Disocclusion Mitigation for Point Cloud Imposters

Chad Mourning[*] David Chelberg[†] Ronaldo Vigo[‡] Derek Ziegler[§]
Ohio University

Figure 1: *The four rendering methods tested, from left to right: full fidelity, decimated, imposters, and imposters with disocclusion mitigation.*

Abstract

Image based imposters suffer from common errors called disocclusion artifacts where portions of the scene that should be occluded by real geometry are visible when using image based imposters. These artifacts are the result of parallax error created by camera motion where regions of a mesh that were not visible at the time of imposter generation have become visible. This document presents an analysis of a computationally inexpensive on-line technique [Mourning et al. 2014] to resolve these disocclusions by stretching existing imposter [Maciel and Shirley 1995] texture information over new geometry bridging the gap between imposters. [Mourning et al. 2014] only presented automatic metrics showing improved image quality compared to traditional techniques; in order to corroborate the findings in [Mourning et al. 2014], human trials were performed to determine if human subjects found a similar increase in image quality. Results show a statistically significant improvement in image quality over traditional imposters.

Keywords: level-of-detail, point clouds, image-based impostors

Concepts: •**Computing methodologies** → **Image-based rendering;**

1 Introduction

Disocclusion artifacts are a common source of error in scenes using image-based imposters as a level-of-detail solution to reduce the computational complexity of the scene. The work presented in [Mourning et al. 2014] aimed to mitigate these errors by resolving the disoccluded space and using the texture information of the surrounding imposters. In the original work, [Mourning et al. 2014] presented an analysis based on six automatic metrics; additional study has been conducted with 105 human subjects performing 3 experiments to evaluate the overall image quality of the new approach compared to other common approaches.

*e-mail:cm492997@ohio.edu
†e-mail:chelberg@ohio.edu
‡e-mail:vigo@ohio.edu
§e-mail:dz118006@ohio.edu

I3D '16, February 27-28, 2016, Redmond, WA
ISBN: 978-1-4503-4043-4/16/03
DOI: http://dx.doi.org/10.1145/2856400.2876016

2 Methods

A lidar point cloud of size 12,212,086 was split into a 12x5 grid of independent meshes. This scene was rendered 4 times using 4 different rendering techniques. First, to development ground truth, the scene was rendered at full fidelity. Second, the 60 independent meshes were replaced with image based imposters [Maciel and Shirley 1995]. Third, imposters were used once again, but augmented with the disocclusion mitigation strategy presented in [Mourning et al. 2014]. Finally, the scene was rendered decimated at a $1:30$ ratio, to match frame rate with the imposter based techniques, serving as a comparison level-of-detail approach.

A quick summary of the disocclusion mitigation strategy presented in [Mourning et al. 2014]:

- Calculate the world space positions of screen space twin points of interface between meshes when projected into the imposter plane.
- Create bridging geometry to resolve the possible dissoclusion space.
- Calculate texture contribution weights of adjacent meshes.
- Assign new texture coordinates to imposter and disocclusion mitigation geometry.

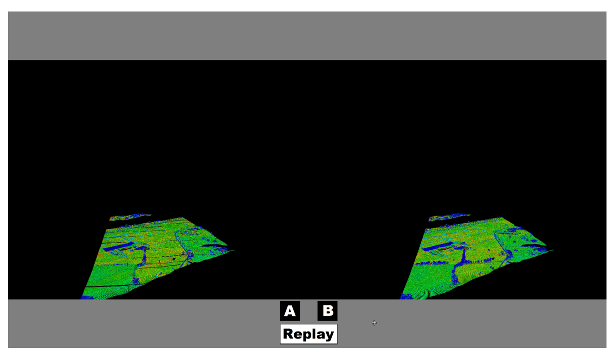

Figure 2: *An example screenshot of a subject making a judgement during experiment 2.*

A 900 frame flightpath was flown over the dataset to capture disocclusions resulting from motion-induced parallax. This flightpath was flown for each of the 4 rendering techniques at each of 3 different resolutions. The frames of each rendering were concatenated into three 5 second video clips which were shown to human subjects in three distinct experiments. The first experiment was an evaluation of image quality from 1 to 9. The second was a forced choice judgement in quality between two simultaneous clips; all randomized pairings of clips were tested. The third was a judgement between two simultaneous clips with ties allowed. All scenes were rendered in the STEAMiE engine [Nykl et al. 2008]. All experiments were developed using the Psychophysics Toolbox Version 3 [Brainard 1997]. For reference, the final scene was reduced from 12,212,086 vertices to roughly 1280 vertices.

In experiment 1, video segments could be replayed up to 1 additional time, and the user was required to make a selection from 0-9 in a fixed amount of time or a *no result* would be recorded. During the experiment each subject would view each of the 36 video clips multiple times.

