Immersive 3D Environment for Remote Collaboration and Training of Physical Activities

Gregorij Kurillo* Univ. of California, Berkeley

Ruzena Bajcsy† Univ. of California, Berkeley

Klara Nahrsted‡ Univ. of Illinois, Urbana Champagne

Oliver Kreylos§ Univ. of California, Davis

Abstract

In this paper we present a framework for immersive virtual environment intended for remote collaboration and training of physical activities. Our multi-camera system performs full-body 3D reconstruction of human user(s) in real time and renders their image in the virtual space allowing remote users to interact. The paper features a short overview of the technology used for the capturing and reconstruction. Some of the applications where we have successfully demonstrated use of the system in combination with the tele-immersive virtual environment are described. Finally, we address current drawbacks with regard to data capturing and networking and provide some ideas for future work.

Keyboards: 3D reconstruction, immersion, real-time systems, remote collaboration, tele-immersion.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, Augmented, and Virtual Realities; I.4.10 [Image Processing and Computer Vision]: Image Representation—Volumetric

1 Introduction

Many applications of immersive virtual reality [3, 8] feature avatars to represent the human user inside the computer generated environment. An avatar is often designed as a simplified version of the user's physical features, while its movements are controlled via tracking of the user in the real world. When using such limited number of markers, only partial information on user's movement can be obtained while the actions of the avatar in the virtual environment (VE) have to be simulated through complex dynamics equations [3]. Due to the limitations of the model appearance and movement ability, the interactive experience with other users is much more restricted. In addition, the users are required to wear tracking equipment (e.g., body suit with markers, tracking devices etc.) which may interfere with their movement. In contrast to avatars, only full body 3D reconstruction can realistically represent the user's appearance and full dynamics of movement, including subtle movements such as facial expressions, chest movement during breathing and movement of hair or clothing.

Several attempts have been made in the past to develop real-time 3D reconstruction systems to capture the human body [7, 1, 2]. Human body can be captured using multi-camera system which allows extraction of depth information. Based on the computational approach to obtain 3D information in real-time using the vision technology, different methods have been proposed: (1) silhouette-based reconstruction via visual hulls, (2) voxel-based method with space sampling, and (3) image-based reconstruction with dense stereo depth-maps obtained from slightly displaced views.

*e-mail: gregorij@eecs.berkeley.edu
†e-mail: bajcsy@eecs.berkeley.edu
‡e-mail: klara@cs.uiuc.edu
§e-mail: kreylos@cs.ucdavis.edu

IEEE Virtual Reality 2008
8-12 March, Reno, Nevada, USA
978-1-4244-1971-5/08/$25.00 ©2008 IEEE

In this paper we present a multi-camera system based on dense stereo mapping that allows 360-degree full-body 3D reconstruction from images using 48 cameras [2]. The captured 3D data are transferred using TCP/IP protocol to the rendering computer which combines point clouds obtained from multiple stereo views into a real-time 3D model. The data can be displayed locally or transferred to a remote site. The presented framework has been successfully used in remote dancing applications in collaboration with University of Illinois at Urbana Champagne [5] and learning of Tai Chi movements with a pre-recorded teacher [6][1]. The presented tele-immersive framework has also been tested in the remote manipulation of a virtual object between two users at UC Berkeley and UC Davis.

2 Overview of Technology

Our tele-immersion apparatus consists of 48 Dragonfly cameras (Point Grey Research Inc, Vancouver, Canada), arranged in 12 clusters to cover 360 degree view of the user(s). The arrangement of cameras allows usable workspace coverage to about 2.0 x 2.0 x 2.5 m^3. Each cluster consists of three black and white cameras intended for stereo reconstruction and a color camera used for texture acquisition. The cameras of each cluster are connected to a dedicated personal computer (dual or quad Intel Xeon, 3.06 GHz, 1 GB RAM) which collects images and performs 3D reconstruction. To synchronize image acquisition on all 48 cameras, the cameras are triggered by external triggering. For 3D reconstruction algorithm the image is first resized and rectified. Next, the background is subtracted from the image and the edges of all foreground objects are extracted to identify regions with similar features (i.e., texture). Region based correlation and triangulation is performed to obtain the depth map. The acquired depth map is combined with the color map and sent through the network using TCP/IP protocol.

