

# Depth Perception in View-Dependent Near-Field Spatial AR

Markus Broecker<sup>1</sup>Ross T. Smith<sup>1</sup>Bruce H. Thomas<sup>1</sup>

<sup>1</sup> Wearable Computer Lab  
 School of Information Technology & Mathematical Sciences  
 University of South Australia  
 Email: markus.broecker@unisa.edu.au  
 ross.smith@unisa.edu.au  
 bruce.thomas@unisa.edu.au

## Abstract

View-dependent rendering techniques are an important tool in Spatial Augmented Reality. These allow the addition of more detail and the depiction of purely virtual geometry inside the shape of physical props. This paper investigates the impact of different depth cues onto the depth perception of users.

*Keywords:* Augmented Reality, Spatial AR, Projector-based Rendering, Depth Perception

## 1 Introduction

Spatial Augmented Reality (SAR) provides a very direct form of mixed reality by augmenting physical models of objects using projected light (Raskar et al. 2001). This approach offers therefore more affordances than other types of augmented reality: the user is able to interact with physically existing models and there is a sense of ‘presence’ usually lacking in see-through displays (Bennett & Stevens 2005).

SAR operates in the following way: physical models, which act as projection surface, have a 3D geometric representation and each projector has a corresponding virtual camera. Projectors are aligned to geometry using pose-estimation techniques. After alignment, projectors augment the physical models by projecting the rendered view from their respective camera.

## 2 View-Dependent Rendering

View-dependent rendering provides perspective-correct rendering in a SAR environment from a single user’s perspective. It is used to provide additional detail to coarse physical models (Menk et al. 2011) or to create purely virtual geometry and space (see Figures 2 or 1). As there is no intrinsically known ‘central camera position’, the user’s position has to be tracked.

Creating these virtual geometries is a two-step process:

1. The user’s position is known from tracking. The view of the purely virtual geometry (the inside of the box in this example) is rendered from the user’s position and stored in a frame buffer.

Copyright ©2014, Australian Computer Society, Inc. This paper appeared at the Fifteenth Australasian User Interface Conference (AUIC 2014), Auckland, New Zealand. Conferences in Research and Practice in Information Technology (CRPIT), Vol. 150, Burkhard Wünsche and Stefan Marks, Ed. Reproduction for academic, not-for-profit purposes permitted provided this text is included.



Figure 1: The virtual inside of a box and virtual wall depth.

2. When the geometry of the box is rendered for each projector, projective texture mapping techniques, using the same parameters as before, are used to project the virtual content onto the geometry of the physical object.

This creates a view-dependent image for a single tracked viewer and independent of any projector’s position or projection parameters.

## 3 Depth Perception

Many individual depth cues are used by the visual system to provide an estimate of relative and absolute distances of objects within an image. Individual depth cues can be sorted by strength into three distinct ‘action distances’ (Cutting & Vishton 1995). In SAR, only near-field and mid-field actions are of interest.

Using view-dependent techniques in a SAR system creates two sets of conflicting depth cues: there are the ‘real’ depth cues, provided by the physical model of the box and the environment and there are the purely ‘virtual’ depth cues from the virtual content. It is desirable that the virtual depth cues’ strength is at least as strong as the real depth cues so that the virtual content is perceived as strongly as the physical prop.

## 4 User Studies

Two user studies were designed to measure depth perception of projected virtual spaces rendered using view-dependent techniques. To do so, an indirect measuring task was created. A box (as seen in Figure 1) had a virtual window cut into one side. The inside was purely virtual and depicted nine pyramids

at varying heights and positions within a  $3 \times 3$  grid. Participants were asked to select the apex of a highlighted pyramid by placing the tip of a tracked pen on the top surface of the box at the location where they suspected the projected position of the apex would lie. The selection was blind – no feedback of the selection was provided to the participant. Task distance error was measured and was used to indicate the effectiveness of a depth cue. The order of selection, the relative positioning of pyramids to each other and the order of the conditions tested for were randomised to minimise any learning effect. Participants were not trained to perform the selection task prior to the study.

