Feature Article

A Framework for Haptic Broadcasting

Jongeun Cha, Yo-Sung Ho, Yeongmi Kim, and Jeha Ryu
Gwangju Institute of Science and Technology

Ian Oakley
University of Madeira

This article presents a comprehensive exploration of the issues underlying haptic multimedia broadcasting. It also describes the implementation of a prototype system as a proof of concept.

The entertainment domain, particularly content broadcasting, develops rapidly with the early adoption of new technologies. A central theme of these developments is the creation and widespread deployment of more realistic or immersive display systems. Consumers are eager to experience engrossing content capable of blurring the boundaries between fiction and reality; they actively seek an engaging feeling of being there, known as *presence*. In virtual environments, one factor that contributes to increased feelings of immersion is the sense of touch.[1] In this article, we discuss how such haptic cues can contribute to the sensation of presence in the context of a broadcast scenario.

We develop two themes in this article. The first theme revolves around the idea that touch cues can increase immersion simply through the provision of additional passively experienced information about a scene. This case is exemplified by the popularity of motion-based rides in amusement parks and the motivations underlying the Percepto system for tactile simulation, which appeared in movie theaters in the 1950s.[2] The second theme relates to interaction, an important aspect of presence, which is largely overlooked in the domain of broadcasting. When viewers have the ability to interact naturally with an environment, or are able to affect and be affected by environmental stimuli, they are likely to become more immersed in that environment.[3]

While several interactive technologies have been applied to broadcast scenarios, these technologies have focused on functionality, such as display of additional (often textual) information related to several factors: the content; the Internet connection; and the participation activity, such as polls.[4] These services are abstract, information-orientated, and provide an indirect form of interaction with the content; they are unlikely to raise levels of presence. This article presents the case that presence can be boosted more effectively by relying on direct-interaction paradigms that are based on haptic feedback.

Indeed, researchers have argued that as the human haptic system uniquely encompasses both perception and action, for many users touch interaction has a fundamental role in the creation of truly immersive virtual experiences.[1] However, broadcasting is restricted by distinct constraints (especially in comparison to those restricting the display of general virtual environments) and has received scant, but emerging,[5] attention in the haptic literature. In light of these issues, this article focuses on the construction of a haptic-enabled broadcasting system and presents a comprehensive exploration of the issues underlying haptic multimedia broadcasting as they relate to the processes of content creation, transmission, viewing, and interaction. The article also describes the implementation of a prototype system as a proof of concept and discusses two possible scenarios in which this system could be used.

In terms of the technical aspect of this multimedia framework, we use MPEG-4, an IEC/ISO standard for streaming multimedia objects in broadcast-specific applications. We incorporate haptic data into the framework by extending the MPEG-4 Binary Format for Scene (BIFS).[6]

Potential scenarios in haptic broadcasting

In contrast to the user experience of virtual environments, users generally experience broadcast programs linearly: these programs have a beginning, middle, and end and are experienced only in that order. Although interactive narrative is an active research area, it's not yet close to transitioning to commercial broadcast systems, and currently viewers have little opportunity to reorder the shows they watch.

1070-986X/09/$25.00

Published by the IEEE Computer Society

Figure 1. Potential scenarios in haptic broadcasting: passive haptic playback such as (a) learning handwriting and (b) feeling the punch between boxers; and active haptic interaction such as (c) a touching scene and (d) manipulating an object in a scene.

The broadcast domain is therefore distinct from, for example, virtual reality applications or computer games, which can present users with a wide range of possibilities and respond richly to user decisions. So, in this effectively choiceless context, any research designed to add an interactive modality at the perceptual level must address the fundamental question of how interactivity can be achieved and represented.[7] If a viewer cannot make choices, what form of interaction can he or she engage in?

One possibility is to allow viewers to experience and interact with the presentation of displayed scenes but not their sequence.[8] This could be accomplished in several ways. In a passive haptic-playback scenario, for example, no direct interaction would take place and prerecorded haptic cues associated with the content could be displayed to the viewers in much the same way as audio and video are played and shown to listeners and viewers. For example, Figure 1a depicts an educational program in which a calligraphy expert is demonstrating her skills. She asks viewers to grip the pen-like, kinesthetic, force-feedback device and starts to write a character. The viewers are physically guided through the path taken by the expert as they sketch the character.[9] In addition, Figure 1b shows how a boxing event could be augmented with a representation of the impact of a punch. We could measure the impact's magnitude with an accelerometer embedded in the glove, and viewers could feel the intensity of the punch via vibrating motors embedded in their seat.

Passive haptic-playback scenarios involve either recording some aspects of haptic data when capturing audiovisual scenes—using the haptic equivalent of a microphone in the form of sensors, such as accelerometers or advanced 3D scanners[10]—or creating it by hand in a postproduction phase, and then presenting these stimuli to viewers. In these scenarios, a kinesthetic haptic device, such as SensAble's Phantom (see http://www.sensable.com) or Immersion's CyberForce glove system (see http://www.immersion.com), could be held or worn on the viewer's hand or arm to provide recorded movements. Alternatively, an array of tactile stimulators could be worn or even embedded in the furniture to deliver stimuli to the user's skin. In both cases, users wouldn't need to make an explicit action to experience the stimuli; contact with the display equipment would be sufficient.

On the other hand, if a scene contained 3D information, an active haptic interaction scenario in which viewers could actively explore the contours and surfaces of the displayed objects and manipulate them, would be enabled. For example, Figure 1c shows a program in which viewers can touch an actress's face as another onscreen character is doing so. In this way, viewers might identify more strongly with the program and feel as if they have become the actor. Technologies such as 3DV Systems' Z-Cam (see http://www.3dvsystems.com), a combined video and depth camera, enable capturing this type of live 3D scene. A second possible active scenario, shown in Figure 1d, is to let viewers explore the feel of items on a home shopping channel. This scenario would include not only examining the surface properties of, for instance, clothes, but also exploring the dynamic feel of electronic devices by manipulating the various controls.

