

Enhancing Perception of Complex Sculptural Forms using Interactive Real-time Ray tracing

Andy Lomas
Goldsmiths, University of London
London SE14 6NW, UK
a.lomas@gold.ac.uk

This paper looks at experiments into using real-time ray tracing to significantly enhance shape perception of complex three-dimensional digitally created structures. The author is a computational artist whose artistic practice explores the creation of intricate organic three-dimensional forms using simulation of morphogenesis. The generated forms are often extremely detailed, comprising tens of millions of cellular primitives. This often makes depth perception of the resulting structures difficult. His practice has explored various techniques to create presentable artefacts from the data, including high resolution prints, animated videos, stereoscopic installations, 3D printing and virtual reality.

The author uses ray tracing techniques to turn the 3D data created from his morphogenetic simulations into visible artefacts. This is typically a time-consuming process, taking from seconds to minutes to create a single frame. The latest generation of graphics processing units offer dedicated hardware to accelerate ray tracing calculations. This potentially allows the generation of ray traced images, including self-shadowed complex structures and multiple levels of transparency, from new viewpoints at frame rates capable of real-time interaction. The author presents the results of his experiments using this technology with the aim of providing significantly enhanced perception of his generated three-dimensional structures by allowing user-initiated interaction to generate novel views, and utilizing depth cues such as stereopsis, depth from motion and defocus blurring.

The intention is for these techniques to be usable to present new exhibitable works in a gallery context.

Generative art. Computational art. Depth perception. Real-time interaction. Ray tracing. Graphics processing units.

1. INTRODUCTION

The author is a practicing computational artist, whose work explores how complex organic forms can be created through digital simulation of morphogenetic processes. Inspired by Alan Turing's use of simple equations to create rich self-organizing patterns (Turing 1952), the author's work focuses on creating simplified models of growth at the level of individual cells and exploring the emergent forms that can be created from these low-level rules (Lomas 2014a).

A key focus of the author's work is to push computational limits for how much intricacy can be created from simulation processes. Typically he runs initial simulations until they produce 1,000,000 cells, which he considers a seedling level, to get an indication of what combinations of parameter values may yield interesting results. These are then iteratively refined using a combination of interactive genetic algorithms (Todd & Latham 1992, Bentley & Corne 2002) and machine learning methods (Lomas

2016, McCormack & Lomas 2020). The resulting parameter values are then re-run on the simulation system with the intention of generating more intricate structures, typically with 10,000,000 to 100,000,000 cells (Figure 1).

The data from these simulations is used to create artefacts using a variety of means, including 2D prints of rendered images, high definition video files, stereoscopic installations and 3D printed sculptures (Lomas 2019a, Lomas 2019b) (Figures 2, 3 & 4). In particular, the author is interested in how realizing the data into different types of artefact changes the user's perception of the data. The author often creates a number of artefacts from the same original data set, each of which visualize different elements from the simulation. He is interested in exploring how perception is affected depending on whether an artefact is in 2D or 3D (such as using stereopsis), whether it is still or animating, and whether it is a physical 3D artefact (such as using 3D printing) or

virtual 3D one (such as using stereopsis or virtual reality).

The data from the author's simulations are typically extremely complex three-dimensional organic shapes which lack simple perspective cues such as parallel lines. This means that depth perception can be particularly challenging.

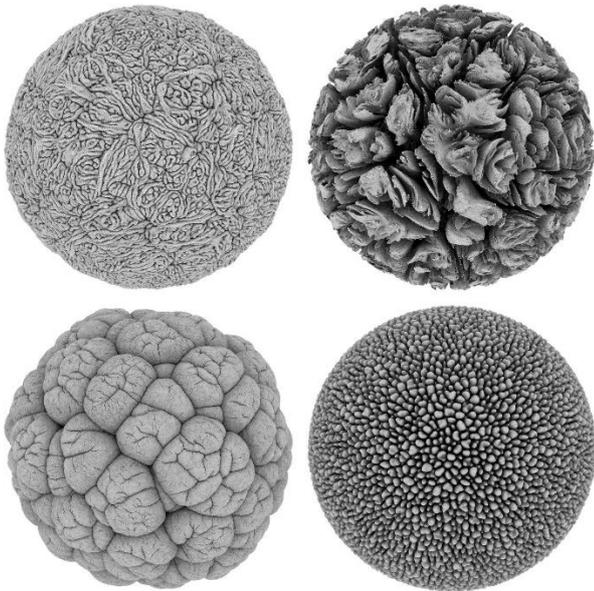


Figure 1: Examples of Cellular Forms



Figure 2: Morphogenetic Creations exhibition, Watermans, 2016

A significant amount of research has been conducted into perception depth (Howard 2012), identifying many different cues that humans use. Stereopsis, the ability of our brains to compare different images from our left and right eyes to create perception of depth, is probably the best known of these. However, it can be argued that in many cases other cues are more important to create depth perception. These cues are typically divided into monocular and binocular cues (Wikipedia Contributors 2019).



