A Framework for Haptic Broadcasting

Jongeun Cha¹, Ian Oakley², Yo-Sung Ho³, Yeongmi Kim¹, and Jeha Ryu¹

¹ Human-Machine-Computer Interaction Lab., Mechatronics Department, Gwangju Institute of Science and Technology (GIST), Gwangju, Republic of Korea
² Lab:USE, Department of Mathematics and Engineering, University of Madeira, Funchal, Portugal
³ Visual Communications Lab., Information and Communications Department, GIST, Gwangju, Republic of Korea

jcha@discover.uottawa.ca, +1-613-562-5800 x 2148
ian@uma.pt, +351-291705117
hoyo@gist.ac.kr, +82-62-970-2211
kym@gist.ac.kr, +82-62-970-2417
ryu@gist.ac.kr, +82-62-970-2389

Abstract: This article presents a comprehensive exploration of the issues underlying haptic multimedia broadcasting from the processes of contents creation through transmission and finally to viewing and interaction. It also describes the implementation of a prototype system as a proof of concept and demonstrates two possible content scenarios using this system. To achieve this, MPEG-4, an IEC/ISO standard for streaming multimedia objects in broadcast specific applications, is adopted as a multimedia framework and haptic data are incorporated into the framework through extensions of MPEG-4 BIFS (Binary Format for Scene). The implementation of this framework is described and two demonstrations that use it, “Home Shopping” and “A Movie with Tactile Feeling”, are introduced.

Keywords: Haptics, Haptic media broadcasting, MPEG-4
Entertainment, particularly in content broadcasting, is a domain that develops rapidly and is an early adopter of new technologies. A central theme of these advances 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 itself and reality; they actively seek an engaging feeling of “being there” known as presence. In virtual environments, one factor that contributes to increasing feelings of immersion is the sense of touch [1] and this paper discusses how such haptic cues can contribute to this sensation in the context of a broadcast scenario. Two themes are developed. The first is 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 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 a number of interactive technologies have been applied to broadcast scenarios, these have focused on functionality, such as display of additional (often textual) information related to the content, connection to the internet and participation in activities, such as polls [4]. These services are abstract, information-orientated and provide a very indirect form of interaction with the content; they are unlikely to raise levels of presence. This article argues that presence can be boosted more effectively by engaging direct interaction paradigms based on haptic feedback.

Indeed, it has been 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 (see [1] or [5] for a review). However, the domain of broadcasting is restricted by distinct constraints (compared to those restricting the display of general virtual environments) and has received scant, but emerging [e.g. 6, 7], attention in the haptic literature. Correspondingly, this article focuses on the construction of a haptic enabled broadcasting system. More specifically, this article presents a comprehensive exploration of the issues underlying haptic multimedia broadcasting from the processes of contents creation through transmission and finally to viewing and interaction. It also describes the implementation of a prototype system as a proof of concept and demonstrates two possible content scenarios using this system. To achieve this, MPEG-4 [8], an IEC/ISO standard for streaming multimedia objects in broadcast specific applications, is adopted as a multimedia framework and haptic data are incorporated into the framework through extensions of MPEG-4 BIFS (Binary Format for Scene). The
implementation of this framework is described and two demonstrations that use it, “Home Shopping” and “A Movie with Tactile Feeling”, are introduced. Before embarking on a description of this technology, this article introduces and grounds the idea of haptic broadcasting by considering potential use scenarios and defines and discusses haptic media, a concept analogous to audiovisual media, the structure of which fundamentally shapes the kinds of haptic sensations that can be represented in a multimedia scene.
1. Potential Scenarios in Haptic Broadcasting

In contrast to virtual environments, broadcast programs are generally experienced linearly: they have a beginning, middle and end and are experienced only in that order. Although interactive narrative is an active research area [9], it is not yet close to transitioning to commercial broadcast systems, and currently viewers have little opportunity to reorder the shows they watch. 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 choice-less context, any research intending to add an inherently (at the perceptual level) interactive modality such as haptics [10] must address the fundamental question of how interactivity can be achieved and represented. If a viewer cannot make choices, what form of interaction can he or she engage in?

