

# 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@cvmr.aau.dk, lwl@cs.aue.aau.dk}

## Abstract

*This paper explores the use of virtual objects augmented into visualizations of real scenes based on Image-Based Rendering. It has been observed that people tend not to fully exploit the possibilities for moving around and visually explore the scenes. We conjecture that by adding virtual objects we can heighten the amount of motion parallax generated from even small movements and thus motivate people to move actively explore scenes.*

**Keywords---** Visual augmentation, image-based rendering, spatial perception, motion parallax.

## 1. Introduction

In this paper we present 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 one can do 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 mobility in scenes. In our present system the user can only move freely inside a circle with a radius of about 2 feet. Such limited mobility 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 different 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.

The paper is organized as follows. First we present the technological background for this work describing the IBR approach and how visual augmentation can be done. Then we briefly describe an experiment performed using still images to explore people's characterization of various types of spaces. Then we present our two ways of augmenting scenes, and finally we describe an experiment we are currently setting up to evaluate the ideas behind this work.

## 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.

### 2.1. Image-based rendering

Image-Based Rendering (IBR) is as explained 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 images are acquired by moving the acquisition camera along the circumference of a circle, with the camera lens pointing outward of the circle. Typically we acquire around 700 images for a full 360 degree scan of the scene. Currently our setup allows us to acquire images in circle with a radius of 60 centimeters, although this is a mechanical constraint and not a principle IBR constraint.

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) or 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.

The size of the REX is limited by the amount of images we can hold computer memory for IBR. Currently 2 GByte is the memory limit in standard PCs, allowing us to run the system with approximately 400 images which in turn gives sufficient image resolution for a REX of radius 60 centimeters. 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.

## 2.2. Augmentation of IBR scenes

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] [4]

## 2.3. Experiences with 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 perpendicular 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.

A major reason for people not performing motion parallax generating visual exploration movements is naturally that their allowed range of movement is limited to the currently quite small REX (a circle of radius 60 centimeters). Nevertheless, even for such small REXs many scenes can cause noticeable parallax.

The work presented in this paper is thus motivated by a desire to explore the use of augmentation to stimulate people to fully exploit the available REX and to move in ways that enable them to experience the spatial layout of the displayed scene, i.e., ways that generate motion parallax.

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

In real world locations there will almost always be a lot of vertical structure in the distance, but since our REX is so small, and since no objects are allowed inside the circle of image acquisition our scenes do often not contain vertical structures very close to the center of the REX (and thus very close to the user). Therefore we need to use visual augmentation with virtual objects to add such close vertical structure.

## 3. Categorization of space

To gain insight into how consistently people characterize spaces according to various space defining parameters we performed an initial experiment using still images. We wanted to investigate if people consistently and objectively perceived spaces in similar ways.

Images of nine different locations were shown to approximately 45 people. For each image the test person was asked to characterize the depicted location according to 4 different axes: 1) constructed or organic, 2) full or empty, 3) big or small, 4) open or closed. Four of the nine location images are shown in Figure 1.

The test showed that perception of space on the chosen four axes is remarkably consistent across test persons. It is beyond the scope of this paper to present and discuss the full set of results from this experiment, but we believe the experiment clearly showed that it is meaningful to categorize spaces on the chosen axes. But more importantly we use the results from this experiment to hypothesize that deliberately chosen additions of objects or structure to a location can change the location's categorization on one or more axes. And in the context of our IBR-based Virtual Environments we can do just that: augment objects and structure into scenes.

That is, the fundamental hypothesis for the work presented in this paper is that in many locations a few key structures or elements actually tip the characterization of the location from one end of an axis to the other.

Moreover, we will use as hypothesis that subtle augmentations can be used to greatly enhance how intriguing users feel a location is. In this context subtle

augmentations are additions to the scene of objects and/or structures which do not completely dominate the scene but rather support the scene and create interesting visual phenomena. Looking at the location examples in Figure 1 we see objects and structures, which do not necessarily in themselves *define* the scene, but simply add depth and character (e.g., the window grid in the lobby scene and the curved pillars in the street corner scene). As such we will just use effects architects, sculptors, designers, and scenographers have known and exploited for millennia, namely that multiple depth layers and foreground vs. background can be used to create stimulating visual experiences that motivate visual exploration of the scene. It is our hope that motivating such exploration can enhance the users' sense of presence in the virtual representation of the scene.

