Apparent Depth in a 3 Dimensional Model Using Techniques in Stereoscopic Rendering and Depth of Field Effect

Mason Stilwell, Jack Hamby, and Conrad Parker, St. Olaf College

This paper examines the use of blur in simulating depth of field in a stereoscopic setting using OpenGL, in order to create a more interactive and navigable 3D model. We introduce important terms and ideas relevant to stereoscopy and depth of field. We also discuss several recent papers that have studied the use of depth of field effects to create more comfortable and natural images, as well as effective blurring rendering methods. We then explain the methods used in order to obtain our data in terms of both our work with stereoscopy and blur effects. We describe the hardware we're using to project in 3D as well as how that influenced the creation of our software. We examine in detail our use of OpenGL's glAccum() function to generate focal blur through a modified accumulation-buffer technique. We took user perception data from a participant study in which the subjects are asked to match real stereographic photographs with stereographic renderings of the same scene based on focal distance. Our results indicate that our simulation of depth of field was mostly ineffective. We postulate that the abstract nature of our rendered scene confused our test subjects and hindered our study, and we recommend that future research apply focal blurring techniques to more realistic virtual scenes.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation/*
methodology; H.1.2 [**Models and Principles**]: User/Machine Systems—*Human Information Processing*; I.5.1 [**Pattern**
Recognition]: Models—*Computer Graphics*

General Terms: Computer Graphics

Additional Key Words and Phrases: Stereoscopic, OpenGL, Depth of Field, Rendering, glAccum

1. INTRODUCTION

There has been a huge push in recent years to create more immersive multimedia applications that seek to replicate some physical place and give the user presence within this environment. Applications that utilize simulated reality are still being realized as this technology continues to emerge. Companies like Matterpoint and Floored are using these ideas in their creation of 3 dimensional models from real spaces. In order to make these models more immersive, many of these models are being made to work with VR devices. However, eye fatigue is an often cited issue that arises from complete immersion [Leroy et al. 2012]. Discomfort seemed to occur in subjects when they were exposed to high spacial frequencies with very high levels of disparity. One potential method to reduce this strain is to alter the depth of field (DoF), which is essential for reducing eye strain [Perrin and Perret 1998]. Our work, which also attempts reduce spacial ambiguity, focuses more on improving user *perception* of 3D space rather than user comfort. In this paper, we examine techniques in stereoscopy and simulated depth of field as applied to a 3D floor space model and the effect on the user's perception of the 3D space.

1.1 Stereoscopy

The primary method of stereoscopy examined in this paper is a two-projector passive system.While this sytem requires the creation of two seperate windows, this method can be easily manipulated to create a single window with the two eye projections on each side for use with a VR system. In talking about these methods in more detail, it becomes necessary to define some terms:

Interocular Distance (IOD). The distance between the centers of rotation of the eyeballs of an individual or between the oculars of optical instruments.

Binocular Disparity. Difference in image location of an object as seen from the left and right eyes

Accommodation. The change in optical power of the eye via focal length in order to maintain focus on an object at varying depths.

Frustrum. Region of space in the modeled world that may appear on the screen.

Parallax. Displacement in the apparent position of an object as viewed along two different lines of sight.

Our approach to stereographics requires the creation of stereo image pairs, which simulates typical binocular disparity in humans that act as a depth cue. Unfortunately we must also account for accommodation. Since the individual pairs are still 2 dimensional, the accommodation cues for the eyes will be inconsistent with the perceived 3 dimensional image cue created from the binocular disparity [Held et al. 2012]. Typically, our visual system can accommodate this discrepancy to a maximum eye separation of 1/30 the focal length.

1.2 Depth of Field

Depth of Field refers to the expanse of an image that appears in acceptable focus. In a real optics system, the depth of field is controlled by three quantities:

(1) The distance at which the lens is focused.
(2) The relative aperture (or f-stop).
(3) The focal length of the lens.

Professional photographers and cinematographers can control these adjustments to direct the view of the audience, often by pulling the focus to only one area of the screen. Computer graphics don't have these controls, which leads to everything being in sharp focus. This feels unnatural to the user [Grosset et al. 2013]. Adding these adjustments in our computer graphics should render the scene more naturally. Additionally, we believe it can be used effectively to draw the attention of the user towards some item in the scene and help eliminate text cues. In doing so, the user experience will become more intuitive and compelling.

Optical depth of field comes with the limitation that focus is determined by the three qualities stated above. This means that focus is determined with distance of the subject from the camera, where that plane of focus can move laterally from the camera. A computer simulated DOF can overcome this, since we can apply blur based on user perimeters independent of depth. For instance, we could use this to focus on a single subject out of a set of three that all exist in the same plane, while maintaining focus on an object on a completely different plane [Kosloff and Barsky 2007]. For our purposes, we focus on how to use blur to create a DOF effect. There exist several methods for blurring that could add to users perception of space. The first method, conveniently dubbed the ”depth of field” technique, simulates the blurring of objects in front of, or behind the eye's focal point. The second method is ”peripheral blur” which blurs all objects outside of the main area of focus (objects within the peripheral vision)[Hillaire et al. 2007]. In our research we focus on the former.