Figure 3: Vase Forms, London Design Festival exhibit at the V&A, 2018

Monocular cues:

- Motion parallax.
- Depth from motion.
- Kinetic depth effect.
- Perspective.
- Relative size.
- Familiar size.
- Absolute size.
- Aerial perspective.
- Accommodation.
- Occultation.
- Curvilinear perspective.
- Texture gradient.
- Lighting and shading.
- Defocus blur.
- Elevation.

Binocular cues:

- Stereopsis.
- Convergence.
- Shadow stereopsis.

In previous work the author has explored the use of a number of these cues in artefacts he has created for exhibition. These include:

Lighting and shading:

- 2D images, such as the author's Cellular Forms (Lomas 2014a) (Figure 1) and Plantlike Forms (Lomas 2014b), rendered using ambient occlusion.

Depth from motion:

- Rotation applied to 3D objects in animated videos. This was used for the author's Hybrid Forms (Lomas 2015) and Mutant Vase Forms (Lomas 2017).

Stereopsis:

- Animation and still images with separate rendering for the left and right eyes. These have been presented using a variety of techniques including installations using mirrors to act as a Wheatstone Viewer, anaglyph rendering using red and cyan, and installing digital parts into antique Victorian Brewster stereoscopes. The author has used stereopsis in a number of installations of his Hybrid Forms as well as for his Mutant Vase Forms (Figure 4).

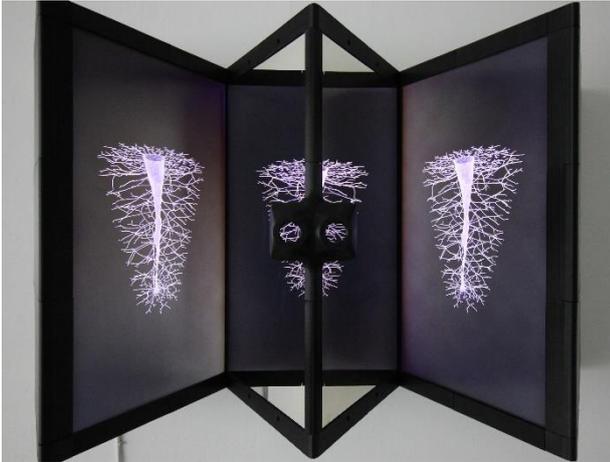


Figure 4: Mutant Vase Forms: stereoscopic installation

Defocus blur:

- For the 2D prints the author created for his Cellular Forms and Plantlike Forms he used depth information generated during rendering to apply a small amount of defocus blur to the high-resolution images to enhance depth perception. This was done as post-process using Foundry's Nuke software.

The author has also performed several experiments into how data from simulations could be viewed in high-end virtual reality systems such as the HTC Vive. These systems can provide depth perception using both stereopsis and motion parallax but can be limited by the rendering techniques suitable for real-time rendering, such as using OpenGL.

2. INTERACTIVE REAL-TIME RAY TRACING

In order to render the data from the cellular simulations, as well as to provide a method for simulating light rays hitting cells during the growth simulations, the author implemented a simple backwards ray tracer (Glassner 1989) that runs on the GPU using CUDA. This has been the main rendering technique that the author has used for a number of years, including the creation of his Cellular Forms (Lomas 2014a), Hybrid Forms

(Lomas 2015) and Mutant Vase Forms (Lomas 2017).

Typical render times using the author's own GPU ray tracing software vary from a few seconds for high definition video (1920x1080) frames to around a minute for very high resolution (8192x8192) images for prints. While these render times are suitable for generating pre-rendered videos or still images for prints, they are significantly longer than would be acceptable for interactive use where render times of 1/30th of a second or less are typically necessary.