![Diagram showing examples of potential scenarios in haptic broadcasting](image)

(a) Learning handwriting               (b) Feeling the punch between boxers
Passive haptic playback

(c) Touching scene                  (d) Manipulating an object in a scene
Active haptic interaction

Figure 1. Examples of potential scenarios in haptic broadcasting
One key possibility is to allow viewers to experience and interact with the presentation of displayed scenes, but not their sequence [11]. This can be envisaged in several ways. In a passive haptic playback scenario, no direct interaction takes place and prerecorded haptic cues associated with the content could be displayed to the viewers in much the same purely experiential sense that audio and video currently are. For example, Figure 1(a), depicts an educational program in which a calligraphy expert is demonstrating her skills. She asks the viewer to grip the pen-like kinesthetic device (force feedback device) and starts to write a character. The viewer is physically guided through the path taken by the expert as they sketch the character [12]. In addition, in Figure 1(b), shows how a boxing event could be augmented with a representation of the impact of a punch. The magnitude of the impact can be measured by an accelerometer embedded in the glove and viewers can 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 [13]) or creating it by hand in a post-production phase, and then presenting these stimuli to viewers. In this scenario, a kinesthetic haptic device such as PHANToM (http://www.sensable.com) or CyberForce glove system (http://www.immersion.com) could be held or worn on the viewers hand or arm in order to provide recorded movements. Alternatively, an array of tactile stimulators could be worn or even embedded in the furniture to deliver cutaneous stimuli to the user’s skin. In both cases users need not make an explicit action in order to experience the stimuli; being in contact with the display equipment is sufficient.

On the other hand, if a scene contained 3D information (including elements of its dynamic behavior), an active haptic interaction scenario, in which viewers can actively explore the contours and surfaces of the displayed objects and manipulate them, would be enabled. For example, Figure 1(c) shows a program in which viewers are able to touch an actress’s face as another on-screen character is doing so. In this way viewers may identify more strongly with the program, indeed feeling as if they have become the actor. Technologies such as the Z-Cam, a combined video and depth camera, offer the possibility to capture this kind of live 3D scene. A second possible active scenario, shown in Figure 1(d), is to enable viewers to explore the feel of items on a home shopping channel. This 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 depth scene or full 3-D mesh of the objects of interest and their haptic properties. To best experience this content, viewers must use a combination of a kinesthetic device and a tactile array. Such a
device could take the form of a kinesthetic device enhanced with finger-tip sized tactile array on its end-effector and which would support the active exploration of a scene with rich information about the reaction forces felt and textures encountered. This is clearly a more demanding scenario, 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 [10] which draws a related distinction between the act of feeling something, an active exploratory process and the experience of being touched, a passive and very different occurrence. However, mapping these concepts onto the domain of broadcasting has led to the lines between these theoretical categories becoming somewhat blurred. This is most clear in the passive haptic playback scenarios described above. While few would dispute that an entirely prerecorded sequence of tactile cues presented to the skin represents a pure passive tactile experience, this is not the case for the movements of the pen interface of a kinesthetic device which has been constrained to follow a path. In this latter case, a user is free to explore the motions (or indeed shape) of the pen in their hand in any way they choose. They might include holding the device loosely, tightly or changing their grip on it. 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 perceptual experience of the user; in this theoretical sense such a system inherently creates active perceptual experiences. In light of this fact, the usage of the terms passive haptic playback and active haptic interaction in this paper can be understood in the concept of the kind of application level behavior they support and which is exemplified through the brief scenarios described above. In short, passive haptic playback signifies the delivery of prerecorded streams of information where active haptic interaction involves scenes which a user can choose to explore entirely as they wish.