Active haptic-interaction scenarios involve the capture and display of either a 2.5D scene or full 3D mesh of the objects of interest and their haptic properties. To experience this content most effectively, viewers must use a combination of a kinesthetic device and a tactile array. Such a kinesthetic device would have an enhanced finger-tip-sized tactile array on its end-effector and would support the active exploration of a scene with rich information about the reaction forces felt and textures encountered. This type of scenario is clearly more demanding, both for the authors who are creating the content and for the viewers who are experiencing it.

This distinction between active and passive scenarios links back to fundamental work on haptic perception,[7] which draws a related distinction between the act of feeling something, which is an active exploratory process, and the experience of being touched, which is a passive occurrence. However, mapping these concepts onto the broadcasting domain has led to the blurring of the lines between these theoretical categories. This is made clear in the passive haptic-playback scenarios described previously. An entirely prerecorded sequence of tactile cues presented to the skin represents a purely passive tactile experience. But when it comes to movements a kinesthetic device constrained to follow a pen's path, a user is more free to explore the motions in any way they choose. Such exploration might include holding the device loosely or tightly, for example. Although none of these actions would alter the nature of the cues delivered to and by the mechanics of the haptic device, they would clearly alter the user's perceptual experience; in this sense, such a system inherently creates active perceptual experiences.

In light of these observations, we use the terms *passive haptic playback* and *active haptic interaction* in this article in terms of the type of application-level behavior the device supports. In short, passive haptic playback signifies the delivery of prerecorded streams of information, while active haptic interaction involves scenes that a user can choose to explore.

It's worth noting that the precise definition of the term *haptic* remains in some dispute.[11] This article follows the majority of the literature in computer science (for example, see Biggs and Srinivasan[12]) and uses *haptic* as an umbrella term for two subcategories of feedback: tactile and kinesthetic. Tactile refers to information sensed through nerve receptors in the skin, while kinesthetic refers to information sensed through movement and force applied to the muscles and joints. A haptic device can include a tactile output component, a kinesthetic output component, or, in the most advanced configuration, both. Displaying haptic information is an emerging field, and while progress is being made steadily, many issues relating to the machine display and human perception of haptic cues remain active research topics.[13]

Haptic media

A media format defines a media type's structure—that is, the data type and its internal representation details. In the context of a broadcast environment, the media format describes not only the content displayed to a viewer, but also how the content is arranged in time, and, by extension, related to associated media forms. For example, audiovisual media is composed of two distinct channels, often created, edited, and encoded at separate times and using separate methods. Video might be captured in a location shoot, while sounds are recorded in a studio or extracted from a library. These two data sources must be composed coherently to produce meaningful content. The role of a media format is to technically enable this composition. Audiovisual media formats are well-established and deal well with the different update rates and the demands placed by streaming over network links. However, there is little work that explicitly addresses haptic media.

While there have been many virtual environment systems and APIs that have included representations for haptic information of various sorts—such as the Reachin API (see http://www.reachin.se) and H3D (see http://www.h3d.org)—they have generally described only static scene information. It's far less common to see examples where haptic information includes a temporal component. Those that exist tend to have a dedicated, single purpose and are simplistic in their implementation. For instance, in the domain of telehaptics, users experience a stream of haptic cues in the form of force data delivered sequentially to a haptic device.[14] Similarly, in haptic training, a novice sequentially receives a stream of force or position data recorded by an expert.[15] Using a similar paradigm, access technologies, such as a tactile vision substitution system, provide systems that map video input from cameras to patterns of tactile activation in a grid of vibrating elements. These systems all share the same basic approach: haptic information is captured, undergoes some immediate transformation to match a display device's specific characteristics, and then is rendered immediately in synch with other media or stored to disk as a stream. While these approaches are useful within their particular application area, they don't constitute a general approach because they are tied to high quality of service communication links and specific input and output technologies. There are clearly many forms of haptic data that these systems can't express.

Because our goal with this article is describing a system that conveys rich and general-purpose haptic-media elements, we need to define exactly what the term *haptic media* encompasses. Given the varied, and frequently incompatible, nature of available haptic technologies, defining *haptic media* is complicated because the definition must be grounded in a methodology for expressing sensations applicable to a broad range of devices. To help clarify our use of the term, we make a distinction between two key media categories: linear and nonlinear haptic media. We relate each of these categories to sensation types that commonly available haptic hardware can display.

Linear haptic media refers to haptic sensations that progress sequentially in time, such as displayed movements or cutaneous patterns of touches that create experiences of passive haptic playback. Essentially, linear haptic media encompasses events such as human touches, impacts, alarms, direction cues, textures, and onscreen motions (such as the bouncing of a ball). These sensation types are best displayed by technologies that provide an array of tactile elements distributed over an area of the skin. By varying which elements are activated over time, a system can create a rich range of sensations. Several technologies, including small mechanical motors, transducers, piezoelectric strips, air jets, and direct application of electric currents, can help create these sensations. This article defines one key type of the linear haptic media to be a sequential series of actuation intensities deployed to control a grid of tactile stimulators spread over an area of skin. For the best effect, these intensities must correspond directly to audio and video events in the content.

Nonlinear haptic media, on the other hand, offers viewers interactivity and allows them to touch a haptically displayed object and experience force and texture information that produces a compelling, active, haptic interaction. In this case, viewers feel objects by exerting effort: they must explore the environment to feel the haptic cues. These cues can take the form not only of force and shape information but also of the surface properties of objects' texture: friction, roughness, and stiffness. For greater realism, nonlinear haptic media must encompass object dynamics, such as how they move and behave in response to user input. Object dynamics can include general factors such as mass and inertia, as well as more specific factors, such as spring constants defining the travel distance and sponginess of a virtual button. Such nonlinear haptic media is arguably best displayed on commercially available kinesthetic devices, such as Phantom devices. This article defines this media as including a complete description not only of the shape of all objects, but also of their surface properties and dynamic behavior.

