

Haptic Interaction in Realistic Multimedia Broadcasting

Jongeun Cha¹, Jeha Ryu¹, Seungjun Kim²,
Seongeun Eom², and Byungha Ahn²

¹ Human-Machine-Computer Interface Lab., Dept. of Mechatronics,
Gwangju Institute of Science and Technology,
1 Oryong-dong, Buk-gu, Gwangju 500-712 Republic of Korea,
{gaecha,ryu}@gist.ac.kr,
<http://dyconlab.gist.ac.kr>

² System Integration Lab., Dept. of Mechatronics,
{zizone, seueom, bayhay}@gist.ac.kr,
<http://si.gist.ac.kr>

Abstract. In this paper, we discuss a haptically enhanced multimedia broadcasting system. Four stages of a proposed system are briefly analyzed: scene capture, haptic editing, data transmission, and display with haptic interaction. In order to show usefulness of the proposed system, some potential scenarios with haptic interaction are listed. These scenarios are classified into passive and active haptic interaction scenarios, which can be fully authored by scenario writers or producers. Finally, in order to show how the haptically enhanced scenario works, a typical example is demonstrated to explain specifically for a home shopping setting.

1 Introduction

Rapid development of computing and telecommunication technology such as enhanced CPU speed and power, low cost memory, and ultra fast communication network has led to digital multimedia age, where viewers can be immersed in some 3D visual and audio contents. Moreover, viewers can interact even haptically with the multimedia contents beyond passive watching and listening. Traditionally, these interaction and immersion are possible only with fully virtual worlds (VR), in which the world is filled with synthesized objects. In addition, only one viewer or small number of viewers can share the virtual contents at the same time [1].

In the area of broadcasting system, new technology is being developed and is available in terms of digital multimedia broadcasting through the air or through the Internet. Main multimedia contents are, however, limited to 2D video and sound so that feeling of full immersion is still far from the reality. Interactivity is also being pursued in these days in a very simple form such as selection and retrieval of 2D AV contents. In the report of the Media Interaction group of Philips Research in Eindhoven, recent efforts for the interactive television are

well summarized from its concepts and history to storytelling application [2]. If the broadcasting network in the future can be completely integrated with communication network like Internet, useful techniques in the network services such as chatting program, server-client system, web casting, device communication can be technically available for the broadcasting system. All these, in some scenarios, may make us enjoy attractive bi-directional services by immersing dynamically into the broadcasting productions that may include sense of touch if viewers want to fully interact with more realistic multimedia contents.

With ATTEST project, which started in March 2002, the development of the first commercially feasible European 3D-TV broadcast system has been in progress. In [3], 3D-video chain of ATTEST including 3D content creation, encoding, transmission, and display stages is described. They have been trying to use head tracking to drive the display optics and develop two 3D displays, one for a single viewer and one for multiple viewers. At the same time, O'Modhrain and Oakley [4,5] discussed the potential role that haptic or touch feedback might play in supporting a greater sense of immersion in broadcast content. Presenting Touch TV, they showed two potential program scenarios: the creation of authored haptic effects for children's cartoon and the automatic capture of motion data to be streamed and displayed in the context of a live sports broadcast.

Unlike the simple addition of touch-enhanced contents to the broadcast media in some scenarios, in this paper, we are investigating more comprehensive realistic multimedia broadcasting system that can include haptic interaction in addition to 3D audio-visual contents. More specifically, firstly, we present a top-level view of creating, editing, transmitting, and displaying with viewer interaction fully immersive multimedia contents in a broadcasting system through the Internet. This is a new attempt beyond a multicasting system that utilizes small number of shared computational platforms servicing for ten to hundred viewers only. We describe each stage of the proposed system in terms of data type and generic processing algorithm, etc. Secondly, in order to present usefulness of the proposed system, we have listed some possible scenarios with haptic interaction. Producers who may be science teachers/professors, educationers, geographers, artists, etc can develop specific sense-of-touch-added scenario. Role of engineers is to provide these producers with content creation tools such as sensors embodied in a real or virtual object, multimedia authoring tools, interaction techniques and devices with haptic sensation. Then viewers can enjoy the immersive interaction dynamically to get their indirect experience as well as additive information. Finally, we present a typical application example to show how the haptically enhanced scenario works. This simple demo system utilizes Augmented Reality (AR) techniques, which show excellence in synthesizing seamless videos in real-time, multimedia streaming technology, and a 6 degree-of-freedom haptic device for a homeshopping setting.