It is also worth noting that the precise definition of the term haptic remains in some dispute among the various research communities that employ it [14]. This paper follows the majority of the literature in computing science [e.g. 15] and uses haptic as an umbrella term for two sub-categories of feedback: tactile and kinesthetic. Tactile refers to information sensed through nerve receptors in the skin, while kinesthetic 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, one of both. This display of 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 (see [16, 17] for recent reviews).
2. Haptic Media

A media format defines the structure of a media type: the kind of data and its internal representation details. In the context of a broadcast environment, it describes not only the contents that are displayed to a viewer, but also how they are 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 recorded in a studio or extracted from a library. These two data sources must be composed coherently (and, of course, artfully) in order to produce meaningful content, and it is the role of a media format to technically enable this. 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 which explicitly addresses haptic media. While there have been many virtual environment systems and APIs which have included representations for haptic information of various sorts, such as Reachin API (http://www.reachin.se) and H3D (http://www.h3d.org), these have generally described only static scene information. It is far less common to see examples where haptic information includes a temporal component. Those that have appeared have tended to be dedicated to a single purpose and simplistic in their implementation. For instance, in the domain of telehaptics [18, 19] users experience a stream of haptic cues, in the form of force data delivered sequentially to a haptic device. Similarly, in haptic training [20, 21], a stream of force or position data recorded from an expert is presented sequentially to a novice. Using a similar paradigm, access technologies, such as TVSS (tactile vision substitution system), have included systems which map video input from cameras to patterns of tactile activation in a grid of vibrating elements. Although streaming haptic data synchronously with other media forms, these systems all share the same basic approach: haptic information is captured, undergoes some immediate transformation to match the specific characteristics of a display device, and then is rendered immediately in synch with any other media (or alternatively stored to disk as a stream). While these approaches are useful within their particular application area, it is clear they do not constitute a general approach as they are tied to high quality of service communication links and specific input and output technologies. There are clearly many further forms of haptic data that they are incapable of expressing.

The main technical contribution of this article is a system which can convey rich and general purpose haptic media elements. Consequently, it needs to define exactly what the term encompasses. Given the varied, and frequently incompatible, nature of available haptic technologies, this definition is further complicated as it has to be grounded on a methodology for expressing sensations which are applicable to
a broad range of devices. To achieve this compromise, this article makes a distinction between two key
media categories: linear and non-linear haptic media. It also relates each of these to the kinds of
sensations that can displayed on commonly available haptic hardware.

Linear haptic media refers to haptic sensations which progress sequentially in time, such as displayed
movements or cutaneous patterns of touches to the skin which create experiences of passive haptic
playback. Essentially, linear haptic media encompasses events such as human touches, impacts, alarms,
direction cues, textures and on-screen motions (such as the bouncing of a ball). These kinds of sensations
are best displayed by technologies such as an array of tactile elements distributed over an area of the skin.
By varying which elements are activated over time, a rich range of sensations can be created. A number of
technologies exist to achieve this, including small mechanical motors, transducers, piezoelectric strips, air
jets and the direct application of electric currents [22]. This article defines one key type of the linear
haptic media to be a sequential series of actuation intensities which can be deployed to control a grid of
tactile stimulators spread over an area of skin. These intensities should correspond directly to audio and
video events in the content.

Non-linear haptic media, on the other hand, offers viewers interactivity and allows them to touch and
explore a haptically displayed object experiencing both force and textural information and producing a
compelling active haptic interaction. In this case, viewers feel objects only by exerting their own agency:
they must explore the environment to feel the haptic cues. These cues can take the form of not only force
and shape information but also the surface properties of objects’ texture: friction, roughness, stiffness.
Non-linear haptic media must also encompass object dynamics: how they move and behave in response to
user input. This can include general terms, such as mass and inertia, but also more specific ones, such as
spring constants defining the travel distance and sponginess of a virtual button. This kind of
representation is arguably best displayed on commercially available kinesthetic devices, such as the
PHANToM devices, and this article defines it to include not only a complete description of the shape of
all objects, but also their surface properties and dynamic behavior.

3. 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. This article
presents, therefore, a comprehensive exploration of the issues underlying haptic multimedia broadcasting,
from the processes of contents creation through transmission and finally to viewing and interaction. It
firstly describes mechanisms by which content producers can acquire or author haptic information to
accompany an audiovisual scene and manipulate this data (by cutting, pasting, sticking, or synthesizing). It then describes a system for representing and transmitting this data (over the airwaves, down cables, or through the internet). Finally, it discusses how a viewer experiences the haptic media: what devices can they use and what sensations can they feel when they are watching a haptic-enabled program?