Haptic broadcasting framework

Incorporating haptics into the broadcasting pipeline entails more than the development of a method and instrument enabling the integration of haptic effects with traditional audiovisual media. We therefore present a comprehensive exploration of the issues underlying haptic multimedia broadcasting, from the processes of content creation through transmission and finally to viewing and interaction.

Technological platform

Haptic sensation is well established in, for example, virtual environment research. In such systems, a typical haptic application implements all necessary software elements for the delivery of the haptic sensation: it drives the haptic device, represents and manipulates the haptic content, and deals with user interaction. In broadcast and many multimedia scenarios, there is a clear distinction between the content and the multimedia player. The two are independent; a player interprets and displays the content. However, previous research focused on incorporating haptics into multimedia applications with limited haptic interaction modes specific to particular systems and scenarios.[8,10] This research reflects the lack of a systematic framework to deal with multimedia data sets that incorporate haptic information.

Researchers have made several attempts to construct a multimedia framework supporting the easy creation and distribution of haptic applications. For example, the Reachin API provides a fully extensible programming framework for building haptic interactive applications based on the Virtual Reality Modeling Language. This approach allows the creation of virtual worlds featuring a range of media types, including video, audio, graphics, and haptics that can be sent via the Internet. H3D supports a similar framework based on X3D. However, these frameworks are not appropriate for broadcasting because they are based on a simple download-and-play delivery system rather than one that has the ability to handle media-data streams. In essence, users of these existing systems must download entire media clips before viewing them, a potentially lengthy process that is in direct opposition to the immediately available streaming-content paradigm typical in broadcast scenarios.

To achieve haptic-streaming media, we use the MPEG-4 framework, which supports not only streaming data for a range of media objects, but also flexible interactivity designed for broadcast specific applications. A key difference in MPEG-4, when compared to prior audiovisual standards, is its object-based, audiovisual representation model. For example, an object describing an animated moving head can encode that movement using mathematical parameters, while a coincidently displayed video scene can remain composed of adaptive pixel values. MPEG-4 also supports the harmonious integration of these varied data types, providing a unifying system that enables novel feats, such as the seamless interaction between a cartoon character and an actor in a studio. This flexibility makes MPEG-4 an ideal technology for supporting haptic broadcasting. It allows for encoding and decoding of haptic media as independent objects that can be synthesized and synchronized with other audiovisual media.

Technically, haptic media can be described using MPEG-4 BIFS, which allows it to be spatiotemporally coordinated with other audiovisual media in a scene. At the most basic level, BIFS describes interactive 3D objects and worlds. However, it also provides a crucial feature for haptic interaction: an update mechanism. A BIFS scene can be updated in many ways. For example, new objects can be added and existing objects modified, deleted, or replaced. This mechanism can transform a static binary scene into a data stream that can be sent over a network and synchronized with other streams (such as, video, audio, and metadata). Consequently, BIFS theoretically supports the transformation of haptic media into a stream synchronized with other audiovisual media streams. However, the current MPEG-4 specifications don't cover haptic media, an omission we address in this article.

Due to the range and complexity of haptic information, a haptic broadcasting system needs to support a wide variety of data types. These include the haptic surface properties of stiffness, friction, and roughness (sufficient to represent many haptic experiences) along with the dynamic properties of movement stiffness, inertia, direction, and travel distance. These data types encapsulate the behavior of virtual controls, such as buttons, sliders, or joypads, that might appear as part of virtual object models. The system must also include tactile video in the form of a gray-scale video component. This format represents a grid of intensities of tactile stimulation that can be mapped to an area of the skin over time and rendered through devices, such as tactile arrays composed of a grid of vibrating elements.[16] Figure 2 shows how such a 2D set of intensities might be represented, albeit with some shape distortion, on a glove-based tactile array. Finally, two aspects of movement data not described in this article are poses, such as positions and orientations of objects or bodies, and their related forces and torques.

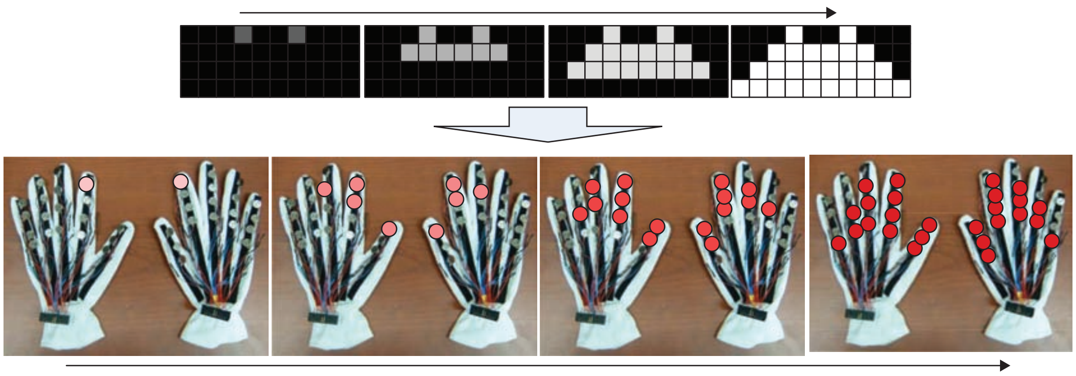


Figure 2. Illustration of mapping tactile video to glove-based tactile device. The top portion of the figure shows four frames in a grid format and the bottom portion shows the glove format. Darker red circles on the gloves indicate active tactors with higher intensity.

Figure 3 shows the complete diagram of a haptic broadcasting system designed to support this range of data types based on the MPEG-4 framework. It depicts the entire process from content creation through consumption. Content creation results in the generation of media types that require different representations and must be seamlessly merged. In MPEG-4, linear media of any modality is typically assigned a dedicated stream due to its large size. These streams are compressed with suitable domain-specific encoders (such as MP3 for audio) and transmitted progressively on the basis of information relating to synchronization and presentation. Tactile video corresponds to the linear media type. Due to its relatively small size, nonlinear media, such as scalar values in a virtual model, can be encoded so that they can be transmitted immediately and stored locally using MPEG-4 BIFS. Haptic surface properties and dynamic properties correspond to the nonlinear media type.

