VECIMS 2004 – IEEE International Conference on
Virtual Environments, Human-Computer Interfaces, and Measurement Systems
Boston, MD, USA, 12-14 July 2004

Interaction of Free Viewpoint 3D Video with a Haptic-Based Environment for Training Applications

Mauricio Orozco, WonSook Lee and Abdulmotaleb El Saddik
Multimedia Communications Research Laboratory
School of Information Technology and Engineering, University of Ottawa
{morozco, abed}@discover.uottawa.ca, wslee@site.uottawa.ca

***Abstract** – Traditional 2D video sequences have been characterized as a static and rigid set of scenes. This mature technology faces two major drawbacks: 2D video scenes as a set of standstill and inflexible sequences, and the limited data available to the viewer, that is, no free selection of viewpoint. A novel image media channel has been proposed to record dynamic visual scenario in the real world defined as 3D video. Creativity and imagination has made possible a closer interaction between humans and computers. This interaction depends on haptics, that is, the sense of touch given by computer technology. Virtual Reality environment is evolving towards more realistic graphical representations. In this paper we will present a novel approach to integrating free viewpoint 3D video with haptic devices in an environment suitable for training.*

***Keywords** – Haptics, 3D modeling, Virtual Environments*

I. INTRODUCTION

Generating three-dimensional images that appears to surround the user has made virtual reality a more appropriate interactive medium to respond to major application's demands in technology. Free viewpoint 3D digital video sequence provides users with the possibility of changing the viewpoint of a presented video during its playback.

Virtual Reality is evolving towards improved environment realism. Recently, research is using deformable models to enhance haptic rendering in virtual environments. Innovative devices can reproduce human behavior and human features from real pictures.

Virtual reality environments integrated with haptics systems is becoming an active area of research although many issues are still in development. As it is the case with virtual reality environments, more realistic models need to be extracted from real scenes or real objects. A novel image media channel has been proposed to record dynamic visual scenarios from the real world defined as 3D video. The full shape of objects and their actions can be recorded from multiple surrounding video cameras. There are some issues that need to be resolved for making 3D video more feasible as suggested by [12]. Among these issues are: computation speed requirement, fidelity of data acquisition, width of observable area, data compression, editing capabilities and visualization performance.

A virtual training environment could be enhanced by the interaction between free viewpoint 3D Video architecture and haptics devices. The overall architecture of this prototype is integrated in four components, which are described in this paper.

This paper is structured as follows: In Section 2 we present a study about current state of the art research in areas related to the proposed system. In Section 3 we provide the overall architecture. Section 4 describes the impact and the potential applications of free Viewpoint 3D Video with a Haptic-Based Environment. Section 5 concludes the paper and gives insight into future work.

II. RELATED WORK

The process of extracting physical models from a multi-viewpoint involves the following three stages:

- 2D video recording: Multiple synchronized and calibrated cameras are set to take images from multiple viewpoints
- 3D reconstruction: Different techniques have been studied for the generation of the virtual model. Visual reconstruction techniques can be classified according to the way that generates novel views. There are two main groups model-based and image based.
- Generating synthetic views of the model. The acquired model is rendered by the view of a virtual camera.

View synthesis of arbitrary viewpoints plays an important role in virtual reality and augmented reality due to its stunning realism for the acquired images, but current technology is still evolving. The integration of the sense of touch (haptics) with 3D environments such as 3D video requires a comprehensive physical model of the 3D environment. Deformable models provide more realistic force feedback for use in virtual environments with haptics. Several approaches focusing on virtual view generation and deformable models have been proposed, as we will discuss in the following sections.

A. Virtual View Generation

View synthesis has been a very important research field for the development of virtual reality applications. Visual reconstruction approaches can be classified in two important classes of research [14].

The first class is image-based model, which employs morphing or warping techniques. The main feature of these methods is that a set of images is used to interpolate intermediate views. Pixel dependency or image flow characterizes this class. Amongst the research carried out using the image-based techniques are the following: The light field rendering technique based on the idea of generating new views from arbitrary camera positions by combining and re-sampling the available images. [10]. A similar method was developed in the work titled the Lumigraph [3] which overcome some limitations from the previous approach such as a depth correction extension for accurate reconstruction from limited number of source cameras. Both methods are based on a dense set of images on a plane, where a volume occupancy method is used to retrieve the geometry from multiple viewpoints. Another example is the linear interpolation approach, to generate intermediate views with a need for pixel correspondence [2]. Other approaches use stereovision techniques. For example, McMillan employs a Plenoptic function describing all information that can be perceived from an observer's point of view [13]. His approach, however, cannot be extended to real-time applications due to its planar views dependency.