3.1 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 contents and deals with user interaction. When the application behavior needs to be changed, a software update is produced. However, in broadcast (and many multimedia) scenarios, there is a clear distinction between the contents and the multimedia player. The two are independent; a player interprets and displays the contents. However, previous literature [11, 13, 23] focused on incorporating haptics into multimedia deals with limited haptic interaction modes specific to particular systems and scenarios. This reflects the lack of a systematic framework to deal with multimedia datasets that incorporate haptic information. There have been 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 VRML (Virtual Reality Modeling Language). It 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 as they are based on a simple download-and-play delivery system rather than one which has the ability to handle streams of media data. In essence, users of these existing systems need to 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, this article focuses on 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. It also supports the harmonious integration of these varied data types, providing a unifying system which can enable 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. Haptic media can be encoded/decoded as independent objects but be easily synthesized and synchronized with other audiovisual media.

Technically, haptic media can be described using MPEG-4 BIFS (Binary Format for Scenes), which allows it to be spatio-temporally coordinated with other audiovisual media in a scene. At the most basic level, BIFS describes interactive 3-D 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: new objects can be added, existing objects modified, deleted or replaced. This mechanism transforms a static binary scene into a data stream that can be sent over a network and synchronized with other streams (video, audio, and meta-data) [24]. Consequently, BIFS theoretically supports the transformation of haptic media into a stream synchronized with other audiovisual media streams. However, the current MPEG-4 specifications do not consider haptic media, an omission addressed by the work in this article. It describes extensions to the MPEG-4 BIFS system incorporating new types of node which can display haptic media and incorporate haptics.

![Figure 2](image.png)

Figure 2. Illustration of mapping tactile video to glove based tactile device. Four frames are shown in grid format on the above and in glove format on the bottom. Darker red circles on the gloves indicate active tactors with higher intensity.

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) and the dynamic properties of movement stiffness, inertia, direction and travel distance. These are intended to encapsulate the behavior of virtual controls,
such as buttons, sliders or joypads that might appear as part of virtual models of objects. It need also include tactile video in the form of a grey-scale video component. This format represents a grid of intensities of tactile stimulation which 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 (e.g. [25]). Figure 2 shows how such a 2D set of intensities might be represented, albeit with some distortion of the shape, on a glove based tactile array. Finally, two key aspects of movement data, although not implemented in this article, are poses (positions and orientations of objects or bodies) and forces and torques that represent one part of the human body.

Figure 3. Diagram of the haptic broadcasting system based on standard MPEG-4 framework. Authors extensions to this standard diagram noted in bold text.
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 the creation of the content through to its consumption and is briefly reviewed here. First, media content creation processes take place. These result in the generation of media types which require very different representations and which must be seamlessly merged. In MPEG-4, linear media (of any modality) are typically assigned a dedicated stream due to their large size. These streams are compressed with suitable domain specific encoders (such as MP3 for audio) and transmitted progressively based on information relating to synchronization and presentation. Tactile video corresponds to this type. However, non-linear media, such as scalar values in a virtual model, can be encoded so that they will be transmitted immediately and stored locally using MPEG-4 BIFS due to their relatively small size. Haptic surface properties and dynamic properties correspond to this type. Larger non-linear media, such as bump-map textures describing roughness, are typically encoded in a unique stream but transmitted immediately. During MPEG-4 data transmission, all this information is multiplexed into a single stream by the sender, sent through a supported network (e.g. Real-Time Transport Protocol (RTP) or MPEG-2 Transport Stream) and then de-multiplexed 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. Additional details of each of these components which relate to haptic broadcasting are included in the following subsections.