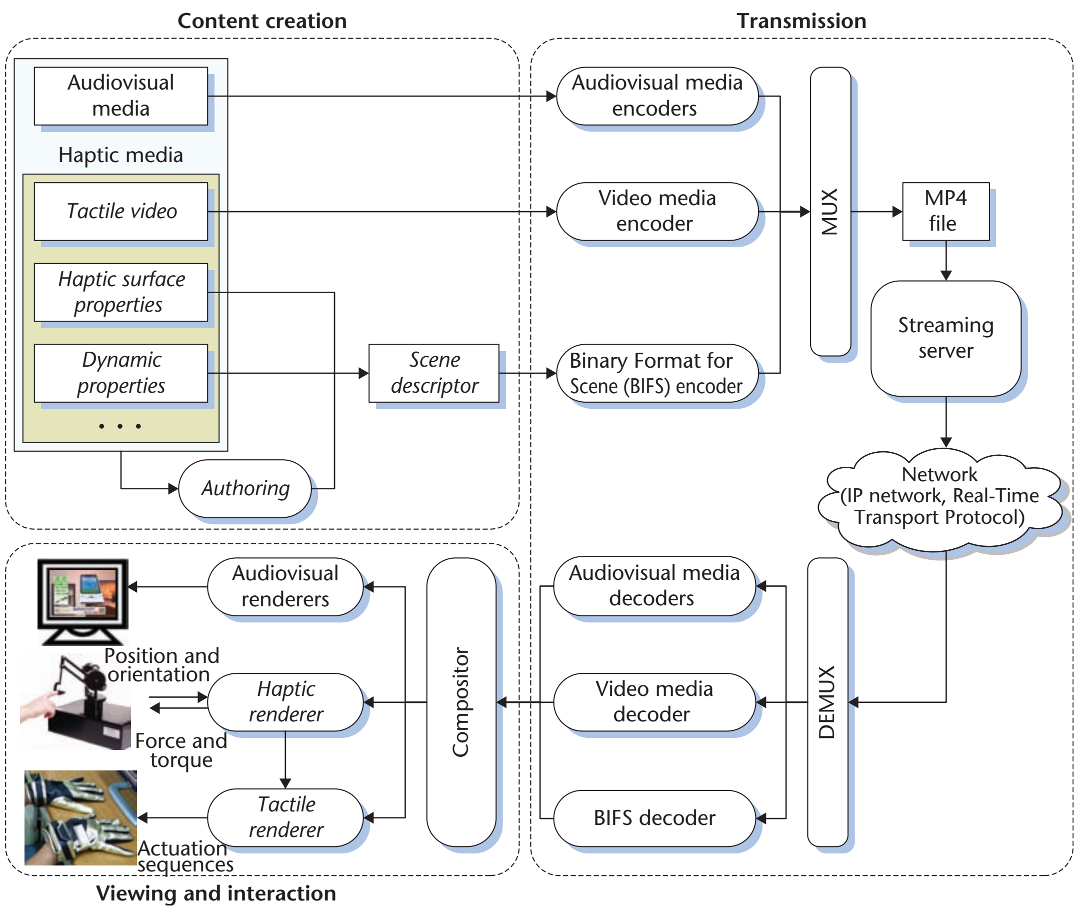


Figure 3. Diagram of the haptic broadcasting system based on standard MPEG-4 framework. The authors' extensions to this standard diagram are noted in italicized text.

Larger nonlinear media, such as bump-map textures describing roughness, are typically encoded in a unique stream but transmitted immediately. During MPEG-4 data transmission, the sender multiplexes all this information into a single stream, which is sent through a supported network (for example, Real-Time Transport Protocol or MPEG-2 Transport Stream) and then demultiplexed by the receiver. During viewing and interaction, a compositor parses and forms the scene by sending all media elements to relevant video, audio, or haptic renderers and display devices.

Content creation

While standard tools exist for capturing audio and video media, it's less obvious how the same objective might be achieved for haptic media. Consequently, this article provides an overview of the mechanisms by which haptic media can be created. There are three main approaches: data can be recorded using physical sensors, generated using specialized modeling tools, and derived automatically from analysis of other associated media.

There are a few studies on the automatic capture of haptic surface properties, such as stiffness, friction, and roughness.[17] In addition, researchers have acquired the dynamic properties of haptic buttons by measuring and analyzing the force profiles of real physical buttons.[18] We can measure movement data, for instance, with a 3D robotic arm equipped with force-torque sensors or with a motion sensor (such as an accelerometer) embedded inside some object of interest. For example, a boxing glove equipped with such a device could broadcast information about the forces its wielder is experiencing. Moreover, we can capture full or partial 3D information from a scene using 3D scanners to produce detailed 3D meshes of real objects or, for a more rapid approach, using a depth video camera, such as the Zcam, to capture 2.5D information at video frame rates.

Tools to generate audiovisual media in the form of 3D modeling environments are commonplace, but few research projects have attempted to create a modeling tool that can integrate haptic properties into a 3D scene. Two examples are K-HapticModeler[19] and Haptic Application Markup Language Authoring Tool (Hamlat).[20] Both of these tools' interfaces support 3D-scene construction and allow assigning both haptic surface properties and dynamic movement properties to parts of that scene.

Furthermore, a few researchers have developed software tools that play video and simultaneously show and record a user's position to enable tracing movement in scenes (such as orchestral conducting) that involve dynamic human motion.[16] We can play back this spatiotemporal path on a kinesthetic device, effectively providing a trace of the original user's movements in synch with the audiovisual content. With this tool,[16] we can record tactile video through gestures made on touch sensors. such as touch pads or touch screens: as a video plays, a user can input patterns of tactile sensation by making movements on the surface sensor, creating a sequential stream of 2D information. We can use this data to represent patterns of tactile activation and synchronize it with the audiovisual content. However, this tool is relatively basic and there is substantial work to be done in developing haptic editing tools that are as sophisticated as current video and audio equivalents.

