Immersive Guided Tours for Virtual Tourism through 3D City Models

Rüdiger Beimler, Gerd Bruder, Frank Steinicke

Immersive Media Group (IMG)
Department of Computer Science
University of Würzburg
E-Mail: {ruediger.beimler,gerd.bruder,frank.steinicke}@uni-wuerzburg.de

Abstract: Since decades, computer-mediated realities such as virtual reality (VR) or augmented reality (AR) have been used to visualize and explore virtual city models. The inherent three-dimensional (3D) nature as well as our natural understanding of urban areas and city models makes them suitable for immersive or semi-immersive installations, which support natural exploration of such complex datasets.

In this paper, we present a novel VR approach to leverage immersive guided virtual tours through 3D city models. Therefore, we combine an immersive head-mounted display (HMD) setup, which is used by one or more *tourists*, with a touch-enabled tabletop, which is used by the *guide*. While the guide overviews the entire virtual 3D city model and the virtual representations of each tourist inside the model, tourists perceive an immersive view from an egocentric perspective to regions of the city model, which can be pointed out by the guide. We describe the implementation of the setup and discuss interactive virtual tours through a 3D city model.

Keywords: Virtual environments, virtual cities, guided tours

1 Introduction

In recent years, virtual environments (VEs) have become more and more popular and widespread due to the requirements of numerous application areas in particular in the 3D city visualization domain. Two-dimensional desktop systems are often limited in cases where natural interfaces are desired, for example, when navigating within complex 3D scenes. In such cases virtual reality (VR) systems, which make use of tracking technologies and stereoscopic display of three-dimensional synthetic worlds, support better exploration of complex datasets. These VR systems allow users to explore virtual worlds in an intuitive and immersive manner. In immersive virtual 3D city environments people can visit, for instance, tourist landmarks by natural locomotion in the space provided by the range of tracking sensors, or can explore larger VEs by using 3D input devices.

In this paper, we introduce a first prototype for a novel VR approach to leverage multi-user immersive VR technology for guided virtual tours through 3D city models. This approach is realized by combining several cost-effective technologies and techniques. Therefore,

(a)

(b)

Figure 1: Photo of (a) a guide interacting with the touch-enabled tabletop, and (b) a rendering from within the VE from a tourist's point of view.

we combine immersive head-mounted display (HMD) setups, which are used by the *tourists*, with a touch-enabled tabletop, which is used by the *guide*. The views and movements of the tourists are tracked by a Kinect-based interface and are then streamed and mapped onto virtual avatars within a VE, which are displayed to the guide. The interactions of the guide with the VE are captured via a touch-based interface and applied to the position and orientation of the tourists' avatars. The collaborating users receive real-time audio-visual feedback.

The paper is structured as follows: Section 2 resumes background information. Section 3 introduces our setup and explains how guided virtual tours through 3D city models are performed. In Section 4 we discuss the results. Section 5 concludes the paper and gives an overview of future work.

2 Background

There has been an increasing demand of virtual 3D city representations during the last years for a variety of application fields [RSH05, SHR06, SRH06]. When navigating in an unknown environment, e. g., a foreign city, way-finding tends to be a complex task, which can be supported by prior virtual exploration. Virtual city models provide several advantages over traditional physically crafted city models, especially when they are explored using stereoscopic 3D displays in conjunction with head-coupled perspective rendering [BKLP04]. Burigat and Chittaro [BC07] discuss the importance of vantage points to overview the whole setup, and propose solutions to aid users by providing visual navigation aids. To improve the degree of realism, virtual city 3D models can be populated with virtual humans, which can be animated either by a crowd simulation algorithm or manually by a performer or an animator to simulate natural phenomena of crowd behavior [TGMY09].

Very recently, developments in the field of consumer sensor and display hardware have provided the means to implement cost-efficient multi-user immersive VEs. Examples for tracking sensors are the Nintendo Wii remote and the Microsoft Kinect, as well as display technologies such as the Sony HMZ-T1 or the Oculus Rift HMD. In this context, different multi-user immersive VEs, in which users assume different roles have been realized in the scope of the 3DUI Contest in 2012 [3DU12]. A similar multi-user approach using immersive and semi-immersive setups has been proposed by Beimler et al. [BBS13] for character animations.

3 Guided Virtual Tourism Setup

In this section we describe our collaborative virtual tourism setup. In the following subsections, we present the setups of the guide as well as the tourists and describe their roles.