2 Haptically Enhanced Multimedia Broadcasting System

From the high-level view, a general broadcasting system can be divided into four parts: capture, edit, transmission, and viewing. A producer captures a scene

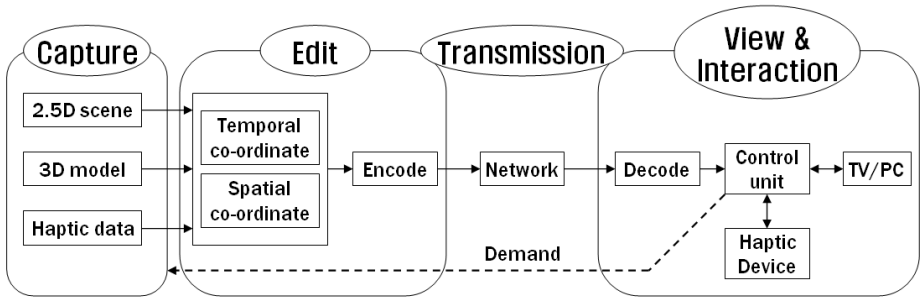


Fig. 1. Haptically enhanced multimedia broadcasting chain

with a camera and edits that scene to make a broadcasting content by cutting, sticking, synthesizing computer graphics and audio contents, and so on. Then the authored program is transmitted to the viewer via the airwave, satellite, cable, or Internet. The viewer passively watches the program in front of the TV with a remote control in his hand for simple interaction, for example, changing the channels.

Figure 1 shows the proposed haptically enhanced multimedia broadcasting system. In this figure, contrary to the traditional broadcasting system, the video media, which is the sequence of the 2.5D scenes plus virtually synthesized computer graphics models, has the depth information in addition to the 2D image. It depicts the geometry of the captured scene in terms of 3-dimensional coordinate from the camera view, not the arbitrary view. 3D computer graphics models can be easily registered with the 2.5D scene in a 3-dimensional space. In addition to the video media, haptic data, the authored, recorded, or physically sensed media representing the kinesthetic and tactile information, e.g. material property data at each pixel for texture tactile feeling or object weight data associated with an object-of-interest for force sensation, is combined to give the viewer haptic effects. The edited hyper media, which can be defined as 3-dimensional audio-video media synchronized with the haptic data, is transmitted through encoding and decoding operations to the viewer site via the Internet. The control unit receiving the media renders the stereo images and 3D sound to the display device by processing the 3-dimensional audio-video media and controls the haptic device to give a haptic interaction to the viewer. In this way, the viewer can actively interact with the 3-dimensional hyper media as well as can feel the haptic effects. Besides viewers can also demand additional data via the bi-directional Internet channel.

In the capture stage, all scenes are captured with the depth information. A depth camera like Zcam TM [7] yields conventional 2D-video accompanied with depth-per-pixel via a direct depth-sensing process. Multiple 2D cameras can generate the same information even though the overall processing is more complicated than the direct depth sensing [6]. A 3D scanning and subsequent reconstruction method or modeling with a CAD program may capture virtually synthesized 3D models. These two kinds of data are composed of the 3-dimensional

video media. 3D audio sound may also be produced by new technology. In the meantime, some authored haptically enhanced data such as vibration of bee wings in the kid's animation, kicking force of a soccer ball, which may be obtained in reality by a real accelerometer embedded inside a ball, etc may also be captured along with the audio-visual data [5]. Notice that in order to synchronize the haptic data with the audio-visual data, haptically related data such as position, velocity, and acceleration data of the corresponding object of interest should also be recorded.

In the edit stage, the captured data are coordinated temporally and spatially to a program. Since the handled data is basically 3-dimensional, the composition operation needs to be managed in the 3-dimensional space. The 3D model is synthesized to the captured 2.5D scene by the z-keying or the Augmented Reality technique. The z-keying is the process to put the 3D model anywhere in the captured scene by giving depth. The Augmented Reality technique is a process to place the 3D model on a specific position, utilizing a specific feature in the captured scene. In addition, the captured haptic data is synchronized with the video-audio data utilizing the haptically related data in this stage.

