Bringing real objects, spaces, actions, and interactions into social VR

Misha Sra* Chris Schmandt†

MIT Media Lab, Cambridge, MA USA

ABSTRACT

We present a novel multiuser virtual reality (VR) system where the physical world is used as a template for the placement of walls, furniture, and objects in the virtual world so as to create a correspondence in scale, spatial layout, and object placement between the two spaces. Through this association between the real and virtual world, users are able to walk freely while wearing a head-mounted device, avoid obstacles like walls and furniture, interact with each other and with objects just like they would in real life, bringing us closer to the realization of realistic collaborative virtual environments. Preliminary deployment during our lab's semi-annual open house shows that the system has potential to offer a high degree of presence in VR.

Index Terms: H.5.1. [Information Interfaces and Presentation (e.g. HCI)]: Multimedia Information Systems—: Artificial, augmented, and virtual realities

1 INTRODUCTION

Although current virtual worlds offer high levels of visual and auditory realism, two kinds of sensations from everyday situations are seldom included in their natural form: full-body movement and touch. Upcoming consumer VR devices are starting to allow for full-body movement in living room sized spaces using cameras or laser emitters with corresponding photosensors. However, movement in space introduces the problem of colliding with physical objects. One solution, implemented by the yet unreleased HTC Vive is called the Chaperone which overlays the physical world onto VR on a button click and also warns users as they approach obstacles. In the real world, touch conveys rich information like contact surface geometry, roughness, slippage, and temperature in addition to force feedback which provides information on object surface compliance, weight, and inertia [2]. But, interactions with virtual objects are often mediated by hand-held input devices that provide vibro-tactile feedback. In this paper, we describe our approach to creating a VR system that allows users with full-body avatars to walk in the real/virtual spaces and interact with virtual objects through physical proxies. For testing our system, we created two VR scenes by mapping a hallway to a virtual bridge (see Figure 1) and the open space outside our offices to a floating island (see Figure 2).

Research has shown that a closer connection between the user's mental body model and the virtual body representation can enhance presence [6]. Building upon this idea, users in our system are represented by a full-body avatar that is controlled by natural movements of each person in the real world to increase their sense of being "in" a place. Interaction in the virtual world requires feedback about the state of the virtual body (avatar), and its relationship to the environment [6]. By providing a virtual body that mirrors all real world movements, including grasping objects, we create a direct proprioceptive feedback loop. Our object tracking system allows users to feel touch by employing *passive haptics* i.e., when users touch or manipulate an object in the virtual world, they simultaneously also touch or manipulate a corresponding object in the physical world. All these elements enable users to interact with each other as well as with their environment almost as naturally as they would in the real world. We envision users at home or work, using our system to dynamically create virtual spaces for interaction with remotely located friends or colleagues, creating VR games to play with co-located friends, demarcating movement space that can be used as the physical space for navigation in any VR game or for visiting architectural sites anywhere on the planet.

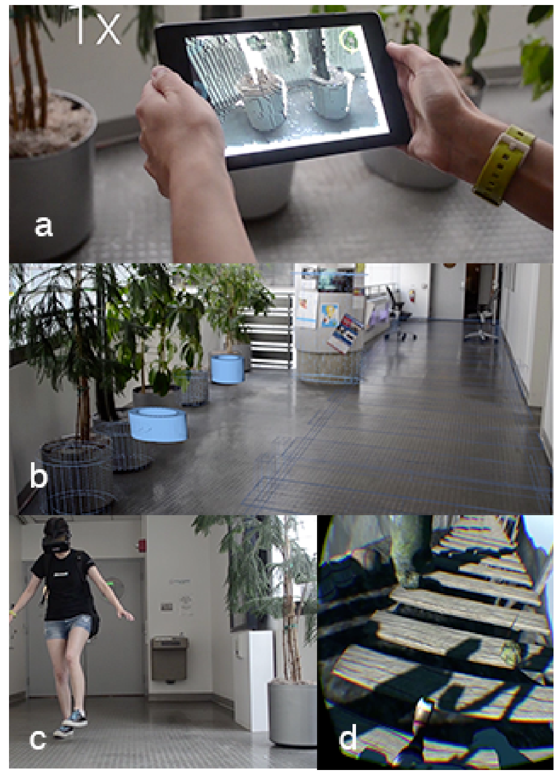

Figure 1: a) 3D scanning real world space using the Google Tango device. A mesh is generated from the scanned point cloud data and used to create the virtual scene. b) Using the mesh as a template, virtual objects like barrels and bridge (in blue) are created and placed in the same position and orientation as the real world planters and hallway. c) User wearing Oculus Rift lifting up their left foot to start walking in the hallway. d) User seeing (first person view through the HMD) their avatar's left foot lift up in sync with their real foot.

2 SYSTEM

We start with a 3D reconstruction of the real world space (see Figure 1a) where we plan to situate the virtual environment, using the Google Project Tango device[1] and Voxxlr[2] (see Figure 1). Tango

*e-mail: sra@media.mit.edu

†e-mail:geek@media.mit.edu

[1]Project Tango: https://www.google.com/atap/project-tango/

[2]Voxxlr: http://voxxlr.com/

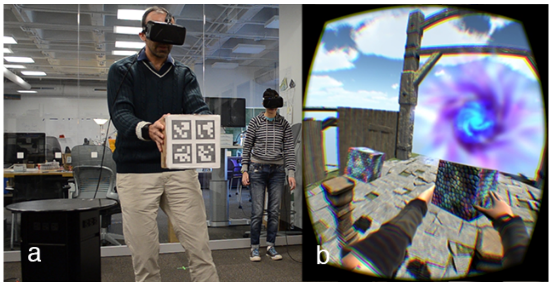

Figure 2: a) A user grasping a physical box while simultaneously grasping a virtual box. b) View through the user's HMD showing their avatar grasping a virtual box. Both user's have full-body avatars that they control through their body movements in the real world.