Variations of ray tracing, such as bidirectional path tracing, are the current standard for high-quality computer-generated imagery for visual effects and animation. For interactive use, such as for computer games, rasterization methods that make use of the GPU (such as using OpenGL) are more generally used. However, the latest generation of GPUs, such as the NVIDIA RTX series, offer hardware support for accelerated ray tracing. These can be programmed using a number of APIs including DirectX, Vulkan and NVIDIA's OptiX library (Parker et al. 2010).

Dedicated hardware for ray tracing has the potential for significantly speeding up render times, including the possibility of generating high definition renders sufficiently fast to enable interactive use. The NVIDIA GeForce RTX 2080 Ti has a quoted performance of 10 billion ray traces per second (NVIDIA 2020), approximately 1,000 times the number of ray-traces per second that the author has been able to achieve with his own custom GPU ray tracing software. This has the potential of allowing the author to progress from presenting artwork using pre-rendered animations to allowing the user to directly manipulate forms and view them from novel viewpoints as they are rendered in real-time.

As discussed in the previous section, there are a wide range of different cues that humans use for depth perception. A potentially interesting area for research is whether user-initiated interaction, such as the user actively controlling the orientation of a 3D object or the focal distance of the camera, can significantly enhance depth perception when used to affect cues such as depth from motion and defocus blur.

3. METHOD

The author created a custom C++ rendering application that reads in binary particle data files generated by his cellular growth simulations. The application renders this data using NVIDIA's OptiX library to perform GPU based ray tracing calculations. The hardware used for these tests was

a PC running Ubuntu 18.04 with an NVIDIA GeForce RTX 2080 Ti graphics processor.

Two rendering methods were implemented, matching the rendering techniques that the author had previously used for images of his Cellular Forms (Figure 5):

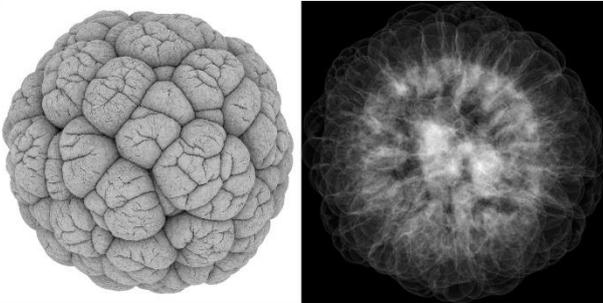


Figure 5: Ambient occlusion and X-Ray renders from the same Cellular Form data

- (i) Ambient occlusion. This is a technique where objects are rendered as perfectly diffuse (Lambertian) opaque white surfaces illuminated by an omni-directional uniform light source (Miller 1994). The 3D form is revealed through self-shadowing of the surface. This type of rendering visualizes the structure of the external surfaces.
- (ii) X-Ray rendering, where rays from the camera accumulate density from each cell that they hit as the ray travels along a straight path through the structure. This creates an X-Ray like image, typically with multiple layers of transparency. This type of rendering allows visualization of internal structures.

For rendering, each cell was represented as a sphere. This was implemented using a custom program written using NVIDIA's OptiX API to calculate the intersection of rays with sphere primitives. Individual bounds were computed for each sphere to create the acceleration structure used for rendering.

Defocus blur was implemented as an option in the renders using a simple stochastic model: jittering the eye position within a circle and tracing the ray towards a fixed target on the focal plane for each pixel (Cook et al. 1984).

Rendering tests were conducted using one eye-ray per pixel for both ambient occlusion and X-Ray renders. The ambient occlusion renders used 16 shadow rays for each intersection point. If the user doesn't change the rotation or the focus settings additional frames are rendered with the origin of the eye ray jittered within each pixel. These frames are combined to give continuous progressive refinement of the displayed image.

The author tested the application with two different configurations to display the output:

- (i) A single high definition (1920x1080) computer monitor.
- (ii) A mirror based stereoscopic viewer. This comprises two 4k (3840x2160) resolution computer screens that are viewed through front surface mirrors arranged as a Wheatstone stereoscope (Wikipedia Contributors 2020) (Figure 6).