#### 4.1 Difference between Real, Static and Head Tracked Perspectives

A first study investigated the difference of depth perception between a completely physical mockup of the virtual content, a static perspective (with head-tracking disabled) and the head-tracked perspective which provided depth parallax. Participants had no time constraint during selection. Three participants had previous experience with SAR installations, the others not. Thirteen valid data sets ( $N=13$ ) were collected with 468 data points for the virtual and 234 data points for the real condition. The participant's mean age was 28 years (min 17, max 64) and the gender distribution was three females and ten male participants. All participants can be classified as 'expert' computer users with more than 50 hours of computer use per week, but only three had previous exposure to spatial AR systems. The results are listed in Table 1.

Condition	Mean	Median	SD
Real	14.53	14.98	3.97
Head tracking	16.24	16.79	3.51
Static	17.75	17.87	4.47

Table 1: Aggregated distance error in the first user study in mm sorted by mean error.

We found no significant difference between the three conditions ( $F(2, 24.0) = 3.3, p = 0.054$ ).

#### 4.2 Head Tracking with Additional Depth Cues

The second study compared head tracking to different other additional depth cues. The following four conditions were tested for: head tracking by itself and head tracking with one of the following: texturing on virtual content (Figure 2), shadows and virtual wall depth ('tunnel effect', see Figure 1). These conditions were chosen after an initial pilot study. Participants were asked to perform the selection as fast and accurately as possible. Eleven data sets were collected ( $N=11$ ) with 396 data points per condition. Seven participants neither had experience with SAR installations nor participated in the first user study but all of the participants can be classified as 'expert' computer users. The participant's median age was 26 years (min 20, max 35); the gender distribution was one female participant and ten male. Table 2 shows the results of the second study.

One-way repeated measures ANOVA showed that there were no significant differences in accuracy between the different depth cues ( $F(3, 30) = 1.658, p = 0.197$ ). Further post-hoc analysis revealed no other

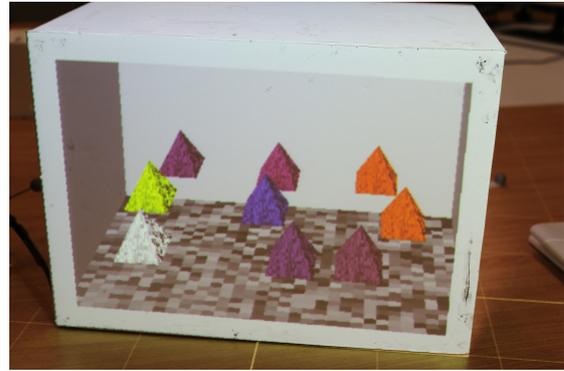


Figure 2: Textured pyramids.

Condition	Mean	Median	SD
HT + shadows	22.99	23.47	4.95
Head tracking	24.43	25.67	6.93
HT + texture	25.85	25.85	7.68
HT + wall depth	25.85	26.42	7.89

Table 2: Aggregated distance error values of the second user study in mm sorted by mean error.

significant differences between the error values, based on different depth cues. Therefore, different depth cues did not improve distance error over head tracking alone, which confirms the findings of the first user study. The increased error in all conditions, compared to the first user study can be explained by Fitt's Law, as there was no time-constraint for selections in the first study.

## 5 Discussion

This study was unfortunately unable to conclude that a certain depth cues significantly improves depth perception of virtual geometry. The performance of head tracking (parallax) improved selection error compared to the static perspective, however not significantly. Further work should investigate the impact of stereoscopic depth cues in near-field SAR.

## References

- Bennett, E. & Stevens, B. (2005), The Effect that Touching a Projection Augmented Model has on Object-Presence, in 'Information Visualisation, 2005. Proceedings. Ninth International Conference on', IEEE, pp. 790–795.
- Cutting, J. E. & Vishton, P. M. (1995), Perceiving Layout and Knowing Distances: The Integration, Relative Potency and Contextual Use of Different Information about Depth, in W. Epstein & S. Rogers, eds, 'Handbook of perception and Cognition.', Vol. 5: Perception of Space and Motion, pp. 69–117.
- Menk, C., Jundt, E. & Koch, R. (2011), 'Visualisation Techniques for Using Spatial Augmented Reality in the Design Process of a Car', *Comput. Graph. Forum* **30**(8), 2354–2366.
- Raskar, R., Welsh, G., Lim Low, K. & Bandyopadhyay, D. (2001), Shader Lamps, Technical report, MITSUBISHI ELECTRIC RESEARCH LABORATORIES, Eurographics Rendering Workshop.