The reconstruction algorithm has been paralleled to exploit the multi-core technology allowing us the frame rate of about 5-7 FPS. The size of each data package depends on the image coverage and ranges from 25 to 50 KB/s. Data streams from all the clusters are combined into a 3D model inside the VE using point-based rendering. The renderer can also receive data streams from a remote site allowing a remote user to be present in the same virtual space as the local user. In addition the users can share different synthetic objects or data which could be manipulated through the network. 3D video can also be recorded and played simultaneously with real-time data.

3 Applications of Technology

In this section we provide short description of several applications where we have demonstrated practical use of our multi-camera system in combination with the tele-immersive VE.

3.1 Tai Chi Learning

In this study we have examined learning of Tai Chi using immersive VE as compared to learning from 2D video. We have captured three moves performed by a Tai Chi teacher. His 3D recording was projected into the virtual space simultaneously with the real-time data from a student. The students saw themselves on the screen

[1]Movies and images available at: http://tele-immersion.citris-uc.org

with the teacher as shown in Fig.1(a). In the study [6] we have demonstrated that immersive virtual reality provides better learning of physical movements than a two-dimensional video. Similar approach could also be used for rehabilitation where a therapist's movement were recorded or streamed live while a remote patient would try to perform the exercises.

3.2 3D Remote Interaction

For remote interaction experiment Fig.1(b) we have streamed 3D video data from UC Berkeley to UC Davis through TCP/IP connection. The VE on both sides featured a floating 'jello'-like object which dynamically deformed when pushed or pulled in any direction. The position of the two remote users' hands was tracked using electro-magnetic tracking system and this data was transferred in a separate stream between the two sites. The display was rendered inside a Cave environment at UC Davis side to feature a life-size rendering of the user from UC Berkeley (Fig.1(b)). The two dislocated users were able to smoothly interact with the featured object. The experiment demonstrated feasibility of remote interaction with 3D objects inside our tele-immersive environment.

3.3 Tele-immersive Dancing

Projection of dislocated digitized users into the same virtual space in real time offers variety of use in art and dance (Fig.1(c)(d)). Over the past two years we have conducted several local and remote experiments [5] with dancers. Our 3D reconstruction does not require the dancers to wear any markers or body suits allowing them to freely move in the space. The system can capture dynamics of hair and clothing without relaying on models. The dancers had to accommodate to technical limitations of the system, such as frame rate and transmission delay. At the same time they introduced new techniques to take advantage of the technology, such as digital transformations (e.g., rotations), disappearing (e.g., moving beyond stereo range), dissolving into 'particles' (e.g., moving too close to the camera). Tele-immersion environment also introduced the novel concept of 'virtual touch' where the feedback for touch relays on visual information from the VE.

4 Conclusion and Future Work

One of the major bottlenecks of our current system is the speed of the reconstruction and the hardware requirements needed for this purpose. Currently, we are upgrading our system with an improved and more robust stereo reconstruction algorithm using finite elements and triangular wavelets which will produce less noisy point clouds with higher frame rate (expected 15 to 20 FPS). We are also evaluating different approaches to calibration of the cameras to make the reconstruction more accurate and robust. Furthermore, we are investigating requirements for a portable version of the tele-immersion system, such as reduced computer power, algorithms adaptive to environmental changes, easy-to-use calibration methods.

Our initial research work [5] has been focused on artists as users of the tele-immersive technology described in this paper. The environment provides different digital effects (e.g., transformations and deformations of rendered images) that can be applied in real-time to manipulate what is displayed to the audience. The presented multi-camera system offers new possibilities for learning and training individuals to perform physical movements (e.g., physical therapy and exercise) [6]. 3D data captured by our system can be used to analyze subject's movement. Extracted kinematic data can be applied as a feedback to the user during learning of motions [4]. The users can be immersed inside computer generated existing or non-existing environments, such as ancient buildings and future architectural designs to allow interactive exploration. The system could be combined with head-mounted display and head tracking to provide higher level of immersion. Tele-immersion in 3D can further enhance video conferencing for business and personal meetings. The users can share different synthetic objects inside the same virtual space, such as CAD designs, laser scanned objects or other type of data sets. Finally, there are many applications of social networking and entertainment where the users could interact in real-time inside a common VE, such as games supporting physical interaction, interactive music video, 3D karaoke.