During experiment 2 every ordered pair of video clips with the same resolution were played and the subject was forced to make a judgement on which video had higher image quality. This resulted in 144 possible pairings (pairs of matching clips were included). In experiments 2 video pairings could be replayed up to 2 additional times, and the user was required to make a make a selection of A or B in a fixed amount of time of a *no result* would be recorded. Experiment 3 followed the same rules as experiment 2, except users had the option to select a tie in addition to the A and B options.

Figure 2 shows an example screenshot of what a human subject would see during experiment 2 after the end of a video clip. On the left is the final frame of the scene rendered with imposters without disocclusion mitigation and on the right is the scene rendered with full fidelity. Beneath the renderings are buttons to select the left scene, the right scene, or replay the videos, along with the selection cross-hairs.

3 Results

Figure 3 shows the means (numeric) and standard deviations (error bars) for each of the four rendering methods, with values from 1-9 matching the selectable options during experiment 1. The *p*-values for pairwise t-tests between the rendering methods in order of increasing quality are also provided, with each of the pairs separated by a statistically significant margin.

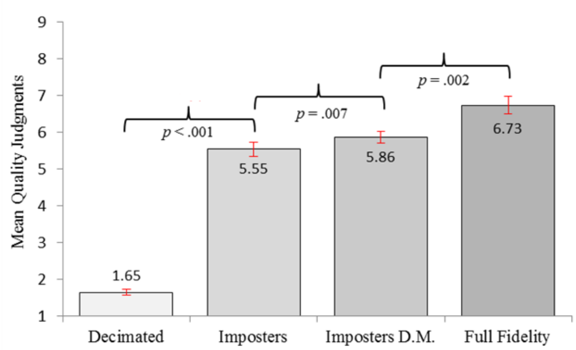

Figure 3: *The means, standard deviations, and pairwise t-test* p-*values of each group.*

Table 1 provides a results matrix for each comparison combination for all subjects when subjects were forced to select either video A or video B. The $p - values$ were calculated using a χ^2 analysis.

Table 2 provides a results matrix for each comparison combination for all subjects when ties were allowed. The $p - values$ were calculated using a χ^2 analysis.

Replays and no-responses were both infrequent events occurring in less than 1% of instances.

Fmt 1	Fmt 2	Tot 1	Tot 2	NR	p-value
Full Fid.	Dec.	696	5	1	$\leq .001$
Full Fid.	Imp.	504	197	1	$\leq .001$
Full Fid.	w/DM	471	228	3	$\leq .001$
Dec.	Imp.	13	688	1	$\leq .001$
Dec.	w/DM	13	689	0	$\leq .001$
Imp.	w/DM	187	510	5	$\leq .001$

Table 1: *Results for the forced choice simultaneous playback image quality test. All pairings were separated statistically significantly.*

Fmt 1	Fmt 2	Tot 1	Tot 2	Ties	NR	p-value
Full Fid.	Dec.	728	5	4	1	$\leq .001$
Full Fid.	Imp.	366	306	40	5	.021
Full Fid.	w/DM	319	355	40	5	.19
Dec.	w/DM	13	714	11	0	$\leq .001$
Dec.	w/DM	4	725	7	2	$\leq .001$
Imp.	w/DM	84	354	294	6	$\leq .001$

Table 2: *Results for the choice with ties simultaneous playback image quality test.*

4 Conclusions

The human subject trials support the results of the automatic metrics evaluated in [Mourning et al. 2014]. All experiments showed a statistically significant improvement when incorporating disocclusion mitigation into rendering. The gap between imposters with and without disocclusion mitigation grew with resolution of output images, implying increased usefulness in the future. When ties were allowed, no statistically significant difference was found between the full fidelity rendering and imposters with disocclusion mitigation, however this is likely triggered by cognitive laziness rather than extreme similarity in the scenes.

5 Future Work

Future work would involve higher level experiments, such as finding objects in scenes with and without disocclusion mitigation, experiments with expert subjects, experiments testing quality of the mitigation strategy under 6 degrees of freedom, and evaluation of how the subjects respond when the technique is used on stereoscopic displays.

References

BRAINARD, D. H. 1997. The psychophysics toolbox. *Spatial vision 10*, 433–436.

MACIEL, P. W. C., AND SHIRLEY, P. 1995. Visual navigation of large environments using textured clusters. In *Proceedings of the 1995 symposium on Interactive 3D graphics*, ACM, New York, NY, USA, I3D '95, 95–ff.

MOURNING, C., NYKL, S., AND CHELBERG, D. 2014. Disocclusion mitigation for image based point cloud imposters. *Advances in Visual Computing*, 861–871.

NYKL, S., MOURNING, C., LEITCH, M., CHELBERG, D., FRANKLIN, T., AND LIU, C. 2008. An overview of the steamie educational game engine. In *Frontiers in Education Conference, 2008. FIE 2008. 38th Annual*, IEEE, F3B–21.