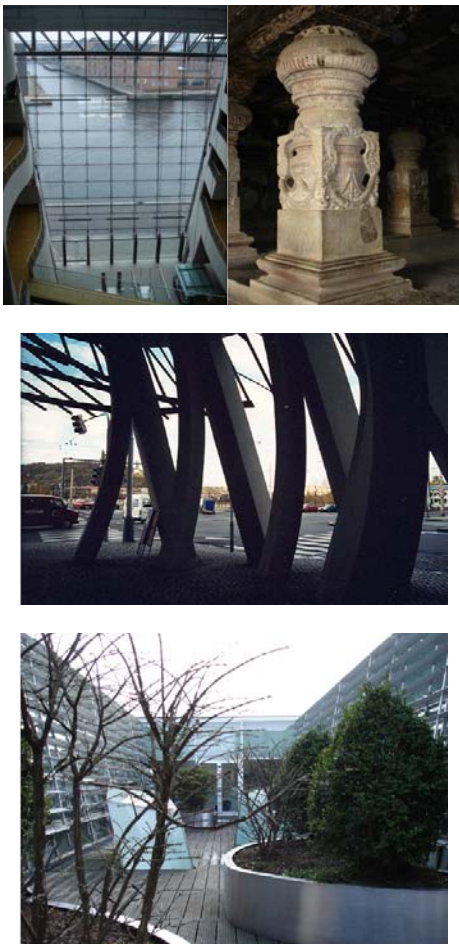


Figure 1: Four images from test of how people characterize space. All four examples were consistently characterized as constructed. Regarding the remaining three axes the top left location was by a clear majority of the test persons characterized as empty, big, and open. The top right as full, big, and closed. The middle location as empty, big and open. The bottom location was characterized as full, small, and 50/50 regarding open/closed.

## 4. Shaping spatial perception through augmentations

As described in the preceding sections we wish to shape users' spatial perception through augmentations and there are multiple reasons for doing it. In this section we first recapitulate those reasons and then we propose two different classes of augmentations.

### 4.1. Motivations for augmentations

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. The second reason for augmenting is that we believe the augmentations themselves and the parallax *they* generate relative to the background can help the observer to better perceive the 3D structure of the scene and perhaps better judging the absolute distances. The third and equally interesting reason is that it allows us a potential for transforming the overall perception of the location, e.g. tip the characteristics from an open to a closed location, i.e., we can essentially engage in space design based on real world locations.

For all augmentations we believe it important that they fit the context of the scene so as not to stand out and draw attention to themselves merely by being "out of place".

### 4.2. Types of augmentation

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. The two types are described below, and an example of each category is shown in Figure 2.

- 1) **Outside-in augmentations.** This type of augmentation basically occupies the center of the observer's area of moveability (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.
- 2) **Inside-out augmentations.** This type of augmentation essentially surrounds the observer. The observer is animated to perform movements perpendicular 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.

A positive side-effect of both types of augmentation is that they serve as a positional grounding for the observer, i.e. the observer can use the augmentation to orientate him or her-self relative to the REX and thus have an intuitive feel for the space available for visual exploration. This in itself may further stimulate head and body movements.



Figure 2: The top image shows an example of an outside-in augmentation, namely a solar clock augmented into an outdoor scene acquired at a viewpoint near a monastery. The bottom image shows an example of the inside-out type of augmentation, in this case a pavilion inserted into a botanical garden scene.

## 5. Design of experiments

We are presently setting up experiments with using the two types of augmentations in captured scans of real world scenes. Figure 2 showed two different scenes (an open and a closed scene) and examples of both the outside-in and the inside-out augmentations. We will experiment with all four combinations of scenes and augmentations (the two augmentations have been chosen to fit the context of both scenes).

There will be both quantitative and qualitative results from the experiments. The quantitative results will be based on recorded tracker data which our system allows to capture during a test. Based on these recordings we will study in what ways the two types of augmentations influence people's movement patterns.

The qualitative results will be based on questionnaires to be completed by test subjects. These questionnaires will address subjects such as what people have noticed in the scene, their impression of scene size, whether they feel being of normal size or if they feel smaller or bigger than normal. We will also ask people to characterize the scenes according to the characteristics described in section 3 to see if the augmentations influence people's overall perception of the displayed location.

## 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 user's 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.

## Acknowledgements

Two research projects have contributed to this research and these will be acknowledged in the final version of the paper.

## References

- [1] V. Bruce, P. R. Green. *Visual Perception: Physiology, Psychology and Ecology*. 2<sup>nd</sup> edition. Lawrence Erlbaum Associates, publishers. 1990.
- [2] E. Trucco, A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall. 1998.
- [3] Anonymous. Estimating Positions and Radiances of a Small Number of Light Sources for Real-Time Image-Based Lighting. In *Proceeding EUROGRAPHICS 2003 – Annual Conference of the European Association for Computer Graphic*, 37 – 44, September 2003.
- [4] Anonymous. The Importance of Shadows in Augmented Reality. In *Proceedings Presence 2003, 6<sup>th</sup> Annual International Workshop on Presence*, October 2003.