For the automatic generation of haptic media from other media sources, some associations are relatively obvious, while others are potentially more rewarding. For example, the trajectory of a soccer ball or the forces exerted on a race car as it corners could be automatically extracted from video or animations using image-processing techniques. Research in this area remains in its infancy and requires further attention before we can reach concrete conclusions about it.

After a system produces haptic media, it requires editing and synchronization, spatially and temporally, with audiovisual media. This stage, essentially one of editing and arranging, fundamentally relies on the media format properties.

Transmission

Though MPEG-4 BIFS is an extensible system for representing media objects, it doesn't include haptic media. New BIFS nodes can help achieve the necessary functionality and support the synchronized representation and transmission of a media stream that includes haptic media. Our approach enables spatiotemporal relationships between haptic and audiovisual media. Figure 4 gives an overview of the new nodes for representing haptic surface properties, dynamic properties, and tactile video along with their encoding method for transmission.

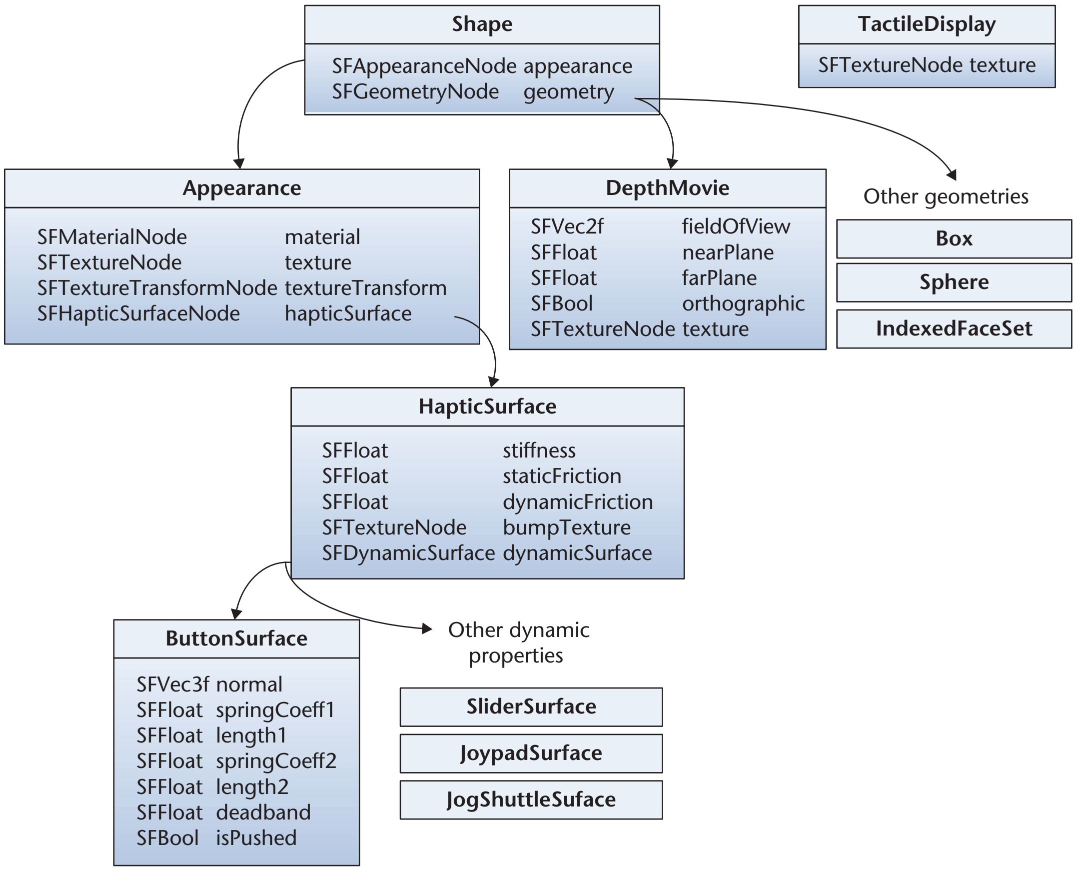


Figure 4. Haptic media nodes in Binary Format for Scene.

Haptic surface properties. The standard MPEG-4 BIFS Shape node represents graphical 3D objects. It makes a distinction between appearance and geometry by implementing individual nodes for each. This is a flexible and efficient approach that enables combining one geometrical form with numerous surface appearances (or vice versa). As can be seen in Figure 4, the Appearance node incorporates several nodes related to graphical texture. To add haptics, we extend it with a new field, hapticSurface. This field can point to the HapticSurface node that contains basic haptic parameters, such as stiffness and friction in the form of scalar values. It also enables a bump-map surface (essentially an array of height values used to perturb the surface of the shape) in the form of a 2D image. These parameters are sufficient to express many key haptic experiences and are similar to those implemented in commercial haptic toolkits, such as the Reachin API. Finally, we include a dynamicSurface field and encode the scalar parameters of haptic surface properties using a standard BIFS encoder for numerical data while the bump-map images rely on any image encoder.

Dynamic properties of haptic widgets. On the basis of other research, we have drawn up several haptic widget nodes intended to represent controls on virtual objects,[21] but in the interests of brevity we only describe one in this article. The ButtonSurface node consists of a basic force profile with three distinct stages that can be encoded with travel distances, spring coefficients, and deadband force parameters.[21] When it's pushed along its vector of travel and exceeds a preset travel distance, it raises a flag and generates an event that can be handled within the BIFS framework and can impact another BIFS object. All fields defined in the ButtonSurface node are small scalar values and can be encoded and compressed using the standard BIFS encoder.

Tactile video. To provide tactile feeling, actuation intensities corresponding to each actuator need to be set in a time line. The actuation intensity can include the amplitude, frequency, or a combination of both of these factors. The actuators themselves can take the form of a range of technologies, such as vibrating motors, electrodes, and piezoelectric ceramics.