The second class is called model-based. The methodology involves basically a there stage process. The first step is to recorder a scene of the model from some fixed views. This is mainly achieved by a set of cameras located strategically to record particular scenes of the object to be constructed. Then the geometry of the model is recovered according to the stereo-vision technique. Finally the model could be rendered from any particular view.

Research carried out using the model-based techniques are the following Wurmlin et al present a framework, which involves data structures for acquisition and construction of a 3D model with a view independent time-varying video [19]. A codec system is based on viewpoint interpolation from a set of equidistant viewpoints [18]. This latter approach is static because it cannot generate random views from the scene. Finally, Virtualized reality or the sense of complete immersion, a concept introduced by Kanade [7], consists of a set of real scenes and stereo process for the generation of depth map information.

B. Object Modeling

In science, modeling is used to recreate a phenomenon. This can be done through a variety of scenarios. In the branch of computer graphics the modeling concept could be defined as computational representation of objects, which can have dimension in the real world, and is best defined as object modeling.

Achieving more accurate models implemented in virtual environment is still a huge challenge. There is a broad range of modeling methodologies for representing physical and non-physical objects. Those approaches vary from volumetric representations, boundary representations, constructive solid geometry (CSG), finite element analysis and implicit function techniques.

Interaction between virtual environment and haptics systems demands models with special features or behaviors such as dynamic global and local deformations. Cani-Gascuel and Desburn [1] explain that deformations can be elastic or inelastic based on the recovery shape after removing the applied forces. They also illustrate that most deformable models are based on nodal approaches; these can be computationally very expensive. Global approaches reduce computational costs; however they have some deformation limitations [1].

Deformable Super quadratics (DeSuq) provides a full and compact geometric representation of an object with few parameters. Local and global deformations are transformed to physics-model domain according to Lagrangarian dynamics [metaxas]

Another approach that can be considered is the biomechanical model, which is based on linear elasticity and finite element theory. This model simulates the deformations of materials having a different behaviour in a given direction that include the notion of anisotropic deformation. This concept is applied to deformable models to represent various forms of interaction with a surgical tool, such as sliding, griping, cutting, [16].

Recent research carried out by Hua makes it possible to handle complicated and arbitrary topologies [5]. Such technique merges implicit functions, parametric representations and physics-based modeling to produce a haptic-based framework. Implicit functions are mappings from a 3D space to real numbers. The points (x,y,z)that represent a surface in a 3D space can be defined by the function f(x,y,z)=0. Such a function defines what is inside or outside of the surface if the function is less than zero or bigger than zero respectively. This property makes implicit functions suitable to perform collision and detection process [1].

The main goal of this research is to integrate object modeling with haptic rendering based on free viewpoint 3D video generation. The potential applications are: teleconferencing, education and medicine, which require the use of more realistic models for surgical training students. Another potential research area is gaming and the entertainment industry. In sites such as museums and archeological areas, virtual reality will be enhanced by more realistic virtual human interaction. Hence, the need of dynamic models is evident.

III. 3D FREE VIEWPOINT VIDEO WITH HAPTICS INTERACTION

Current technologies and approaches of view synthesis and dynamic modeling have been mainly individual problem solutions. This range of solutions can provide the architecture for a new integrated functional platform. The challenge is to integrate those methodologies over the same pipeline process in order to support different applications.

The idea has to be complemented by looking closer to current software engineering models and add new components based on existing approaches to extend these available applications. In addition the framework has to be designed for future extensions.

A. Framework Overview

The basic concept of this proposed architecture is to break down the haptic rendering, 3D modeling and graphical display areas of research into corresponding components. The main requirements to consider during the design of our framework, which deals with volume data visualization and haptics interaction, are:

- Haptic update rate, which is the simulation frequency normally of 1KHz for stable, responsive haptic feedback
- Efficient Contact detection analysis to compute collision impulse and contact forces between the physical interface and the virtual object
- Efficient data updating in order to modify visual and sense of touch interaction, and
- Coherent volumetric data and haptic rendering scheme so that the system can feel rigid and deformable models

These requirements can to be handled separately. In our study, we propose a control system to integrate the different areas of investigation and to manage the overall flow of information between the different components as shown in Figure 1.