3.2 Contents Creation

While standard tools exist for capturing audio and video media, it is less obvious how the same objective might be achieved for haptic media. Consequently, this section provides an overview of the mechanisms by which haptic media can be created. Three key approaches exist: Firstly, data can be recorded using physical sensors. Secondly, it can be generated using specialized modeling tools. Finally, it can be 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 [26]. The dynamic properties of haptic buttons have been acquired by measuring and analyzing the force profiles of real physical buttons [27]. Movement data can, for instance, be measured with a 3-D robotic arm equipped with force-torque sensors or with a motion sensor (such as an accelerometer) embedded inside some object of interest. For instance, a boxing glove equipped with such a device could broadcast information about the forces its wielder is experiencing. It is also possible to capture full or partial 3D information from a scene which can then be used to produce haptic models. This can take place in two ways: detailed 3D meshes of real objects can be produced by 3D scanners or a more
rapid approach relies on a depth video camera, such as the Zcam™ (http://www.3dvsystem.com), which can capture 2.5D information to record dynamic real objects (such as people) at video frame rates. Tools to generate audiovisual media, in the form of 3D modeling environments, are commonplace but there have been a few efforts to integrate haptic properties into a 3D scene in a modeling tool. Two exceptions are the author’s K-HapticModeler™ [28] and HAMLAT [29]. Both of these tools provide interfaces which support construction of a 3D scene and allow both haptic surface properties and dynamic movement properties to be assigned to parts of that scene. Movement data can be created by mimicking and recording the movements in a video scene. A few researchers (including the authors) have developed software tools [25, 30] that play video and simultaneously show and record a user’s position (with a pointing device) to enable tracing movement in scenes (such as orchestral conducting) which involve dynamic human motion. This spatiotemporal path can later be played back on a kinesthetic device, effectively providing a trace of the original user’s movements in synch with the audio-video content. In one of these tools [25], tactile video can also be recorded through gestures made on touch sensors such as touch pads or touch screens. Essentially, as a video is played a user can input patterns of tactile sensation by making movements on the surface such a sensor, creating a sequential stream of 2D information which can be used to represent patterns of tactile activation and which is synchronized with the audiovisual content. However, these tools are relatively basic and this article suggests there is substantial work to be done in developing haptic editing tools with the expressivity of current visual and audio equivalents.

The same is true for the automatic generation of haptic media from other media sources. While some associations are relatively obvious, 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 [e.g. 31]. This topic is also in its infancy and requires further attention before concrete conclusions about its usefulness can be reached.

After haptic media has been produced, it needs to be edited and synchronized both spatially and temporally with audiovisual media. This stage, essentially one of editing and arranging, fundamentally relies on the properties of the media format. The structure of format proposed in this article, in the form of extensions to MPEG-4 BIFS, is described in the next section.

3.3 Transmission

Though MPEG-4 BIFS is an extensible system for representing media objects, it does not include haptic media. This article describes the implementation of new BIFS nodes to achieve this functionality and support the synchronized representation and transmission of a media stream including haptic media.
In this way, it enables spatio-temporal relationships between haptic and the audiovisual media to be produced and stored. The new nodes for representing haptic surface properties, dynamic properties, and tactile video and their encoding method for transmission are described below. An overview of the nodes is presented in Figure 4.

**Figure 4. Haptic media nodes in BIFS**

### 3.3.1 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 which enables one geometrical form to be combined with numerous different surface appearances (or vice versa). As can be seen from Figure 4, the Appearance node incorporates several nodes related to graphical texture. In order to add haptics, we extend it with a new field, hapticSurface. This field can point to 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, it includes a dynamicSurface field, explained in more detail below. The scalar parameters of haptic surface properties are encoded using a standard BIFS encoder for numerical data, while the bump-map images rely on any image encoder.

3.3.2 Dynamic Properties of Haptic Widgets

A number of haptic widget nodes (intended to represent controls on virtual objects) have been implemented based on [32, 33], but in the interests of brevity only one, ButtonSurface node, is described in detail. In virtual haptic pushbuttons, a basic force profile consists of three distinct stages which can be encoded with travel distances, spring coefficients and deadband force parameters [32]. The node also stores an axis vector indicating the travel direction. Essentially, the operation of this node is such that when it pushed along its vector of travel and exceeds a preset travel distance, it raises a flag and generates an event which can be handled within the BIFS framework (and can impact on 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.

3.3.3 Tactile Video

In order to provide tactile feeling through tactile devices, 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 including vibrating motors, electrodes and piezoelectric ceramics. Therefore, each frame of actuation intensity array of rectangular tactile device can be represented with an image and sequences of frames can be represented with a video. This is termed tactile video, is identical in fundamental structure to conventional video media and is implemented in a new BIFS TactileDisplay node. Within this, a standard SFTextureNode stores the actual grey scale tactile video file. Any resolution differences between the stored video and the local tactile device can be resolved through basic image resizing operations at viewing time. This tactile video can be compressed with video encoder.