In simulating depth of field, we use an accumulation-buffer technique to blend layers of the rendered images from multiple pinhole cameras that contain no depth information. We

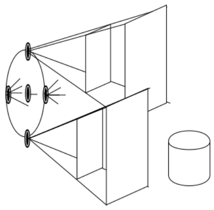

Fig. 1: Our use of glAccum().

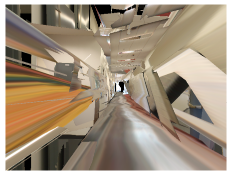

Fig. 2: Our rendering of the Regents Hall of Natural Science.

modify this slightly to include aspects of the composite technique, the layers are composited from front to back, and then blended using alpha blending. [Porter and Duff 1984].

2. METHODS

All rendering was done in OpenGL, using GLFW and GLUT library packages. The rendered 3D space was made from images taken from Regents Hall of Natural Science. Our images come from an ongoing project of the St. Olaf College Computer Science department to create a three-dimensional model of Regents Hall from pairs of photographs. For the past several years, students have been manually identifying roughly planar objects in these pairs and locating them in 3D space.[1] We wrote a program to load this information and create a 3D model of the building as it exists in the project's files.

2.1 Accumulation-Buffer Technique for Simulated DoF

To simulate DoF in OpenGL, we utilize glAccum() to render our scene several times, each from a slightly different camera location. We combine the rendered images into the accumulation buffer to compute an averaged composite of the accumulated objects. We use an integer we call *jitter* to specify the number of copies that are accumulated into the buffer. Each image is rendered from a camera position in a circle around the actual viewpoint and parallel to the projection plane. For instance, if jitter is 24, we draw our scene 24 times, accumulating the data for these 24 images in the accumulation buffer. Once the accumulation is finished we can call glAccum() again with *operation* set to replace to current draw buffer with our accumulation buffer [Woo et al. 1999].

2.2 Asymmetric Frustum Projection for Passive Stereoscopic 3D

To render our scene, we must set up a virtual camera and render two stereo pairs at some offset. The cameras should look along parallel vectors, which are perpendicular to the projection plane. However, we also want to make sure the frustum from the camera to the corners of the projection plane remains symmetric. If we were to simply angle each frustum towards the projection plane, the resulting asymmetry in the projections would introduce vertical parallax into our image. We want to maintain parallelism of the camera view vector so that the viewing frustum of each camera is extended horizontally past the projection plane. This shift should be equal to $\pm\frac{1}{2}(\mathrm{IOD})$. Then, we trim this region from the rendering. For our two-projector system, we created two separate windows. This method can be easily manipulated to create a single window with the two eye projections on each side for use with a VR system.

2.3 Participant Use Study

To evaluate the effectiveness of our method, we rendered a hall in the Regents Hall of Natural Science in our program with several different focal distances. The image pairs

[1] A recent team has explored the use of machine learning to identify objects automatically.

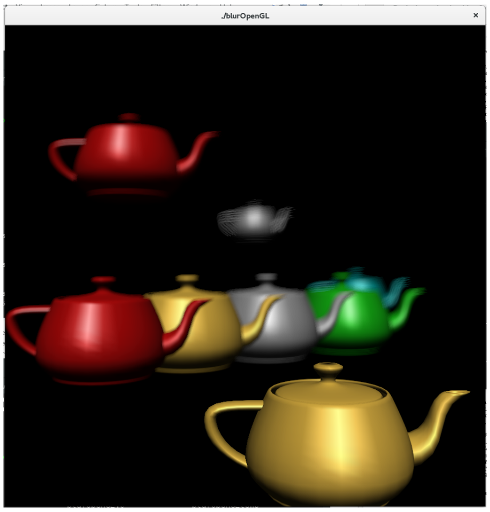

Fig. 3: Rendered teapots in OpenGL demonstrating the use of our blur function

Fig. 4: The left and right sides of our stereoscopic focused image pair.

so generated resembled that in Fig. 4. We also took stereoscopic photographs of the same location in real life at various focal distances. We projected the images stereographically, one pair of each type at a time, and asked participants to match the photos to the rendered images based on focal distance as we cycled through. We also asked them to rate their visual comfort on a scale from -5 to 5.

3. RESULTS

The 3D model produced from these images is still far from a replicate of the real life scene, as the data from Eriol is imperfect and incomplete. The results can be seen in Fig. 5. To generate the numbers on the x-axis, we assigned values of 1, 2, 3, and 4 to the different focal distances we took the photographs at so that higher numbers corresponded to longer distances. The data has a correlation of -0.35, indicating that participants actually matched images in the opposite order on average.