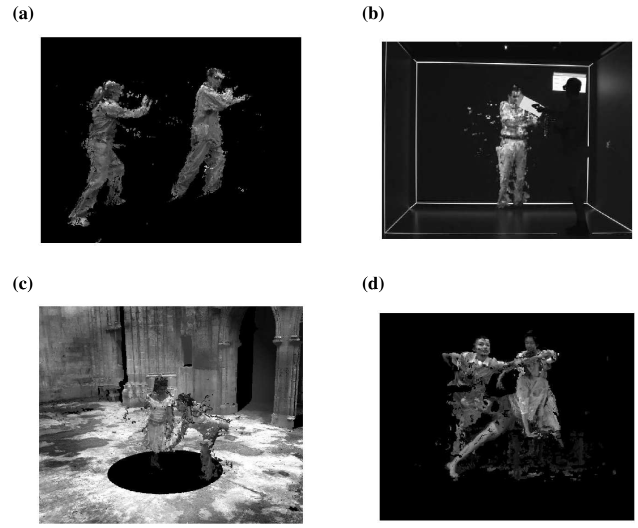

Figure 1: (a) Tai Chi student is rendered in real-time with prerecorded 3D video of the teacher. (b) Remote collaboration experiment between UC Berkeley and UC Davis where two subjects were remotely manipulating the same virtual object. (c) Two dancers are projected inside 3D model of a cathedral which was generated by a laser scan. (d) Two dancers located at Berkeley and another dancer located in Illinois are projected in real-time into the same virtual space.

References

[1] J. Hasenfratz, M. Lapierre, and F. Sillion. A real-time system for full-body interaction with virtual worlds. In *Proceedings of Eurographics Symposium on Virtual Environments*, pages 147–156. The Eurographics Association, 2004.

[2] S. Jung and R. Bajcsy. A framework for constructing real-time immersive environments for training physical activities. *Journal of Multimedia*, 1(7):9–17, 2006.

[3] P. Kalra, N. Magnenat-Thalman, L. Moccozet, G. Sannier, A. Aubel, and D. Thalman. Real-time animation of realistic virtual humans. *IEEE Computer Graphics and Applications*, 18(25):42–56, 1998.

[4] J.-M. Lien, G. Kurillo, and R. Bajcsy. Skeleton-based data compression for multi-camera tele-immersion system. In *Poceedings of International Symposium on Visual Computing (ISVC 2007)*, Lake Tahoe, CA, November 26-28 2007.

[5] K. Nahrstedt, R. Bajcsy, L. Wymore, G. Kurillo, K. Mezur, R. Sheppard, Z. Yang, and W. Wu. Symbiosis of tele-immersive environments with creative choreography. In *ACM Workshop on Supporting Creative Acts Beyond Dissemination, Associated with 6th ACM Creativity and Cognition Conference*, Washington D.C., June 13-15 2007.

[6] K. Patel, J. N. Bailenson, S. Hack-Jung, R. Diankov, and R. Bajcsy. The effects of fully immersive virtual reality on the learning of physical tasks. In *Proceedings of the 9th Annual International Workshop on Presence, Ohio, USA*, pages 87–94, 2006.

[7] S. Wurmlin, E. Lamboray, and M. Gross. 3d video fragments: dynamic point samples for real-time free-viewpoint video. *Computers and Graphics*, 28:3–14, 2004.

[8] Y. Yang, X. Wang, and J. X. Chen. Rendering avatars in virtual reality: integrating a 3d model with 2d images. *Computing in Science and Engineering*, 4(1):86–91, 2002.