In the following we will describe each of the components of our architecture in more details.

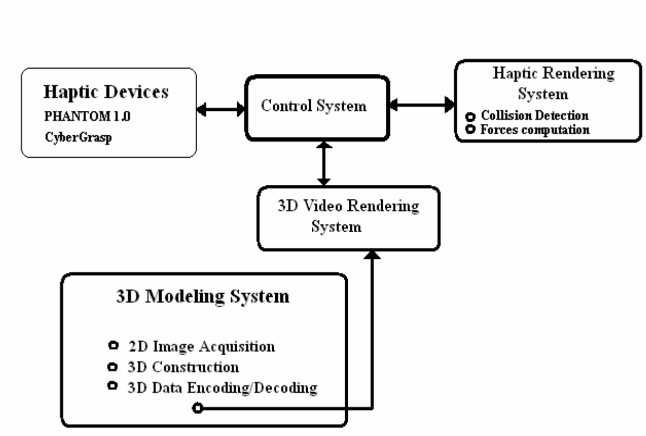

Figure 1: Overall architecture of free viewpoint 3D video system with haptic interaction.

B. 3D Modeling Scheme

The construction of the 3D model begins by capturing a 2D pixel images sequence by the Virtual Viewpoint Real-Time System manufactured by Zaxel Systems, Inc. [20]. Arbitrary views of the object are then generated from the captured information following an image-based visual hull approach similar to the one suggested by Würmlin [Wurmil02]. The information based on a hierarchical spatial structure is encoded and processed in order to obtain a more compact and robust 3D data. The overall pipeline of this component is depicted in Figure 2.

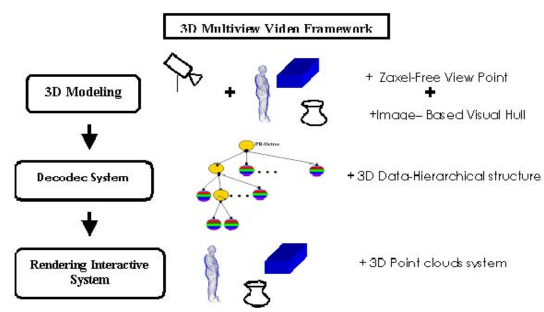

Figure 2: Free Viewpoint 3D Video Framework.

The Virtual Viewpoint Real-Time system consists of 12 digital IEEE 1394 cameras connected to a three server PCs. These PCs are capable of capturing synchronized video from 12 cameras and to process that captured video and transmit it in real-time to a fourth computer, which we refer to as "control PC". Videos can be captured and stored on, e.g., the hard disk so that they can be processed offline.

The Virtual Viewpoint System constructs a coarse 3D model, which relies heavily on the ability to separate foreground objects from the background surrounding. From the information provided by Zaxel system and according to visual-based approach [19] the 3D geometry of a particular object can be approximated using point primitives, and then be rendered from any view. Characterizing surfaces by point clouds supports efficient re-sampling process for intense geometric deformations and topology variation. In addition there is a trend in the computer graphic field to present densely sampled shapes originated from 3D acquisition. Following this approach pioneered by Levoy [9] large set of data could be efficiently rendered with a single data structure based on a hierarchy scheme. Then this hierarchical structure should be adapted in order to accommodate a dynamic set of data, applying the point region quadtree data structure [wurmil02]. In this approach an efficient and progressive encoding procedure has been developed in order to transverse and store the tree structure. This technique presents the foundation of the present research work.

C. 3D Video Rendering System

The rendering interactive stage provides the information for a playback process of the 3D object, which is the final step of this framework, as can be seen in Figure 2. The final 3D constructed model could be treated as an ordinary 3D graphics object with mechanical properties. The prototype can be displayed by using standard graphic libraries such as opengGL. These libraries should be implemented in an application software interface designed to provide some video functionality. This feature controls the frames and their resolution as well as scene viewpoint selection, pausing, amongst others. This framework serves also as an interface between our layered 3D model and the haptic device. The graphical representation of the 3D model is the main issue to implement this haptic rendering recipe.

D. Haptic Rendering component

Haptic Rendering component is the interface between the haptic device and virtual environment. It provides the realistic interaction between the real world and the virtual world. Haptic rendering is a well-studied process, which mainly involves the following steps: contact detection, response force computation, and force calculation display. Based on these principles an overall architecture is proposed to cope with these tasks.