3.3.4 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 is reasonable that the depth video node should be defined in the Shape node. In MPEG-4 BIFS, a DepthImage node is already defined [34]; however, it is not included in the Shape node but an optional component of the Transform node. This existing DepthImage node includes a description of camera pose and 2.5D textural information. However, this article proposes a new DepthMovie node that can be stored in the geometry field and which follows the concept of the Shape node. The color video corresponding to the depth video is stored in the Appearance node through a BIFS standard MovieTexture node. In order to make the depth video touchable, the HapticSurface node needs to be set 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 is used to store the video data using a MovieTexture node. Once again all video media can be compressed with any video encoder.

3.3.5 MPEG-4 Transmission

In the transmitting stage, the audiovisual media are compressed with suitable encoders as normal. The haptic media in image or video format are compressed with any suitable image or video encoder and all haptic media in scalar value format, including a scene descriptor if one is present, are compressed with a BIFS encoder. All this information is then saved in an MP4 file format [8], designed to contain the media information of an MPEG-4 presentation in a flexible, extensible format that facilitates interchange, management, editing, and presentation of the media. This file is transmitted to viewers through a streaming server. MPEG-4 content can be carried over many different transport layers including traditional broadcast methods and IP networks. Therefore, the resulting content can be streamed through a satellite, over the air-waves and through the Internet. At the viewer’s own site, the transmitted content can be demultiplexed and decoded and the media shown in the viewing and interaction stage.

3.4 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 more traditional experiences, such as watching and listening. The decoded media is fed to a compositer process which has access to the BIFS scene graph. Traditionally, the compositer scans the scene graph, determines what audio and visual content should be shown and then passes this to the audiovisual renderers that actually handle the display of the media on an entertainment device, such as a TV.
The system proposed in this article extends the compositors 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, developed by the authors, which consists of a commercial 3 degree-of-freedom kinesthetic device, a PHANToM, combined with a custom fingertip pneumatic tactile device [35]. Recently, thermal perception has also been explored as a mechanism to provide more realistic tactile sensations [36].

![Figure 5. Example of a hybrid haptic device.](image)

The tactile sensation renderer converts tactile information from either the haptic renderer or tactile video (routed through the compositors) 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 which supports passive tactile playback on viewer’s hand, a bespoke glove type device incorporating 76 vibrating motors produced by the authors [25].
Figure 6. Example of a tactile device for passive tactile playback. (left) Completed tactile device packaged with outer glove. (upper right) Vibrating motors on inner gloves. (lower right) Device controller in the glove wrist band

3.4.1 Implementation

The proposed haptic broadcasting system is implemented using GPAC (http://gpac.sourceforge.net), a multimedia framework based on the MPEG-4 Systems standard. It supplies basic modules for en/decoding and multiplexing/demultiplexing various multimedia including MPEG-4 BIFS and for playing the transmitted multimedia audiovisually. In order to stream the content, the Darwin Streaming Server (http://developer.apple.com/opensource/server/streaming) was used. It allows streaming media to be delivered to clients across the Internet using the Real-Time Streaming Protocol (RTSP) that uses RTP as the transport protocol.

To display this data to viewers, an MPEG-4 content player that shows the audiovisual media and provides haptic interaction and playback was implemented based on the Osmo4 player included as part of the GPAC framework. Osmo4 imports all modules for traditional audiovisual media decoders and displays these data through speakers and a standard visual unit. In order to enable haptic interactions, a kinesthetic device driver was imported from the manufacturer, a tactile device driver was developed for the glove-type tactile device manufactured by the authors [25], and the haptic rendering algorithm to touch depth video was developed and included in the player. The haptic rendering algorithm was
implemented with a conventional proxy-based solution for 3D mesh models [37] including dynamic haptic UI widgets. It also included a combination of a modified proxy graph algorithm [13] and a depth image-based haptic rendering module [38] to present the depth video. Since the force can provide surface properties such as stiffness, friction and roughness without any cutaneous feeling [39], the case of applied force rendering was used as a proof-of-concept of the haptic broadcasting system.