In the transmission stage, the finished program is transferred to viewers via Internet. Because the communication channel is Internet, many useful bi-directional interactions between viewers and broadcasters may be possible. Viewers can demand some data from the broadcaster especially in a live broadcasting system.

In view and interaction stage, the control unit processes the video media and the haptic data to produce a stereoscopic scene to the display device and to control haptic device by using haptic rendering algorithm. In haptic interaction stage, viewers may feel the transmitted haptic effects that are synchronized with some specific scenario passively by putting their hands on the vibrotactile display device. Or they can actively touch, explore, and manipulate some transmitted 3D CG objects or background 2.5D scene according to the preplanned path or to the viewers' will.

3 Potential Haptic Interaction Application Scenarios

A distinctive characteristic of the broadcasting program is that it is captured and edited nonlinearly but broadcasted linearly in time domain. The program, as any other television show, has to be waited for, i.e., has a fixed position in the daily schedule. Furthermore, it is not possible to go back, or start the program over. Therefore, in this paper, the haptic interaction occurs in the viewer site only and does not change the path of the program story. But viewer can explore the program or manipulate some digital objects by active touch or by the passive haptic effect that is authored and provided. In this section, we list some potential application scenarios that take advantage of the haptic interaction.

Potential haptic interaction scenarios may be classified into passive or active interaction: Passive haptic interaction scenario just records some haptically related data when capturing audio-visual scene including object-of-interest and sends them to the viewer with interaction time indicator (e.g. caption on the

screen). Then, the haptically-related data controls a haptic device worn on the viewer's hand or arm. In this interaction, therefore, viewers are only passive. Active haptic interaction scenario captures 2.5D audio-visual scene only or together with full 3D virtual objects of interest that are either independent of the 2.5D scene or dependent on it as is the case of Augmented Reality. Then, these data are transmitted to the viewer's control box, where 2.5D scene, virtual objects, and haptic device are synchronized temporally and spatially. After this, viewers can interact actively with the scene or object-of-interest, i. e., can touch or push buttons. Collision detection and response calculations are all done in the viewer's control box in this case. In the following potential scenarios, scenario 1 is passive and fully authored by producer. This passive interaction may be lively broadcasted, e.g. live soccer game. Meanwhile, scenarios 2, 3 & 4 are active and viewers take time to explore or manipulate the virtual object-of-interest. This active mode may be live too.

3.1 Scenario 1: Feeling Haptic Data

In teaching programs of some manipulation techniques, a producer may want the viewer to follow the instructor's movement because it is very useful to learn expert's manipulation technique by viewing his actions as well as by tracking. For example, the expert in pen writing is showing how to write a pretty hand. He asks the viewer to grip the pen-like haptic device and starts to write a character. The viewer is completely guided to move the pen following the expert's writing. In this case, the producer captures a handwriting expert visually as well as haptically by recording the hand poses in real time or off-line. The captured scene and the recorded pose data are edited synchronously to an educational content. The control unit in the viewer's site displays the expert's handwriting and controls a haptic device to follow the recorded hand pose with the viewer wearing the pen-type haptic device. Note that this scenario is record-and-play type haptic interaction. This scenario does not record haptic data but record pose (position and orientation) data in the capture stage. In the interaction display stage, recorded pose data will drive the haptic device with force generation so that viewers feel touch sensation.

3.2 Scenario 2: Touching 2.5D Scene

While viewing a TV program, a viewer sometimes may want to touch a weird shaped thing or an actor's face to acquire the shape or the skin feeling. For example, in a drama, lovers are looking into the eyes of each other and going closer. A viewer may want to touch one of the acting lovers on the face. His face is slowly closed up with a camera and the 'haptic interaction' caption shows up on the corner of the screen. The viewer then touches the actor's face by a haptic device worn on the hands. This touching interaction may be possible if the video media has 2.5 or 3-dimensional information. For this scenario, a program is captured as the 2.5D scene that contains object-of-interest to touch. Since the viewer interacts with the object by the physical force contact, in the

captured scene the object-of-interest should be static or moving slowly. The producer overlays a haptic interaction caption indicating when to touch. When a viewer wants to touch, the control box will perform collision detection and force computation to drive a haptic device worn on viewer's hands.