4. DISCUSSION

Initial results did not support our hypothesis that dynamic DoF manipulation in conjunction with stereoscopic 3D would improve the viewers perception of 3D space. Matching of

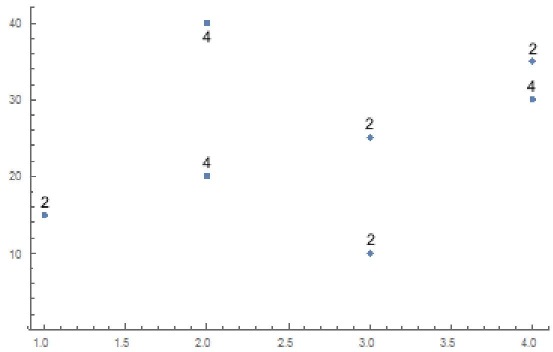

Fig. 5: The results of the image matching study. The x-axis is a crude measure of focal distance on the photos, and the y-axis is the simulated focal distance of the rendered image in feet. The labels indicate number of repeats.

the rendered images with varying blur distances to the taken image with at varying focus distances was indiscriminate and statistically insignificant ($p < 0.05$). This indicates test subjects could not discern the effect of blurring in the rendered image.

Depth of field as a means of enhancing depth has in no way been conclusive in previous studies. Results have ranged from blur being at best a weak depth cue [Mather 1996], to a conditional but sometimes useful cue [Grosset et al. 2013], to very important and underestimated cue [Held et al. 2012]. Similar improvement in users perception of depth has been shown in stereoscopic imaging [Howard and Rogers 1995][Lipton 1997][McAllister et al. 1995], but less than optimal conditions can actually hinder the usefulness of stereoscopy [Pollock et al. 2012]. While yet mostly untested in computer rendering, the combination of these two cues has further been observed to complement one another in human perception of depth [Held et al. 2012]. Our data does not directly contradict the previous body of research, since previous results have been inconclusive. We believe our inconclusive results are due to the incomplete rendering seen from our 3D scene rendering method. The textures observed in these rendered images already have a significant amount of distortion from objects that have incomplete mappings. Given that the rendered scene prior to the application of our methods appears abstract, the use of this approach will not make the scene appear concrete.

Our initial images were promising, and we believe that more research is needed to fully answer our question. To reach a more definitive conclusion, the same methods of study should be applied to more realistic rendering of our scene, and a larger sample size of participants. Participant observation as a means of determining the effectiveness of the approach on aiding interactivity and navigability of the 3D model would also provide a useful metric of data [Kawulich 2005].

REFERENCES

A.V.P. Grosset, M. Schott, G.-P. Bonneau, and C.D. Hansen. 2013. Evaluation of Depth of Field for depth perception in DVR. In *Visualization Symposium (PacificVis), 2013 IEEE Pacific*. 81–88. DOI:http://dx.doi.org/10.1109/PacificVis.2013.6596131

RobertT. Held, EmilyA. Cooper, and MartinS. Banks. 2012. Blur and Disparity Are Complementary Cues to Depth. *Current Biology* 22, 5 (2012), 426 – 431. DOI:http://dx.doi.org/10.1016/j.cub.2012.01.033

Sébastien Hillaire, Anatole Lécuyer, Rémi Cozot, and Géry Casiez. 2007. Depth-of-field blur effects for first-person navigation in virtual environments. In *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*. ACM, 203–206.

Ian P Howard and Brian J Rogers. 1995. *Binocular vision and stereopsis*. Oxford University Press.

Barbara B Kawulich. 2005. Participant Observation as a Data Collection Method. (2005).

Todd J Kosloff and Brian A Barsky. 2007. An algorithm for rendering generalized depth of field effects based on simulated heat diffusion. In *Computational Science and Its Applications–ICCSA 2007*. Springer, 1124–1140.

Laure Leroy, Philippe Fuchs, and Guillaume Moreau. 2012. Real-time adaptive blur for reducing eye strain in stereoscopic displays. *ACM Transactions on Applied Perception (TAP)* 9, 2 (2012), 9.

Lenny Lipton. 1997. Stereographics, Developers Handbook. *StereoGraphics Corporation* (1997), 1–26.

George Mather. 1996. Image blur as a pictorial depth cue. *Proceedings of the Royal Society of London B: Biological Sciences* 263, 1367 (1996), 169–172.

David F McAllister, Susan McKay, and others. 1995. Stereo computer graphics and other true 3D technologies. *Computers in Physics* 9, 1 (1995), 56–56.

Fuchs P. Roumes C. Perrin, J. and F. Perret. 1998. Improvement of stereoscopic comfort through control of disparity and spatial frequency content. *Proceedings of the SPIE Visual Information Processing Conference* (1998).

B. Pollock, M. Burton, J.W. Kelly, S. Gilbert, and E. Winer. 2012. The Right View from the Wrong Location: Depth Perception in Stereoscopic Multi-User Virtual Environments. *Visualization and Computer Graphics, IEEE Transactions on* 18, 4 (April 2012), 581–588. DOI:http://dx.doi.org/10.1109/TVCG.2012.58

Thomas Porter and Tom Duff. 1984. Compositing digital images. In *ACM Siggraph Computer Graphics*, Vol. 18. ACM, 253–259.

Mason Woo, Jackie Neider, Tom Davis, and Dave Shreiner. 1999. *OpenGL Programming Guide: The Official Guide to Learning OpenGL, Version 1.2* (3rd ed.). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

Received February 2009; revised July 2009; accepted October 2009