3.4.2 Demonstration: Home Shopping Scenario

Figure 7. Demonstration of the home shopping scenario as a active haptic interaction

A home shopping scenario was implemented as a practical example of a haptic broadcasting scenario. 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, the video cues to a close up of the device in the form of a virtual model. 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 depth video from the captured shopping host. As can be seen from Figure 7, viewers can touch the displayed virtual objects and manipulate their controls by wearing and moving a thimble attached to the end of a standard kinesthetic device.

3.4.3 Demonstration: A Movie with Tactile Feeling

The display of passive tactile playback was explored in the context of a viewer watching a movie. In one scenario, a Hollywood movie, “Ghost” was augmented with manually authored tactile cues produced in a dedicated authoring tool [25]. In one instance, shown in the Figure 8, the actions of touching an object (e.g. when the ghost’s arms penetrate the door) were mapped to the tactile display, linking the point of view of the on-screen 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.
4. Conclusions and Future Work

Technical advances in multimedia applications have improved the aural and visual quality of multimedia and allowed more immersive experiences. Haptics technologies in human-computer interaction context have improved the performance of tasks in specific applications, such as surgical training, rehabilitation, and military training. However, supporting haptic interaction in the context of entertaining consumer multimedia applications has been overlooked. Addressing this omission, this article provided an overview of some potential haptic broadcasting scenarios, defined haptic media and discussed (and described the implementation of) a method which enables it to be combined with traditional audiovisual media. This is accomplished through adopting a standard multimedia framework, MPEG-4. Some demonstration examples were implemented to show the feasibility of the proposed approach. The work in this article represents a key step towards popularizing the notion of haptic media, and opening the door to a whole range of new possibilities for consumer entertainment.

The proposed definition and implementation, however, are not perfect, requiring future sophistication and major development of the whole haptic broadcasting chain over some specific networks, such as IPTV. Furthermore, this article highlights the need to develop expressive and easy-to-use editing/authoring tools to support and encourage the widespread creation of haptic media. Finally, further development at both practical and conceptual levels is required to fully explore the scope, advantages and usability of this emerging technology. For example, one key omission to the work reported in this paper is the lack of empirical studies clearly demonstrating the value of haptic technology in broadcasting. Although many theoretical aspects of this, such as the importance of touch for engendering feelings of presence, have been examined in the field of virtual environments, it is important to conduct studies
confirming this remains true in the domain of everyday, sit-back home entertainment. To summarize, this paper establishes a framework upon which future work on haptic broadcasting can be grounded, but the both the commercial applications and limits of this domain have yet to be fully understood, let alone attained.

6. 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)

7. References
477-491.
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 degree in Mechatronics from the Gwangju Institute of Science and Technology (GIST), Korea.

Ian Oakley is an assistant professor of Human Computer Interaction at Lab:USE in the University of Madeira, Funchal, Portugal. He is a member of the ACM special interest group on Computer Human Interaction. His research interests focus on multi sensory interfaces. Oakley received a PhD degree in human computer interaction a Joint Honours (First Class) in Computing Science and Psychology from the University of Glasgow, UK.
Yo-Sung Ho is a professor of the Information and Communications Department at the Gwangju Institute of Science and Technology (GIST), Korea. He is a Senior Member of the IEEE and the director of Realistic Broadcasting Research Center (RBRC) at GIST. 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 received a PhD degree in Electrical and Computer Engineering from the University of California at Santa Barbara (UCSB), USA.

Yeongmi Kim is a PhD student of the Mechatronics Department at the Gwangju Institute of Science and Technology (GIST). Her research interests are psychophysics, tactile display design, haptic content authoring, and haptic interaction in broadcasting. Kim received a B.S. degree in Control and Instrumentation Engineering from the Korea Maritime University, Busan, Korea.

Jeha Ryu is a professor of the Mechatronics Department at the Gwangju Institute of Science and Technology (GIST). He is the director of national Haptics Technology Research Center. His research interests include haptics, haptic rendering, haptic interaction control, design and control of haptic devices, as well as haptic-included modeling and broadcasting. Ryu received a PhD degree in Mechanical Engineering from the University of Iowa, USA.