3.3 Scenario 3: Touching and Manipulating 3D Models

Sometime, a producer may want to let the viewer feel something in a program as well as see it to give deeper understanding. For example, in an education channel for the science of dynamics an instructor is explaining the force equation for a spring. After teaching the theory, he uses augmented reality technology to arrange few virtual springs on a real experiment desk and asks the viewer to push the spring and feel the spring force with the haptic interaction caption. Then he repeats the experiment changing the number and the arrangement of the springs. In this scenario, the viewer can have deeper knowledge of the spring properties by handling it directly as well as being taught the theory. For this, a program is captured as 2.5D scene including the clue, like a feature, for composing the realistic looking 3D model using Augmented Reality technique. In the program, the MC puts the features and makes the experimental environment seeing synthesized video in real time. After capturing the program, the producer augments the 3D model, such as a virtual spring, exactly and stably and attaches the haptic interaction caption. A viewer follows the program and interact haptically. This scenario is different from the previous scenario in touching and manipulating 3D virtual objects instead of 2.5D object that is captured from the real scene.

3.4 Scenario 4: Touching and Manipulating 3D Models on Demand

Sometimes, a viewer may want to buy a product-of-interest while watching a TV program. In this case, providing him with haptic sensation in addition to the audio-visual information of the product can help his purchasing decision. For example, while viewing a drama, a viewer may get interested in a camera that the actor is using. He picks up a remote controller and pushes the menu button to get the camera product information. After looking into the specification, he may get into the haptic mode to try to touch the camera. He, then, can manipulate some buttons, or touch the surface feature, or feel the inertia of the downloaded virtual camera. For this scenario, a producer makes an advertisement's content that includes the product information and 3D virtual model of the product and saves them in the server. The viewer demands the additional information for the product, downloads it, and examines the product carefully for purchasing by exploring and manipulating it.

4 Demonstration Example

This section explains a demonstration example of an active haptic exploration/manipulation in a home shopping setting. To explore how the scenario works,

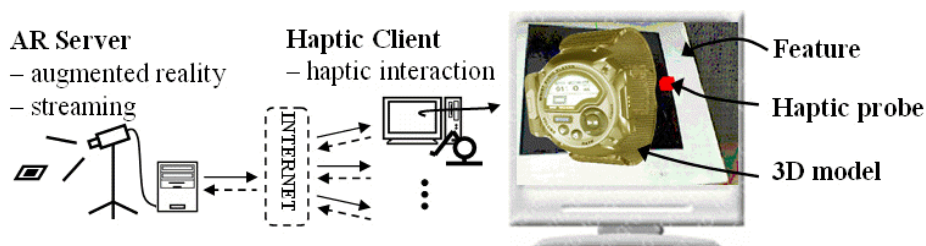


Fig. 2. Demonstration system overview

we have implemented a simple broadcasting system, which is capable of delivering video media on the Internet and giving haptic interaction. The demo system is constructed basically following the stages of the broadcasting chain in Fig. 1, with two major subsystems: AR (Augmented Reality) server and Haptic client. As shown in Fig. 2, the AR server consists of a typical AR system and a streaming server. This server makes it possible to create a broadcasting content and stream it via Internet. Therefore, Capture, Edit and Transmission stage are all performed in this server. The View & Interaction stages are implemented in the Haptic client system. It receives the content and realizes viewer's interaction with a connected haptic device. We have used Augmented Reality techniques based on ARToolKit[9], multimedia streaming technique, and a 6-dof haptic system (PHANTOM[10]) in the demo system. ARToolKit is a software library that can be used to calculate camera position and orientation relative to physical markers in real time. The SensAble Technologies PHANTOM makes it possible for users to touch and manipulate virtual objects through the help of the GHOST SDK (General Haptic Open Software Toolkit) that is a powerful, C++ software tool kit that eases the task of developing touch-enabled applications. Firstly, we explain how the system works technically and then show how the home shopping application scenario can be realized. The AR system captures the real environment scene that contains a known marker and obtains the position and orientation of the marker relative to the camera. Then, the streaming server packages the captured scene and the marker's location and transfers them via Internet. The 3D model data is transferred in advance through the other channel. At viewer's end, the haptic client system receives the transferred data and augments the 3D model to the captured scene at the marker's location. The haptic probe, that corresponds to the handle of a haptic device grasped by a viewer, is graphically overlaid to the augmented scene relative to the camera reference. The viewer is able to interact with the 3D model by moving the handle as watching the scene.

