Tele-immersive Gaming for Everybody*

Ahsan Arefin Zixia Huang Raoul Rivas Shu Shi Wanmin Wu Klara Nahrstedt
University of Illinois at Urbana-Champaign
Urbana, IL
{marefin2, huang21, trivas, shushi2, wwu23, klara}@illinois.edu

ABSTRACT

In this demonstration, we present two 3D tele-immersive games: *light-saber dual* and *block fencing* that merge 3D video representations of participants in real-time to enable remote interactions in a virtual world. The light-saber dual arranges participants in a symmetric setup where both participants interact with each other in a virtual world with similar goals. On the other hand, the block fencing creates an asymmetric setup where participants interact with virtual objects having different goals. Using these two setups, we address the challenges and novelty of our solutions in portable environment setup, data acquisition, multi-stream synchronization, multi-stream session management, mobile device rendering, and overlay communication in the design and implementation of advanced 3D tele-immersive systems.

Categories and Subject Descriptors

H.5.1 [**Multimedia Information Systems**]: [Video, Audio input/output]; C.2.4 [**Distributed Systems**]: [Distributed Applications]

General Terms

Algorithm, Design, Performance

Keywords

Tele-immersion, Gaming, Portability

1. INTRODUCTION

The last decade has witnessed the rising of multi-sensory, multi-modal distributed tele-immersive (TI) environments, which are much more advanced and media-enriched than traditional audio-visual interactive systems such as Skype, NetMeeting, Yahoo Messenger, QQ and others [3]. By capturing full-body human motion in 3D and exchanging the video data over the Internet, these systems can render 3D photo-realistic representations of remote users into a joint virtual space at interactive rates [4]. Technologies are now advancing towards collaborative interactive environments equipped with large number of correlated media sensors, multi-view displays, and advanced haptic-audio-visual user interfaces. They generate highly correlated multi-sensory contents (bundle of streams [1]) that require to be captured in a synchronous and real-time manner, distributed on a view base demand [5] and presented at output device(s) in a real-time and synchronous manner [3]. Current available single-stream based session management and synchronization protocols are not quite adequate to meet these requirements [3].

*First five authors are in alphabetical order

Copyright is held by the author/owner(s).
MM'11, November 28–December 1, 2011, Scottsdale, Arizona, USA.
ACM 978-1-4503-0616-4/11/11.

In this demonstration, we showcase our architecture, design and implementation towards building an advanced 3D tele-immersive environment meeting the real-time requirements. We present two real-time 3D interactive exer-gaming applications: *light-saber dual* and *block fencing* showing the novelty of our solutions in stream synchronization, mobile device rendering, and multi-stream session management, monitoring and adaptation. We introduce haptic sensors in the TI environment to provide game-specific sensing to the participants. In order to understand the performance of our solutions, we provide a visualization of the running session showing the QoS values at each part of the system. Furthermore, to support wider deployment of such 3DTI applications in open areas or home environments, we address the challenges in setting up a portable TI system.

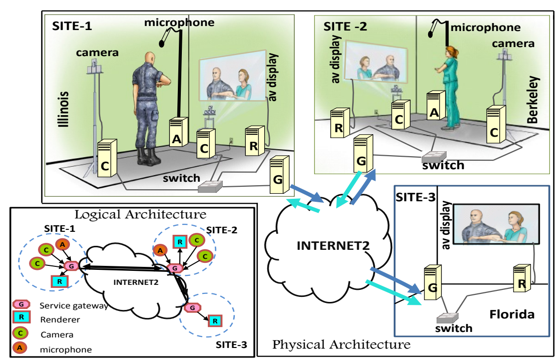

Figure 1: A 3D Tele-immersive Environment with Three Sites.

2. SYSTEM DESIGN

Figure 1 shows an architecture of our portable 3DTI environment with three tiers. In the capturing tier, several 3D cameras are used for capturing. Their host PCs grab the 2D video frames and perform 3D reconstruction in real time and then send the 3D video data to a Rendezvous Points (RP) called gateway node. All cameras are time synchronized by a trigger to capture images at the same time. In the transmission tier, the data is transmitted to/ from remote sites depending on user subscription to specific views. Remote communications are done via gateways. In the rendering tier, gateways send received streams to the local render to display. The inter-stream synchronization is done at the receiver-side gateway.

3D Video Acquisition. To improve portability, we use portable hardware and software components in 3D video acquisition. On the hardware side, we use the stereo camera BumbleBee2 from Point Grey Research Inc. They capture higher resolution images with one-time calibration. A new mesh based 3D video representation [2] is built based on the depth images generated by stereo cameras using point-cloud, which not only improves the image quality but also reduce the required transmission bandwidth to a few Mbps compared to other existing approaches.

Universal Gateway Architecture. The gateway architecture is uniquely designed as a distributed user-level kernel that features bundle of streams as first class object, provides a unified interface for end devices, provides separation of policy and mechanism, and allows dynamically loadable user-defined functions (such as compression, congestion control, encryption, and synchronization).

Inter-stream Synchronization. To achieve the inter-stream synchronization, we synchronize all the sensory devices in the capturing side periodically using NTP, and timestamp every produced media packets. At the rendering side, we develop a coordinated protocol across the received multi-stream from each site, and adapt the receiver buffer based on the media packet timestamp to achieve synchronization, which eventually satisfies 80msec skew threshold under Internet dynamics.