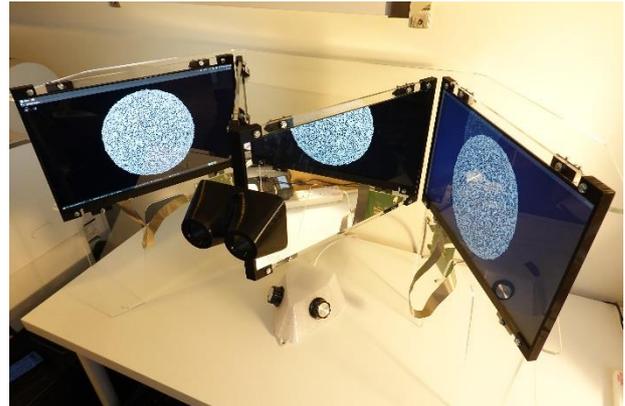


Figure 6: Mirror based Wheatstone stereoscope viewer

User interaction was enabled using two rotary potentiometers connected to an Arduino, with data sent to the host computer over a USB serial connection.

The following user-initiated actions were implemented to allow users to interact with the form they are viewing:

Depth from motion:

- Allowing the user to initiate rotational motion of the viewed object, modifying the orientation of the form in 3D space.

Defocus blur:

- Allowing the user to change the distance to the focal point when rendering using defocus blur.

The two potentiometers were tested in two different configurations: one where both were used for rotation around different axes, and a second where one potentiometer controlled rotation and the other allowed the user to interactively change the focal distance.

The application was tested with data sets from five different Cellular Forms simulations. From each simulation four different files were used with approximately 1,000,000, 4,000,000, 16,000,000

and 67,000,000 cells (the maximum cell count from these growth simulations).

4. RESULTS

Using the test data and rendering a single view with a resolution of 1920x1080 pixels the application achieved the frame rates shown in Table 1.

Table 1: Frame rates for Ambient Occlusion and X-Ray renders at 1920x1080 pixels

Dataset	Number of cells	Ambient occlusion	X-Ray
14_0014_0129	1,027,163	91.2 fps	82.0 fps
	4,035,075	84.0 fps	29.8 fps
	16,105,577	71.6 fps	6.1 fps
	66,858,809	58.3 fps	2.6 fps
17_0012_0242	1,058,572	108.2 fps	37.7 fps
	4,266,966	100.1 fps	18.1 fps
	14,786,136	90.1 fps	0.7 fps
	67,108,864	69.3 fps	0.3 fps
18_0006_0014	1,033,490	93.0 fps	65.7 fps
	4,037,800	81.4 fps	28.2 fps
	16,035,996	70.1 fps	5.6 fps
	66,670,254	59.6 fps	1.7 fps
18_009_0017	1,027,135	82.3 fps	92.6 fps
	4,052,135	68.6 fps	35.5 fps
	16,168,631	52.1 fps	7.0 fps
	67,108,864	32.9 fps	2.4 fps
18_001_0003	1,032,328	91.1 fps	170.9 fps
	4,054,214	62.5 fps	75.4 fps
	16,035,995	42.1 fps	21.2 fps
	66,489,951	23.7 fps	8.3 fps

Based on a frame rate of 30fps or higher as being acceptable for interactive use, these results show that in all but one of the test cases ambient occlusion rendering could be done at acceptable frame rates even with the largest data sets of around 67,000,000 cells.

For X-Ray rendering the results are much more variable, with acceptable frame rates in all of the test cases when using around 1,000,000 cells, but performance degrading significantly with high numbers of cells to the level that with the largest data sets (around 67,000,000 cells) the frame rate was unsuitable for interactive use in all cases.

Calculations for the number of rays per second that the application was achieving were typically higher than 1 billion rays per second, with the highest recorded performance being 1.68 billion rays per second. Impressive though this is, it is significantly lower than the 10 billion rays per second that NVIDIA quotes for the GeForce RTX 2080 Ti. It appears that

when using OptiX these quoted frame rates are only achievable with triangle primitives rather than custom intersection programs (such as the one used for spheres in these tests).

From the tests conducted, it is the author's opinion that allowing the user to directly manipulate the orientation of the forms in 3D space appears to significantly improve the perception of depth, both when testing the system using a single computer monitor and with a stereoscopic display. Essentially the user gets an augmented version of depth from motion since they are consciously instigating the motion themselves. Being able to view forms from novel views also has the potential to increase the engagement of the user with the work, exploring looking at each form from new viewpoints in a similar manner to the way that users can engage with physical sculptures in a gallery setting.

Though the X-Ray rendering couldn't achieve acceptable frame rates with the largest data sets, the ability to actively change the orientation appears to have a particularly strong effect and works well with lower cell counts. The X-Ray visualization typically shows multiple layers of internal structure which can be challenging to interpret even when using stereopsis. Actively manipulating the orientation of X-Ray rendered forms has given the author a much better understanding of some of these structures, revealing aspects of the internal details of the forms that weren't previously apparent to him.