In the example application scenario, we have considered a situation that a shopping host tries to advertise a product: Wrist-held MP3 player. She wants to explain functions and features of the product using visual and haptic sensation. She explains the product by rotating it and making viewers touch surfaces or push buttons. When the host puts a marker-based feature in the camera range, viewers in the haptic client site can watch a scene augmented with the MP3

player 3D model. Then, they grasp the handle of haptic device to interact with it. Watching the haptic probe navigating in the scene according to the each viewer's intention, they can actively explore the outlines of the product and push a functional button to know how it works.

We have implemented the first demo example based on the broadcasting chain, which is haptically enhanced. It makes us fully immersed into the broadcasted world and provides much interest in experiencing the broadcasted contents.

5 Future Work

As discussed in the previous sections, haptic interaction in a broadcasting system requires new data format and processing technique in each stage. For example, passive haptic-related data must be prepared in the capture/edit stage and be transmitted along with the audio-visual data plus some camera-related data. In addition, conventional haptic rendering algorithms were mainly developed focusing on the interaction between 3D models in virtual environments. Moreover, the process to get the 3D model by scanning the real object or modelling with a CAD program is time-consuming. The advent of the ZcamTM[7], however, makes it relatively easy to get the 2.5D model of the real scene because the process is just capturing not modelling. In a scenario where complete 3D model is not needed e.g. scenario exploring only the visible part inside a view fulcrum, the scene data will be 2.5D and a novel haptic rendering algorithm (collision detection and response calculation) for the 2.5D scene is needed.

In this paper, we consider only the interaction between viewers and the broadcasting multimedia. The noticeable feature of the interaction in the haptics is the social presence, the feeling of being socially present with another person at a remote location[8]. In a live broadcasting scenario, an audience may take part in a program and communicate with other people viewing the same program by haptic interaction. Since the communication channel of the broadcasting via the Internet is bi-directional, it seems to be possible. One can pursue to establish the system and the scenarios for the social presence in the future.

6 Conclusions

In this paper, we discussed a top-level structure and brief data structure and processing algorithm of future realistic multimedia broadcasting system that may include sense-of-touch. Also, some potential scenarios taking advantages of haptic interaction were listed in a realistic broadcasting in which the video media is 3-dimensional. Finally, an application example demo system is presented. Addition of the haptic interaction to the conventional audio-visual contents will improve the immersion of the viewers together with rich contents. Moreover, full engagement to the realistic multimedia by haptic interaction can enhance amusement as well.

Acknowledgements. This work was supported by the Ministry of Information and Communication (MIC) through the Realistic Broadcasting IT Research Center (RBRC) at Gwangju Institute of Science and Technology (GIST).

References

- [1] Grigore C. B., Philippe C.: Virtual Reality Technology. Second Edition, by John Wiley & Sons, (2003)
- [2] Bukowska, Magdalena: Winky Dink half a century later : interaction with broadcast content : concept development based on an interactive storytelling application for children. (2001)
- [3] Andre R., Marc O. B., Christoph F., Wijnand I., Marc P., Luc V. G., Eyal O., Ian S., Philip S.: Advanced Three-dimensional Television System Technologies. ATTEST Publication, Padova, Italy
- [4] O'Modhrain, S., Oakley, I.: Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS '04. Proceedings. 12th International Symposium on. (2004) 293–294
- [5] O'Modhrain S., Oakley I.: Touch TV: Adding Feeling to Broadcast Media. in proceedings of the European Conference on Interactive Television: from Viewers to Actors, Brighton, UK, (2003) 41–47
- [6] Grau O., Price M., Thomas G. A.: Use of 3-D Techniques for Virtual Production. BBC R&D White Paper, WHP 033, (2002)
- [7] 3DV Systems. <http://www.3dvsystems.com>
- [8] E. Sallnas, K. Rassmus-Grohn and C Sjostrom: Supporting Presence in Collaborative Environments by Haptic Force Feedback, ACM Transactions on CHI 7(4), ACM Press, 2000, 461–476
- [9] ARToolKit computer vision software: <http://www.hitl.washington.edu/artoolkit>
- [10] PHANTOM, SensAble Technologies, <http://www.sensable.com/>