Mobile Rendering. Integrating mobile computing devices with 3DTI environments to receive and render 3D video streams in real-time are challenging due to their limited computation capability and low data rate of wireless networks. To support this, we have developed a remote rendering framework that converts 3D scene to 2D image frames based on user's current view and transmit them to mobile devices in real-time. To accommodate view change, it uses a 3D warping-based algorithm that responds to user control request locally on mobile devices to avoid the round-trip time of client-server communication.

Session Monitoring. To enable run-time session monitoring, we use an unobtrusive framework to capture and store system and application metadata (frame size, frame rate, reconstruction time, CPU load, etc.). The purpose of session monitoring is to allow applications as well as system administrators to ask run-time queries, understand system characteristics, metadata dependency, detect abnormal behaviors and adapt accordingly.

3. DEMONSTRATION

To show the novelty of our solutions, we present two interactive exer-gaming applications: *Light-saber Dual* and *Block fencing*.

Light-saber Dual. The light-saber duel game was designed to take full advantage of remote interaction. Figure 2 illustrates how the game is played. Participant 1 in Site-1 puts on a lab coat with red patches and takes a green light-saber, while Participant 2 in Site-2 puts on a coat with green patches and holds a red light-saber. Each participant then tries to use the light-saber to hit the opponent's color patches in the virtual space as much and as fast as possible in order to gain points. A collision detection module is developed in the renderer to automatically detect the hitting. Once a successful hit occurs, the sword and patches turn blueish for an electrifying effect; meanwhile, the game points, represented on a bar of the light-saber's color, get updated. Also the participant who is hit feels a haptic feedback (vibration) on his/her sword and the participant who hits see a lightning in his/her sword. Both players have identical goals and roles in the game (symmetric).

Block Fencing. This game is asymmetric, where the mobile device plays an important role. Figure 3 shows how the game is played.

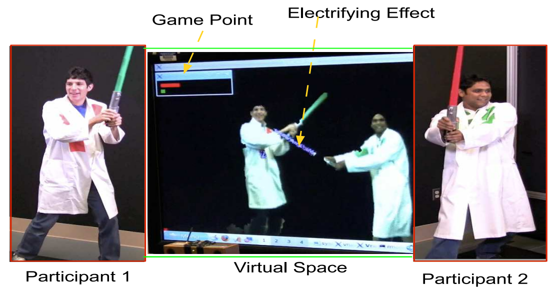

Figure 2: Virtual Light-saber Dual

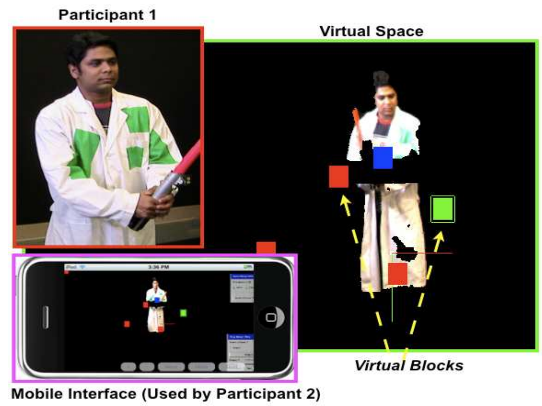

Figure 3: Virtual Block Fencing

Participant 1, captured by a 3D camera in Site-1, can see himself/herself in the virtual space together with a number of graphical red blocks when the game starts. The goal for him/her is to use a physical sword to "touch" all the blocks in the virtual space as fast as possible. As an additional challenge, the touching has to happen in a particular order - the next block to touch will turn from red to green as an indication for Participant 1. If the indicated block is successfully hit, it turns blue. Participant 2 (not shown in Figure 3) uses an iPhone (or iPod) to define the layout and touch-order of the blocks for participant 1. He/she can use a block controller to move the blocks in the virtual world before Participant 1 starts, as well as watch in real-time the actions of Participant 1 on a mobile 3D video viewer. An iPhone is used because of its intuitive touch screen interface. Buttons are design in the 3D video viewer for Participant 2 to change the number of blocks or reset the game.

4. REFERENCES

[1] P. Agarwal, R. Rivas, W. Wu, K. Nahrstedt, and A. Arefin. Bundle of streams: Concept and evaluation in distributed interactive multimedia environments. In *Proc. of ISM*, 2010.

[2] G. Kurillo, R. Vasudevan, E. Lobaton, and R. Bajcsy. A framework for collaborative real-time 3d teleimmersion in a geographically distributed environment. In *Proc. of ISM*, 2008.

[3] K. Nahrstedt, Z. Yang, W. Wu, A. Arefin, and R. Rivas. Next generation session management for 3D teleimmersive interactive environments. *MTAP*, 2011.

[4] W. Wu, R. Rivas, A. Arefin, S. Shi, and K. Nahrstedt. MobileTI: a portable tele-immersive systems. In *Proc. of ACM Multimedia*, 2009.

[5] W. Wu, Z. Yang, and K. Nahrstedt. Dynamic overlay multicast in 3D video collaborative systems. In *Proc. of Nossdav*, 2008.