We can represent each frame of an actuation intensity array of a rectangular tactile device with an image and represent frame sequences with a video. This tactile video approach is identical in fundamental structure to conventional video media and is implemented in a new BIFS TactileDisplay node. A standard SFTextureNode stores the actual gray-scale tactile video file. We can resolve any resolution differences between the stored video and the local tactile device through basic image resizing operations at viewing time. We can compress this tactile video with a video encoder.

Depth video. All visible objects in a scene are described within the Shape node that contains appearance and geometric information. Depth video consists of sequences of depth images that describe a 2.5D geometric scene. Therefore, it's reasonable to define the depth video node in the Shape node. MPEG-4 BIFS already defines a DepthImage node.[22] However, it's not included in the Shape node but as an optional component of the Transform node. This existing DepthImage node includes a description of camera position and 2.5D textural information. We propose a new DepthMovie node that can be stored in the geometry field and follows the concept of the Shape node. The color video corresponding to the depth video is stored in the Appearance node through a BIFS MovieTexture node. To make the depth video touchable, we need to set the HapticSurface node in the Appearance node. The fields of the DepthMovie node are identical to those of the DepthImage node, with the exception that the texture field stores the video data using a MovieTexture node. Once again, we can compress all video media with any video encoder.

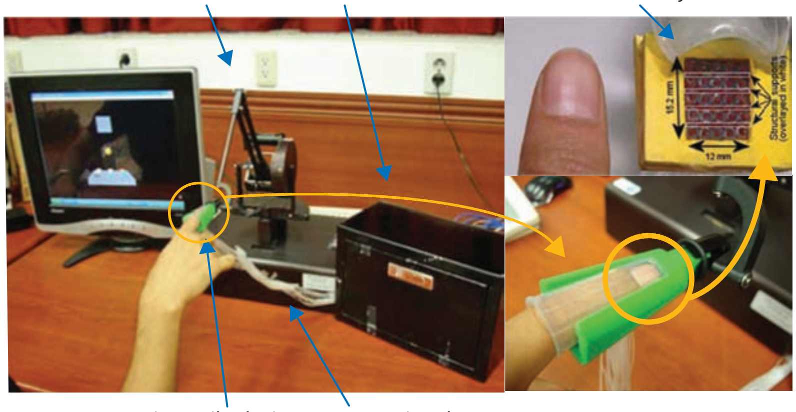


Figure 5. Example of a hybrid haptic device.

MPEG-4 transmission. In the transmitting stage, we compress audiovisual media with suitable encoders. Any suitable image or video encoder compresses the haptic media in image or video format and a BIFS encoder compresses all haptic media (including a scene descriptor if one is present) in scalar value format. All this information is then saved in an MP4 file format,[6] designed to contain the media information of an MPEG-4 presentation in a flexible, extensible format that facilitates interchange, management, editing, and presentation. A streaming server transmits this to viewers. MPEG-4 content can be carried over many different transport layers, including traditional broadcast methods and IP networks, which means we can stream the resulting content through a satellite, over radio and television airwaves, and to Internet nodes. At the viewer's own site, the transmitted content can be decoded and shown.

Viewing and interaction

In the viewing and interaction stage, viewers can enjoy active haptic interaction and passive haptic playback synchronized with an audiovisual scene, as well as have more traditional experiences, such as watching and listening. In this stage, the decoded media is fed to a compositor process that has access to the BIFS scene graph. Traditionally, the compositor scans the scene graph, determines what audio and visual content should be shown, and then passes this information to the audiovisual renderers that actually handle the display of the media on an entertainment device, such as a TV.

Our proposed system extends the compositor process to deal with haptic elements in the scene graph by a similar process of routing them to appropriate renderers. The haptic renderers themselves maintain the current scene's touchable objects and determine what haptic cues to apply. For active haptic interaction, the haptic renderer obtains the viewer's interaction position, computes the interaction force and tactile information generated from the touched objects in the scene, and then transfers those values to a hybrid haptic device incorporating both a kinesthetic and tactile display. Figure 5 shows an example of such a system, which consists of a commercial Phantom, 3-degrees-of-freedom, kinesthetic device combined with a custom fingertip pneumatic tactile device.[23]

The tactile sensation renderer converts tactile information from either the haptic renderer or tactile video (routed through the compositor) into a grid of data corresponding to the size of the local tactile array. It then applies these intensities to the tactile device worn or held by the viewer. Figure 6 shows an example of a tactile device that supports passive, tactile playback on a viewer's hand, a glove-type device incorporating 76 vibrating motors.[16]

We implemented our proposed haptic-broadcasting system using GPAC (see http://gpac.sourceforge.net), a multimedia framework based on the MPEG-4 Systems standard. GPAC supplies basic modules for encoding/decoding and multiplexing/demultiplexing multimedia, including MPEG-4 BIFS, and for playing the transmitted multimedia audiovisually. To stream the content, we use the Darwin Streaming Server (see http://developer.apple.com/opensource/server/streaming). The server allows us to deliver streaming media to clients across the Internet using the Real-Time Streaming Protocol that uses Real-Time Transport Protocol.

To display this data to viewers, we implemented an MPEG-4 content player that shows the audiovisual media and provides haptic interaction and playback based on the Osmo4 player, which is part of the GPAC framework. Osmo4 imports all modules for traditional audiovisual media decoders and displays this data through speakers and a standard visual unit. To enable haptic interactions, we imported a kinesthetic device driver from the manufacturer, created a device driver for the glove-type tactile device,[16] and developed the haptic rendering algorithm. We implemented the algorithm with a conventional, proxy-based solution for 3D mesh models and dynamic haptic user-interface widgets. In addition, we included a combination of a modified proxy graph algorithm[10] and a depth-image-based haptic-rendering module[24] to present the depth video. Because the force can provide surface properties such as stiffness, friction, and roughness without any cutaneous feeling, we used applied-force rendering as a proof-of-concept for the haptic broadcasting system.