3.1 Semi-Immersive Guidance Setup

The *guide* takes the role of a "tourist guide", i. e., the guide points out important sights, landmarks and other places of interest to a user that plays the role of a *tourist*. The guide oversees the whole scene from a bird's eye point of view as naturally supported by the tabletop setup (see Figure 1(a)). Since each tourist is represented in the scene by its corresponding avatar, the guide can overview the VE and interact with the scene and the tourists' avatars from the elevated vantage point [BC07]. In our setup both roles, the tourists and the guide are able to communicate with each other, for instance, if there is a spontaneous question of the tourist, the guide can immediately react on requests. In this regard, the tour guide is able to direct the tourists to virtual landmarks, while keeping track of the whole scene.

Setup The physical setup of the guide is based on the *SmurVEbox* approach [FLBS12, BBS13]. The setup consists of a tabletop screen with dimensions of 62cm $\times$ 112cm that displays back projected images over a mirror mounted at the bottom of the box. The virtual scene is rendered on a Windows 7 workstation, equipped with Intel Core i7 3.40GHz processors and an Nvidia Quadro 4000 graphic card. To render the virtual scene on different screens and displays, we made use of the Unity 3D proprietary game engine with the MiddleVR for Unity framework[1].

To provide one- or multi-finger touch capacity on the surface, we made use of rear diffused illumination (Rear-DI) [MTS10]. A cluster of six high-power infrared LED have been mounted to the bottom of the box to illuminate the surface from below. Since the surface consists of a diffusing material, it disperses the light and reflections can be captured with a camera that is equipped with an infrared band-pass filter at the bottom of the setup. To detect touch points captured by the camera, the video stream was evaluated by a modified

[1] http://www.imin-vr.com/middlevr/

(a) (b)

Figure 2: (a) Photo of a tourist in front of a Microsoft Kinect sensor. (b) The tourist's virtual view on the HMD.

version of NUI Group's Community Core Vision (CCV)[2]. The transfer protocol to handle touch information was the TUIO[3] protocol, which is often used with tangible multi-touch surfaces. In addition, we use the community edition of xTUIO's uniTUIO[4] scripts library to stream multi-touch gestures into Unity 3D.

Interaction Since the virtual scene is displayed to the guide as seen from an elevated vantage point on the tabletop, the guide can naturally interact with the miniature representation of the tourists in the minified virtual 3D city environment. To support intuitive and natural interaction, we mapped single-finger pan gestures as well as two-finger rotate gestures to translations and rotations of the virtual avatar, which provides the guide with the ability to move the tourist's avatar to any desired pose within the virtual city (see Figure 1(a)). We built this interaction in two modes: The guide can translate or rotate the tourist's avatar around the scene within the ground plane, and the guide is able to navigate the whole scene through camera panning or rotating in the same plane. We distinguish among the modes by determining the touched object below the user's finger when interacting with the tabletop.

3.2 Immersive Tourist Setup

The tourist is immersed in the guided virtual tour by donning an HMD in front of a Kinect sensor in a room-sized tracked space. The user perceives the virtual scene from a pedestrian point of view, such as if he or she would experience a city in the real world. The user is free to look around and explore the VE. For a virtual self-representation we map the tracked skeleton information from the Kinect sensor to a rigged, bipedal character in the VE.

[2]http://ccv.nuigroup.com/
[3]http://www.tuio.org/
[4]http://www.xtuio.com/

Setup The physical setup is based on an Oculus Rift HMD with a resolution of 640×800 pixels per eye at 60 frames per second. In our current setup, the HMD is connected to the guide's rendering workstation, which renders both the output screens for a tourist and the guide using the MiddleVR framework. For the tourist's avatar we used a rigged, bipedal character, which we adopted from the Unity Asset store, and refined it using the Autodesk Maya software by placing Joint-Deformer Objects to shape a bipedal skeleton (see Figure 1). For the real-time tracking of the tourist's body movements, we use a Microsoft Kinect sensor as illustrated in Figure 2. The Kinect tracks users with an update rate of 30 frames per second. We stream the captured motion of the tourist's skeleton limbs into the Unity 3D engine and map them onto the tourist's avatar. Therefore, we utilized the Kinect for Windows SDK[5] and the Kinect Wrapper Package by CMU's Entertainment Technology Center[6]. As a result, the avatar behaves accordingly to the motions of the tracked tourist. The tourist's avatar has a virtual scene camera attached to his head node, which is used to provide the user with an egocentric view.

Interaction Since the user's body is tracked with the Kinect sensor, and the user's virtual view is slaved to the corresponding head node of the avatar, the user can navigate naturally in the virtual city by real walking. In particular, the user can move and turn towards locations of interest in the virtual scene from a pedestrian's perspective. However, the low-cost tracking currently imposes some limitations on natural interaction. In particular, the accuracy and precision of the tracked head node as provided by the Kinect sensor is very low. Moreover, the Kinect's resolution and our room-sized workspace impose limitations on the size of the virtual space that the user can interact within. In order to travel longer distances, the tourist can make requests to the guide, who then moves the tourist in the virtual city.