Most research about haptic rendering has been focused on polygonal models. Less research has been oriented to other 3D model representations, such as implicit surfaces and constructive solid geometries (CSG). In our case the model is obtained from view synthesis approach. Cloud points structure represents the nature of this information. It needs to be adjusted by some methodology, for example implicit surface technique, in order to describe the model as implicit surfaces.

The interaction between the 3D model and the haptic device is the first task to be accomplished. Implicit surface representation is based on implicit functions, which map the 3D space to the real numbers domain. These functions present an attractive property, which defines what is inside and what is outside [11]. Following this feature, collision detection algorithm is implemented according to [8] and [1].

The implicit formulation of the model is an implicit isopotential surface generated by a set of skeletons which can be any geometric primitive such as points, curves, parametric surfaces, etc. [1]. This implicit surface defines an extra layer on the top of the base structure of a deformable or rigid object, which exploits the inside/outside property to facilitate the collision detection process. The next step is to compute the response forces, which are obtained by the calculation of the surface normals of the neighborhood around the tool tip and then finding the force direction, in similar way [8]. The force magnitude is then calculated in accordance with a spring-damper dynamic model proposed by A.Yoshitaka [Yoshitaka]. This approach computes the force vector in a way that maintains the contact between the virtual object and the haptics tip at the same position.

E. Control System

This component is the core of this approach. Here the integration tasks of the free viewpoint 3D video system and the haptics rendering process are performed. The haptics interaction process starts after a 3D model image has been displayed through the conventional PC monitor.

The main tasks performed by this component are: firstly to recover the current position and orientation of the haptics device by the implementation of a set C/C++ based libraries. Secondly the control system will check if volume data has been modified through the haptics rendering process in order to update the 3D model into the device display. If a collision or interaction between haptics and virtual object occurs a process for updating such data will be called, according to the Haptics rendering formulation and data structure representation of the 3D object. As final step into the haptic-3D interaction loop, the physical display device needs to be refreshed. This is done by the Haptics rendering component, which updates all surface components into the graphical representation of the object. This provides a force feedback or deformation according to the type of object tested. The control system also needs to be designed to cope with Haptics rendering issues such as latency and to be oriented to work in a collaborative environment.

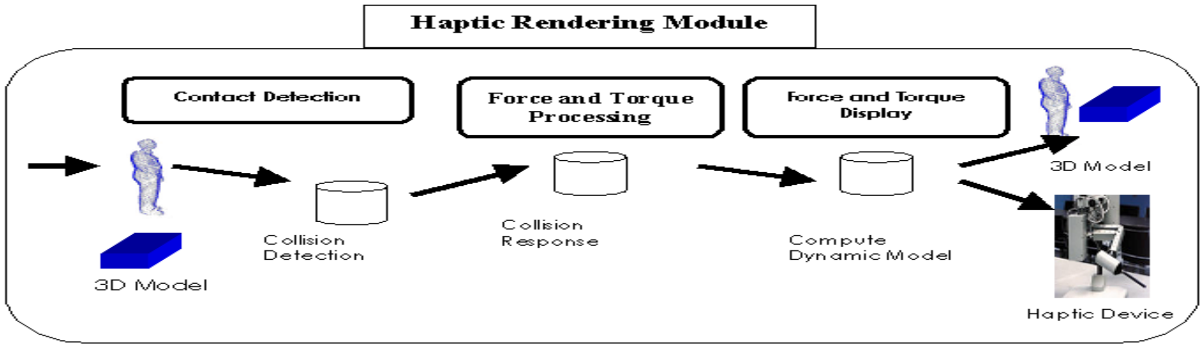

Fig. 3 Haptic Rendering Module architecture

F. Haptics Devices

Haptics devices like the PHATOM Desktop and the Cyber Grasp are used in this prototype. They are integrated in the overall architecture. The mechanism of this integration is through the implementation of the required libraries in C/C++ languages. These libraries permit the interpretation of the geometry of the Haptics scene and also allow the communication between the Haptics devices and the free viewpoint 3D video framework. The implementations of these libraries follow the device drivers and interface functions provided by the software library of the PHATOM and the Cyber grasp technologies.

This proposed architecture has as primary goal to test an improved interaction between current Haptics technologies and real models.

