, "Quality of experience," *IEEE Multimedia*, vol. 65, no. 1, pp. 1–2, 2004.

[65] A. Hamam, M. Eid, A. E. Saddik, and N. D. Georganas, "A quality of experience model for haptic user interfaces," in *Proceedings of the 2008 Ambi-Sys workshop on Haptic user interfaces in ambient media systems*, pp. 1–6, 2008.

[66] M. Waltl and C. Timmerer, "Increasing the user experience of multimedia presentations with sensory effects," in *11th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, pp. 1–4, 2010.

[67] A. Hamam, N. Georganas, and A. El Saddik, "Effect of haptics on the quality of experience," in *IEEE International Symposium on Haptic Audio-Visual Environments and Games (HAVE)*, pp. 1–6, IEEE, 2010.

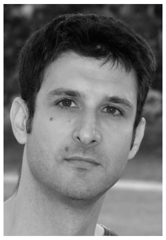
[68] M. Meehan, B. Insko, M. Whitton, and F. P. J. Brooks, "Physiological measures of presence in stressful virtual environments," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 645–652, 2002.

[69] M. Schuemie, P. Van Der Straaten, M. Krijn, and C. Van Der Mast, "Research on presence in virtual reality: a survey," *CyberPsychology & Behavior*, vol. 4, no. 2, pp. 183–201, 2001.

[70] R. L. Mandryk, K. M. Inkpen, and T. W. Calvert, "Using psychophysiological techniques to measure user experience with entertainment technologies," *Behavior & Information Technology*, vol. 25, no. 2, pp. 141–158, 2006.

[71] J. Fleureau, P. Guillotel, and Q. Huynh-Thu, "Physiological-based affect event detector for entertainment video applications.," *IEEE Transactions on Affective Computing*, 2012.

Fabien Danieau is a Ph.D student at Technicolor R&I (France) and in the INRIA VR4I Team (France). His research interests include haptic interaction, multimedia and user experience. In 2010, he received a engineering degree in Cognitics from the Ecole Nationale Supérieure de Cognitique (France).



Anatole Lécuyer is a Senior Researcher at the French National Institute for Research in Computer Science and Control (INRIA), Rennes, France. His main research interests include virtual reality, haptic interaction, and brain-computer interfaces. He has been involved in many collaborative projects related to haptics such as European project "NIW" (08-11) on Augmented Walking, and the European Network of Excellence's "INTUITION" (05-08) project on Virtual Reality in which he led the Working Group on "Haptic interaction". Anatole Lécuyer serves as expert in VR and haptics for public bodies such as the European Commission (EC), the European Research Council (ERC) and the French National Research Agency (ANR). He is secretary of the French Association for Virtual Reality and former secretary of the IEEE Technical Committee on Haptics. He is associate editor of ACM Transactions on Applied Perception (ACM TAP) and the International Journal of Human-Computer Studies (IJHCS).

Philippe Guillotel is a Distinguished Scientist at Technicolor R&I (France), in charge of Video Perception, Processing & Coding research. He received a M.S. degree in electrical engineering and information theory from University of Rennes (France) in 1986, and a degree in electrical engineering, telecommunications and computer science from the Ecole Nationale Supérieure des Télécommunications (France) in 1988. He has worked in the Media & Broadcast Industry for more than 20 years. In addition to his work on video coding systems and compression algorithms, his work at Technicolor has covered video processing, video delivery as well as human perception and video quality assessment studies. He is the author of more than 50 patents and 30 papers.

Julien Fleureau is a Researcher at Technicolor R&I (France) in the Video Processing & Perception laboratory. He received engineering and M.S. degrees in electrical engineering and computer science from the Ecole Centrale de Nantes (France) in 2005, and a Ph.D. degree from the University of Rennes (France) in 2008. His Ph.D studies in image processing and biomedical modeling were crowned by a 2 year post doctoral fellowship at the Laboratoire de Traitement du Signal et de l'Image (INSERM, France). His main research interests are signal processing, datamining and machine learning applied to biomedical and physiological signals analysis.

Nicolas Mollet is a Researcher at Technicolor R&I (France). His scientific interests focus on human-machine interactions, multi-entities and informed environment approaches. He defended his PhD in 2005 (INRIA Rennes). In it he proposed a full authoring environment of virtual maintenance training sessions on military equipment, based on dynamic and behavioral models of objects and scenarios. In 2007 he joined the IIT laboratory (Genova, Italy), where he led a team working on virtual reality applied to multi-robots, through distributed information and behavioral models. He was involved in the European Network of Excellence "INTUITION" project (05-08), with a focus on virtual reality for training. He is a founding member of the European Association of Virtual Reality (EuroVR). He joined Technicolor In 2011, where he is currently leading activities on Augmented Reality, Immersion, and Presence, aiming to create new Technicolor applications and services.

Marc Christie is an Associate Professor at Rennes 1 University and a member of the INRIA Mimetic research team. His main center of interest is Interactive Virtual Cinematography. He defended his Ph.D in 2003 (Nantes University) in which he proposed a declarative approach to camera control using interval-based constraint solving techniques and explored aspects related to camera path planning. He contributed to publications in both computer graphics (Eurographics, SCA) and constraint solving (CP series). With Patrick Olivier (Newcastle University), he authored the first state of the art report on the topic (Eurographics 2006), organized a session on camera control at the Smartgraphics 2008 conference, and presented a course on the topic at Siggraph Asia 2009. Marc Christie is currently working on visibility computation, screen composition, editing and interactive virtual and real cinematography systems.