4 Discussion

The guided tourism system described in Section 3 represents an early prototype that we developed with immersive and semi-immersive consumer-level VR hard- and software. To this point we have not yet conducted a formal user evaluation of the system to show its advantages over traditional Desktop-based guided tours. However, we observed several benefits and limitations during informal tests of the system.

In particular, we observed that interaction with the virtual city model and the tourist's avatar via the tabletop setup provides an intuitive and direct method to move and guide a user through a virtual scene. We observed that users often interact with both hands using the common multi-touch gestures that we implemented on the tabletop: In particular, users often pan the virtual city view with their non-dominant hand, whereas interactions with the avatar are usually performed with their dominant hand at the same time. While this use of

[5]http://www.microsoft.com/en-us/kinectforwindows/

[6]http://wiki.etc.cmu.edu/unity3d/

both hands seems to provide an intuitive form of interaction, interactions with the avatar could be simplified with additional gestures, such as a single-finger move-to gesture.

On the other hand, we observed that the tourist setup has much room for improvements. While users can use the Kinect-based tracking to change the view on the HMD, the quality of the tracking data is very poor, resulting in many incorrectly identified body poses. More recent developments in the consumer market such as the upcoming revised Kinect sensor may help to alleviate these problems without raising the cost of the system out of proportion for multiple users. To this regard, while the current system has only been tested for one tourist, we see the potential to provide views to multiple collocated or remote tourists.

5 Conclusion

In this paper we introduced a setup for collaborative interactive 3D city exploration. We described a prototype in which we implemented the multi-user approach using consumer-level VR hard- and software. With users assuming the roles of a tour guide or tourists, the system allows users to explore virtual 3D cities in a collaborative manner. We observed that a semi-immersive touch-enabled tabletop environment is well-suited to be used by an interactive guide who views the VE from an elevated perspective, whereas an immersive pedestrian view is provided to a tourist via a HMD. We discussed initial observations of our prototype, and identified benefits and limitations.

For future work, we see much potential for tourists to connect to virtual city environments using consumer-level immersive display and tracking hardware using their home system or setups at interactive installations. For interactive multi-user installations, such as realized in museums or art galleries, we believe that incorporating a tour guide interface to steer and direct immersed users to points of interest will provide many benefits in terms of the tourists' sense of feeling present in an interactive virtual world.

References

[3DU12] 3DUI 2012 Contest. In *Proceedings of Symposium on 3D User Interfaces (3DUI)*. IEEE, 2012.

[BBS13] R. Beimler, G. Bruder, and F. Steinicke. SmurVEbox: A smart multi-user real-time virtual environment for generating character animations. In *Proceedings of Virtual Reality International Conference (VRIC)*, 7 pages, 2013.

[BC07] S. Burigat and L. Chittaro. Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids. *International Journal of Human-Computer Studies (IJHCS)*, 65(11):945–958, 2007.

[BKLP04] D. Bowman, E. Kruijff, J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley, 2004.

[FLBS12] M. Fischbach, M. E. Latoschik, G. Bruder, and F. Steinicke. smARTbox: Out-of-the-box technologies for interactive art and exhibition. In *Proceedings of Virtual Reality International Conference (VRIC)*, pages 1–7, 2012.

[MTS10] C. Müller-Tomfelde and J. Schöning. Building interactive multi-touch surfaces. In C. Müller-Tomfelde, editor, *Tabletops - Horizontal Interactive Displays*, Human-Computer Interaction Series, pages 27–49. Springer, 2010.

[RSH05] T. Ropinski, F. Steinicke, and K. H. Hinrichs. A Constrained Road-based VR Navigation Technique for Travelling in 3D City Models. In *Proceedings of International Conference on Artificial Reality and Telexistence (ICAT)*, pages 228–235, 2005.

[SHR06] F. Steinicke, K. H. Hinrichs, and T. Ropinski. A Hybrid Decision Support System for 3D City Planning. In *Proceedings of International Commission II Symposium (ISPRS)*, pages 103–108, 2006.

[SRH06] F. Steinicke, T. Ropinski, and K. H. Hinrichs. Collaborative Interation Concepts for City Planning Tasks in Projection-based Virtual Reality Systems. In *Proceedings of Virtual Reality International Conference (VRIC)*, pages 113–122. IEEE, 2006.

[TGMY09] D. Thalmann, H. Grillon, J. Maim, and B. Yersin. Challenges in crowd simulation. *Proceedings of International Conference on Cyberworlds*, pages 1–12, 2009.