IV. IMPACT AND APLICATIONS

The framework proposed will represent a full visual and haptic feedback system. It is intended to be applied in the medical and educational fields. Medical students would have an available practice tool, with more realistic models, improving their learning experience. Other benefits are: reduction in practicing with animals and pre-operative sessions. A virtualized work place could be used to train workers and new employees. In simulation of emergency situations could help experts to design more efficient safety places of work.

V. CONCLUSIONS AND FUTURE WORK

This paper proposes an idea for developing a framework that can integrate current methodologies and explore new research approaches. It is to build a robust architecture in the field of virtual reality and augmented reality based on haptics interactions for training purposes. Although at this stage the research is still premature the idea is becoming more consolidated through the initial stages of validation and testing. It is also intended that this proposal architecture could serve as prototype for more researchers to develop and carry out their ideas.

VI. REFERENCES

[1] M. Cani-Gascuel, M. Desbrun. Animation of Deformable Models Using Implicit Surfaces. IEEE Transactions on Visualization and Compuetr Graphics, Vol. 3 No. 1 March 1997.
[2] S.E Chen and L. Williams. View Interpolation for Image Synthesis. In proceedings of SIGGRAPH 93, Los Angeles, CA, USA, Aug. 1993
[3] S. J. Gortler, R Grzeszczuk, R. Szeliski and M. F. Cohen. The Lumigraph. In Proceedings of SIGGRAPH96, New Orleans, Louisiana, USA 1996
[4] J. Hua, H. Qin. Haptic sculpting of Volumetric Implicit Functions. In Proceedings of 9th pacific Conference on Computer Graphics and Applications. 2001
[5] J. Hua, H. Qin. Haptic-based Volumetric Modeling Using Dynamic Spline-based Implicit Functions.
[6] Immersion Corporation http://www.immersion.com/products/3d/intereaction/hardware.shtml
[7] T. Kanade, P. Rander and P. Narayanan. Virtualized Reality: Constructing Virtual Worlds from Real Scenes. IEEE Multimedia 1997
[8] L kim, A. Kyrikou, G.S Sukhatme, M. Desbrun. An Implicit Haptic Rendering Technique. Proceedings of the 2002 IEEE/RSJ. EPFL. Lausanne, Switzerland October 2002
[9] M. Levoy. QSplat : A Multiresolution Point Rendering System for Large Meshes.Proceeding SIGGRAPH 2000.
[10] M. Levoy, P. Hanrahan. Light Field Rendering. In Proceedings of SIGGRAPH96, New Orleans, Louisiana, USA 1996
[11] Ming.C. Lin and Stefan Gottschalk, "Collision detection between geometric models: a survey", the Proceedings of IMA Conference on Mathematics of Surfaces 1998.
[12] T. Matsuyama, X. Wu, T. Takai and S. Nobuhara. Real-Time Generation and Hihg Fidelity Visualization of 3D Video. IEEE Transactions, Volume 14, Issue 3, March 2004.
[13] L. McMillan and G. Bishop. Plenoptic Modeling: An Image-Based Rendering System. In Proceedings of SIGGRAPH 95. Los Angeles, California, August 1995.
[14] S. Moezzi, L.Tai and P. Gerard. Virtual View Generation for 3D digital Video. IEEE Multimedia 1997.
[15] SenAble Tehcnologies Inc. http://www.sensable.com/haptics/products/phantom.html

[16] G. Picinbono, J.C Lombardo, H. Delingette and N. Ayache. Improving realism of a surgery simulator : Linear anisotropic elasticity, complex interaction and force extrapolation. The Journal of Visualization and Computer Animation 2002.
[17] R. Ramanathan, D. Metaxas. Dynamic Deformable Models for Enhanced Haptic Rendering in Virtual Environments,. Virtual Reality, 2000. Proceedings.IEEE,18-22 March 2000.
[18] B.L. Tseng and D. Anastassiou. Multi-viewpoint video coding with MPEG-2 compatibility. Columbia University, New York, N.Y. IEEE Transactions on Circuits and Systems for Video Technology
[19] S. Wurmlin, E. Lamboray, O. G. Staadt and M. H. Gross. 3D Video Recorder: a System for Recording and Playing Free-Viewpoint Video. In Proceedings of Pacific Graphics 2002.
[20] Zaxel Systems, Inc. Virtual Viewpoint,
http://www.zaxel.com/virtualviewpoint/index.html