devices are equipped with integrated sensors that can measure the distance from a device to objects in the real world and provide the position and orientation of the device in six degrees of freedom. Unlike the Microsoft Kinect that needs to be tethered to a computer, Tango is a self-contained device that provides depth and pose data necessary for algorithmically reconstructing the shape of the environment such as the floor, the walls, and other obstacles as we walk around untethered in the physical space. The depth data returned by the device is a point cloud that represents point samples of the surfaces in front of the device. This data is transformed into a triangle mesh by creating a polygonal surface representation of the isosurface [1] followed by polygon reduction and hole filling. The watertight mesh is then edited, textured, shaded and converted into a VR environment that the users can explore.

2.1 Texturing

Previous research has focused on virtual objects being modeled on their physical proxy [4]. We do not want to reconstruct a realistic virtual replica of the physical space. Instead, we want to use the position and size of walls, furniture, and objects as a guide for the placement of similar obstacles in the virtual world such that a correspondence between the two spaces exists (see Figure 1b). Additionally, we want to change the visual appearance of virtual objects by texturing the 3D models differently from their real world textures in order to create a virtual space that looks remarkably different from the physical space (see Figures 1c and 1d). Even though the virtual objects will look different from the corresponding physical objects, we aim to match the physical characteristics of the actual objects in the virtual representation to maintain sensory coherence. For example, an empty cardboard box feels light, non-slippery, a little rough, and has a distinct sound when touched or moved. Our virtual box therefore needs to match some of these tactile characteristics to maintain coherence in sensory feedback (see Figure 2). Additionally, there needs to be an approximate size match or the user may not be able to pick up the real box without some trial and error, due to proprioception. Objects that present similar affordances [3] may be the best candidates for virtualizing and using as passive haptics.

2.2 Body and Object Tracking

Depending on the physical space size, shape and the corresponding virtual experience we wanted to create, we used one or more Kinects to track users and objects. By tracking up to 25 joints representing body parts such as head, neck, shoulders, and arms the user is able to realistically control the avatar body by mapping the 3D coordinates of each joint to the corresponding bones in a rigged 3D model. This allows users to control a full body avatar in the virtual environment through corresponding body movements in the real world. Real-walking in virtual environments is more natural and produces a higher sense of presence than other locomotion techniques [7]. We created two avatars to match the steampunk theme of our floating island to increase visual match between the scene and the user's virtual body [6].

Our approach for interacting with physical objects builds on the concept of *passive haptics* [4], i.e. receiving feedback from touching a physical object that is registered to a virtual object. We use markers to track position and orientation of two cardboard boxes on our server using markers due to their simplicity and robustness over feature-based tracking systems [5]. Given the initialized state (e.g., position and size) of our boxes in an RGB frame from the Kinect, the goal of tracking was to estimate the states of the boxes in the subsequent frames and applying the learned transformations (position and orientation) to their virtual counterparts in our VR scene.

2.3 User Experience

To receive initial feedback, we deployed an initial prototype of our system in our lab space for few hours. A total of 34 participants tried our system, for approximately 8-10 minutes each with generally positive reactions. Most people appreciated being able to walk without fear of colliding with things and seeing a full body avatar instead of disembodied hands of their previous VR experiences. Interacting with boxes was more challenging for people who's height varied considerably from the developer's height. Our initial conjecture is that a mismatch between a user's proprioceptive sense and the size of the avatar's body may be the cause. We are designing an experiment to understand and resolve this observed behavior.

In the setup for the hallway/bridge scenario, two users start at opposite ends of the hallway/bridge and walk towards one another. They end with a high five in the real world which corresponds to a high five by their avatars in the virtual world. Each user is tracked by a Kinect and data is synced using a client-server architecture. In the open space/floating island scenario, both users are tracked by a single Kinect placed in front of them and the RGB camera of the Kinect is used for tracking the boxes. We asked both users to start by looking around the VR scene, looking at their hands, feet, and body followed by acknowledging each others presence in the virtual world by waving to one another. Each user was then asked to walk up to the virtual box on the pillar a few feet in front of them, to pick it up and put it on top of the other box. We believe, based on initial feedback from users, that physical movement, full-body avatar, and interaction with other people and objects helped create a natural and immersive VR experience.

References

[1] P. Bourke. Polygonising a scalar field. *Cupertino: http://paulbourke. net Available from: http://paulbourke. net/geometry/polygonise [Accessed 1 April 2011]*, 1994.

[2] M. Bouzit, G. Burdea, G. Popescu, and R. Boian. The rutgers master ii-new design force-feedback glove. *Mechatronics, IEEE/ASME Transactions on*, 7(2):256–263, 2002.

[3] J. J. Gibson. 1986. *The ecological approach to visual perception*, 1979.

[4] H. G. Hoffman. Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments. In *Virtual Reality Annual International Symposium, 1998. Proceedings., IEEE 1998*, pages 59–63. IEEE, 1998.

[5] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*, pages 85–94. IEEE, 1999.

[6] M. Slater, A. Steed, and M. Usoh. Being there together: Experiments on presence in virtual environments (1990s). Technical report, Technical Report, Department of Computer Science, University College London, UK, 2013.

[7] M. Slater, M. Usoh, and A. Steed. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(3):201–219, 1995.