Figure 7 shows a scene from a fictional shopping channel created to explore this concept. While the host is enumerating a PDA's functions and features in Figure 7a, the video cues to a close-up of the device in the form of a virtual model in Figure 7b. The host then provides a spoken guide to the viewers about how to touch and manipulate the product using a haptic display. Viewers are able to touch not only the products in the foreground but also items in the background. As can be seen in Figure 7c, viewers can touch the displayed virtual objects

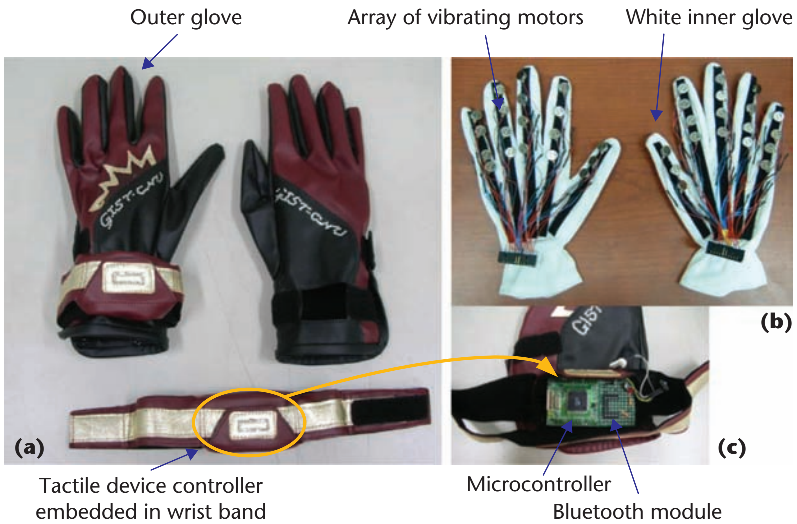
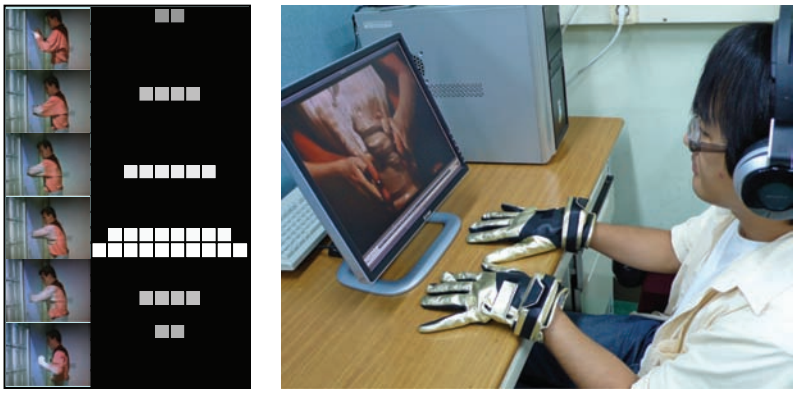


Figure 6. Example of a tactile device for passive tactile playback: (a) completed tactile device packaged with outer glove, (b) vibrating motors on inner gloves, and (c) device controller in the glove's wristband.

Figure 7. Demonstration of the home shopping scenario as an active haptic interaction: while (a) the host discusses a PDA's features, (b) the video cues a close-up of the device, and (c) viewers can interact with the product.



(a) (b)

Figure 8. ***(a) Demonstration of the movie with (b) tactile feeling as the passive haptic playback.***

and manipulate their controls by wearing and moving a thimble attached to the end of a standard kinesthetic device.

In addition to examining the scenario of a home shopping network, we explored the display of passive, tactile playback in the context of a viewer watching a movie. In one scenario, we augmented the movie *Ghost* with manually authored tactile cues produced in a dedicated authoring tool.[16] As shown in Figure 8, we mapped the actions of touching an object (for example, when the ghost's arms penetrate the door) to the tactile display, linking the point of view of the onscreen character to the physical sensations applied to the viewer. Figure 8 illustrates this example and also shows a viewer wearing the glove-type tactile device.

Conclusions and future work

Our proposed definition and implementation are not perfect, requiring more fine-tuning and development for specific networks, such as IPTV. Furthermore, as mentioned previously, there is a great need to develop expressive and easy-to-use editing and authoring tools to support and encourage the widespread creation of haptic media. Finally, further development at the practical and conceptual levels is required to fully explore the scope, advantages, and usability of this emerging technology. For example, there aren't many empirical studies demonstrating the value of haptic technology in broadcasting. Although researchers have examined many theoretical aspects of haptic media in broadcasting, such as the importance of touch for engendering feelings of presence, it's important to conduct studies confirming this value for home entertainment. This article serves as a framework upon which our future research on haptic broadcasting will be grounded. **MM**

Acknowledgments

This research was supported by the Ministry of Knowledge Economy, Korea, under the Information Technology Research Center support program supervised by the Institute of Information Technology Advancement. (grant number IITA-2008-C1090-0801-0017 & IITA-2008-C1090-0804-0002).

References

1. M. Reiner, ''The Role of Haptics in Immersive Telecommunication Environments,'' *IEEE Trans. Circuits and Systems for Video Technology,* vol. 14, no. 3, 2004, pp. 392-401.
2. W.A. IJsselsteijn, ''Presence in the Past: What Can We Learn from Media History?'' *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments,* G. Riva, F. Davide, and W. A. IJsselsteijn, eds., IOS Press, 2003, pp. 17-40.
3. B.G. Witmer and M.J. Singer, ''Measuring Presence in Virtual Environments: A Presence Questionnaire,'' *Presence: Teleoperators and Virtual Environments,* vol. 7, no. 3, 1998, pp. 225-240.
4. M. Bukowska, *Winky Dink Half a Century Later,* tech. report, Media Interaction Group, Philips Research Laboratories, 2001.
5. N. Magnenat-Thalmann and U. Bonanni, ''Haptics in Virtual Reality and Multimedia,'' *IEEE MultiMedia,* vol. 13, no. 3, 2006, pp. 6-11.
6. *Std. JTC1/SC29/WG11 14496-1, Information Technology—Coding of Audiovisual Objects—Part 1: Systems,* ISO/IEC.
7. J.J. Gibson, ''Observations on Active Touch,'' *Psychological Review,* vol. 69, no. 6, 1962, pp. 477-491.
8. S. O'Modhrain and I. Oakley, ''Touch TV: Adding Feeling to Broadcast Media,'' *Proc. European Conf. Interactive Television: from Viewers to Actors,* 2003, pp. 41-47; http://www.brighton.ac.uk/cmis/courses/postgraduate/pgpit/euroitv/euroitv03/Proceedings.htm.
9. B. Plimmer et al., ''Multimodal Collaborative Handwriting Training for Visually-Impaired People,'' *Proc. ACM Conf. Computer–Human Interaction,* ACM Press, 2008, pp. 393-402.
10. J. Cha et al., ''3D Video Player System with Haptic Interaction Based on Depth Image-Based

