Using Visual Augmentations to Influence Spatial Perception in Virtual Representations of Real Scenes

Claus B. Madsen, Lisbeth W. Lorentzen
Laboratory of Computer Vision and Media Technology, Aalborg University, Denmark
{cbm@cvmt.aau.dk, lwl@cs.aue.aau.dk}

1. Introduction

This paper presents an approach to enhancing the perception of depth in Virtual Environments based on Image-Based Rendering. We propose that by augmenting scenes with virtual objects and structures we can stimulate the user's desire to perform visual exploration and thus heighten the user's sense of presence.

In traditional 3D model-based Virtual Environments (VEs) users are free to navigate within the bounds of the model. This allows the user to visually explore the scene to get a sense of its spatial layout and composition, exactly as in the real, physical world. The main problem with such model-based VEs is that it is difficult, bordering on impossible, to photo-realistically recreate complicated, real-world locations due to: 1) the complexity of constructing a 3D model of the scene geometry, and 2) the computational complexity of rendering realistic illumination phenomena in real-time.

Image-Based Rendering (IBR) attacks both these problems. IBR synthesizes the user's current view of a scene from a set of pre-recorded images of a real scene. By using IBR users can move around and visually explore a visual recreation of a real scene, and since the visualization is based solely on images the scene can be arbitrarily complex (i.e., there is no 3D modeling involved), and all illumination phenomena are naturally recreated correctly. IBR is by definition photo-realism. But IBR suffers from a catch: with current computer technology it can only provide limited movability in scenes. In our present system the user can only move freely inside a circle with a radius of about 2 feet. Such limited movability makes visual exploration of scenes a little trivial and the amount of motion parallax that can arise from such small movement is limited.

In static scenes motion parallax is loosely speaking the difference in how points in the scene move across the retina as the observer moves [1] [2] Motion parallax is an extra-ordinarily important cue for perceiving the 3D structure of a scene, more important than stereoscopic vision for distances of more than a few feet, and rivaled only by high level information such as a priori knowledge of the natural sizes of recognizable objects.

This paper explores the use of visual augmentations, i.e., the addition of virtual objects to the scene in order to create stronger motion parallax for small ego movements. We propose that scenes can be augmented in two different ways resulting in different explorative behaviors and different perception of the scene.

2. Background

This study has its background in a research project which uses IBR to enable people to visually explore real world places without actually being there. Furthermore our system allows us to augment virtual objects into scenes.

Image-Based Rendering (IBR) is an alternative paradigm to traditional 3D model-based computer graphics. In IBR views of a scene from arbitrary viewpoints are synthesized from data in a large set of images acquired at some location. In our IBR system approximately 400 images are acquired by moving the acquisition camera along the circumference of a circle, with the camera lens point outward of the circle. Currently our setup allows us to acquire images in circle with a radius of 60 centimeters.

From the acquired set of images we can synthesize views of the scene from any point inside the circle. We call this area Region Of Exploration (REX), because it is within this area the user can explore the scene in all directions (full view sphere). The position and viewing direction of the user is tracked with a commercial tracking system, and correct views are presented to the user in stereo at more than 20 frames per second in either a Head Mounted Display (HMD) in a six sided CAVE. In case of the HMD the system runs on a single standard PC, whereas for the CAVE version a PC per projection surface is used.

As computers can hold more and more memory larger and larger REXs are feasible at no extra computational cost, but IBR will always entail some REX concept, i.e., some finite area within which the user can move freely, but outside which the scene cannot be rendered. IBR's biggest advantage is that no modeling whatsoever is involved. We just set up the acquisition system, scan the scene, and afterwards the images can be used directly for photo-realistic visual exploration of the scanned location.

Since IBR is based directly on recorded images the scenes that can be visualized with this technique have to be static. Moreover the visualization approach cannot handle if there are real world objects inside the acquisition region, and thus inside the REX.

In order to get dynamics, interactivity and/or objects inside REX we need to insert virtual objects (similarly to augmented reality). Augmented objects are visualized using a traditional model-based rendering paradigm, that is, virtual objects are modeled and textured in a commercial modeling package such as 3D Studio Max, saved in a VRML file and loaded into our system at start up time. To get scenario consistent illumination of the virtual objects we model the real scene illumination conditions [3] .

3. Experiences with the system

The system has been tested on hundreds of test persons experiencing one or more of about a dozen scanned real world locations ranging from wide open outdoor scenes over an indoor sub-tropical botanical garden to small office spaces. Generally our experience is that people are impressed with the IBR based approach and its ability to realistically recreate complex real world places in 3D stereo. Yet, our main impression is that people tend not to fully exploit the potential for visually exploring the displayed locations. People do look around in all directions but they do typically not perform much sideways head movement or shift their position perpendicularly to the viewing direction. If no perpendicular movement is performed it is impossible to appreciate that the displayed environment is in fact a full 3D environment with objects and structures at different depths, because then no motion parallax is generated and the only cue to depth differences is the stereo disparities.

The question then is: what space characteristics motivate visual exploration? We conjecture that spaces which generate strong motion parallax are interesting to explore. For the types of motion we are talking about this means scenes that have a lot of vertical structure at different depths, ranging from the center of the REX to far away.

4. Motivating exploration with augmentation

The main reason for inserting augmentations is to animate the user to visually explore the scenario and to engage in movements which generate visual parallax.

For the purpose of this study we have decided to operate with two types of augmentations, and the main hypothesis of this study is that these two types have fundamentally different effects on users' spatial perception of a scene. An example of each type is shown in Figure 1.

Outside-in augmentations. This type of augmentation basically occupies the center of the observer's area of move ability (REX). With these augmentations the observer is pushed out to the border of the REX and is stimulated to circulate the augmentation and thus to visually explore the scene relative to the augmentation. The main visual focus of the observer will tend to be on the augmentation.

Inside-out augmentations. This type of augmentation essentially surrounds the observer. The observer is animated to perform movements perpendicularly to the viewing direction in order to see past the augmentation, and the main visual focus may be evenly distributed between the augmentation and the rest of the scene.

Conclusions

We have proposed the use of visual augmentation for Virtual Environments based on Image-Based Rendering in order to stimulate observers to engage in more active visual exploration of the scenes. Specifically we have argued for categorizing augmentations into two types: outside-in and inside-out, which we believe will invoke different exploration behaviors.

Finally we believe that these two types of augmentation can be applied in similar manner to normal model-based Virtual Environments, which when displayed in either HMD or CAVE also restrict users' area of body movement due to the range of tracking equipment and, in the case of CAVE, the size of the CAVE. That is, all VEs where navigation is performed with normal body movements rather than with interaction devices can exploit the use of special objects placed in the scene at or around the center of the exploration area to shape the manner in which people explore the scene.

Figure 1: Top: solar clock augmented into an outdoor scene (outside-in augmentation). Bottom: pavilion inserted into a botanical garden scene (inside-out augmentation).

Acknowledgements

This research is funded by the BENOGO project under the European Commision (IST-2001-39184), and by the CoSPE project (26-04-0171) under the Danish Research Agency.

References

[1] V. Bruce, P. R. Green. Visual Perception: Physiology, Psychology and Ecology. 2nd edition. Lawrence Erlbaum Associates, publishers. 1990.

[2] E. Trucco, A. Verri. Introductory Techniques for 3-D Computer Vision. Prentice Hall. 1998.

[3] C.B. Madsen, M.K.D. Sørensen, M. Vittrup. Estimating Positions and Radiances of a Small Number of Light Sources for Real-Time Image-Based Lighting. *Proceedings of EUROGRAPHICS 2003,* 37 – 44, September 2003.