Representation,'' *IEEE Trans. Consumer Electronics,* vol. 52, no. 2, 2006, pp. 477-484.
11. G. Robles-De-La-Torre, ''The Importance of the Sense of Touch in Virtual and Real Environments,'' *IEEE MultiMedia,* vol. 13, no. 3, 2006, pp. 24-30.
12. J. Biggs and M.A. Srinivasan, ''Haptic Interfaces,'' *Handbook of Virtual Environments,* K. Stanney, ed., Lawrence Erlbaum, 2002, pp. 93-116.
13. V. Hayward et al., ''Haptic Interfaces and Devices,'' *Sensor Review,* vol. 24, no. 1, 2004, pp. 16-29.
14. O. Wongwirat and S. Ohara, ''Haptic Media Synchronization for Remote Surgery through Simulation,'' *IEEE MultiMedia,* vol. 13, no. 3, 2006, pp. 62-69.
15. G. Srimathveeravalli and K. Thenkurussi, ''Motor Skill Training Assistance Using Haptic Attributes,'' *Proc. WorldHaptics,* IEEE CS Press, 2005, pp. 452-457.
16. J. Cha et al., ''An Authoring/Editing Framework for Haptic Broadcasting: Passive Haptic Interactions Using MPEG-4 BIFS,'' *Proc. WorldHaptics,* IEEE CS Press, 2007, pp. 274-279.
17. G.M. Krishna and K. Rajanna, ''Tactile Sensor Based on Piezoelectric Resonance,'' *IEEE Sensors J.,* vol. 4, no. 5, 2004, pp. 691-697.
18. Y. Kim et al., ''Air-Jet Button Effects in AR,'' *Int'l Conf. Artificial Reality and Telexistence,* LNCS 4282, Springer, 2006, pp. 384-391.
19. Y. Seo et al., ''K-HapticModeler: A Haptic Modeling Scope and Basic Framework,'' *Proc. IEEE Int'l Workshop Haptic Audio, and Visual Environments, and their Application,* IEEE Press, 2007, pp. 136-141.
20. M. Eid et al., ''Hamlat: A HAML-Based Authoring Tool for Haptic Application Development,'' *Proc. EuroHaptics,* LNCS 5024, Springer, 2008, pp. 857-766.
21. T. Miller and R. Zeleznik, ''The Design of 3D Haptic Widgets,'' *Proc. ACM Siggraph,* ACM Press, 1999, pp. 97-102.
22. L. Levkovich-Maslyuk et al., ''Depth Image-Based Representation and Compression for Static and Animated 3-D Objects,'' *IEEE Tran. Circuits and Systems for Video Technology,* vol. 14, no. 7, 2004, pp. 1032-1045.
23. Y. Kim, I. Oakley, and J. Ryu, ''Human Perception of Pneumatic Tactile Cues,'' *Advanced Robotics,* special issue on tactile feedback for humanoids and humans, vol. 22, no. 8, 2008, pp. 807-828.
24. J. Cha, M. Eid, and A. El Saddik, ''DIBHR: Depth Image-Based Haptic Rendering,'' *Proc. EuroHaptics,* LNCS 5024, Springer 2008, pp. 640-650.

Jongeun Cha is a postdoctoral fellow at the School of Information Technology and Engineering in University of Ottawa, Canada. His research interests include haptic rendering algorithms, haptic interactions in broadcasting and augmented/mixed reality, haptic content authoring, and interpersonal haptic communication. Cha received a PhD in mechatronics from the Gwangju Institute of Science and Technology (GIST), Korea. Contact him at jcha@discover.uottawa.ca.

Yo-Sung Ho is a professor in the Information and Communications Department at the Gwangju Institute of Science and Technology (GIST), Korea, where he is also the director of Realistic Broadcasting Research Center. His research interests include digital image and video coding, image analysis and image restoration, advanced coding techniques, digital video and audio broadcasting, 3D television, and realistic broadcasting. Ho has a PhD in electrical and computer engineering from the University of California at Santa Barbara. He is a senior member of the IEEE. Contact him at hoyo@gist.ac.kr.

Yeongmi Kim is a PhD student in the Mechatronics Department at the Gwangju Institute of Science and Technology. Her research interests include psychophysics, tactile display design, haptic content authoring, and haptic interaction in broadcasting. Cha has an MS in mechatronics from the Gwangju Institute of Science and Technology. Contact her at kym@gist.ac.kr.

Jeha Ryu is a professor in the Mechatronics Department at the Gwangju Institute of Science and Technology. He is the director of the national Haptics Technology Research Center. His research interests include haptics, haptic rendering, haptic-interaction control, design and control of haptic devices, and haptic-included modeling and broadcasting. Ryu has a PhD in mechanical engineering from the University of Iowa. Contact him at ryu@gist.ac.kr.

Ian Oakley is an assistant professor of Human Computer Interaction at Lab:USE in the University of Madeira, Funchal, Portugal. His research interests include multisensory interfaces. Oakley has a PhD in human–computer interaction and a Joint Honours (First Class) in computing science and psychology from the University of Glasgow. He is a member of the ACM special interest group on computer–human interaction. Contact him at ian@uma.pt.