The tests using defocus blur were less successful. The stochastic method used traces a single eye ray per pixel for each rendered frame. This means that the same frame rate can be achieved when rendering with defocus blur as when rendering with a simple pinhole camera, but the results are generally initially very noisy before being progressively refined as long as the user doesn't change the rotation or focus settings. The amount of refinement to achieve a good visual quality appears to be in the order of seconds. This is too slow to be useful for enhancing depth perception.

The author is now considering testing a hybrid approach combining an initial render using OptiX without any defocus blur followed by a 2D post process using OpenGL shaders to simulate defocus effects. However, this technique will only work with opaque surfaces (such as the ambient occlusion renders) and not with renders that have multiple levels of transparency (such as the X-Ray renders).

5. CONCLUSION

The tests show that, depending on the rendering technique used, the ray tracing capabilities available

in current graphics hardware allow the rendering of complex organic forms made of millions of primitives at resolutions and frame rates that make direct interaction possible. In particular, the tests show that ambient occlusion renders of forms with over 67,000,000 sphere primitives can be achieved at high definition in real-time.

However, the specific rendering technique used can be critical. Tests rendering forms in a manner that simulated X-Rays passing through multiple layers of transparent material had variable performance, only achieving interactive frame rates with data sets of around 1,000,000 to 4,000,000 cells. This can be expected: the time to render is likely to be proportional to number of primitives that each ray goes through, so complex structures with a lot of internal layers are likely to have low performance.

In the author's view it appears that allowing the user to directly manipulate the orientation of objects in 3D space can significantly enhance depth perception. Tests allowing the user to interactively modify the plane of focus were less successful since the quality of renders that could be achieved interactively were too noisy. More work is planned on trying to improve this.

Real-time rendering of high-resolution images also raises the possibility of using ray tracing instead of OpenGL when working with virtual reality.

Improving the capabilities of GPUs to support ray tracing appears a key focus of the hardware manufacturers, so performance can be expected to significantly increase in the near future.

As ever, more research is needed, but the author believes that the results he is getting are already sufficiently interesting for him to propose using them to create exhibitable interactive artefacts.

6. REFERENCES

- Bentley, P.J. and Corne, D.W (eds.) (2002). *Creative Evolutionary Systems*. Academic Press, London.
- Cook, R. L., Porter, T., & Carpenter, L. (1984). Distributed ray tracing. In *Proceedings of the 11th annual conference on Computer graphics and interactive techniques* (pp. 137-145).
- Glassner, A. S. (ed.) (1989). *An introduction to ray tracing*. Elsevier.
- Howard, I. P. (2012). *Perceiving in depth*. Oxford University Press.
- Lomas, A. (2014a). Cellular Forms: An Artistic Exploration of Morphogenesis. *AISB-50*, London, UK, 1-4 April 2014.
- Lomas, A. (2014b). Plantlike Forms. <http://www.andylomas.com/plantlikeForms.html> (retrieved 13 March 2020).
- Lomas, A. (2015). Hybrid Forms. <http://www.andylomas.com/hybridForms.html> (retrieved 13 March 2020).
- Lomas, A. (2016). Species Explorer: An interface for artistic exploration of multi-dimensional parameter spaces. *EVA London 2016 Conference, Electronic Workshops in Computing (eWiC)*. London: BCS, 12–14 July 2016.
- Lomas, A. (2017). Mutant Vase Forms. <http://www.andylomas.com/mutantVaseForms.html>. (retrieved 13 March 2020).
- Lomas, A. (2019a). Morphogenetic Creations: Exhibiting and Collecting Digital Art. In Giannini, T., & Bowen, J. P. (eds.). *Museums and Digital Culture: New perspectives and research*. Springer.
- Lomas, A. (2019b). Morphogenetic Vase Forms. *ALIFE 2019: The 2019 Conference on Artificial Life* (pp. 523-530). MIT Press.
- McCormack, J. and Lomas, A. (2020). Understanding Aesthetic Evaluation using Deep Learning. *EvoMUSART 2020*, Held as Part of EvoStar 2020, Seville, Spain, 15-17 April 2020. Springer.
- Miller, G. (1994). Efficient algorithms for local and global accessibility shading. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques* (pp. 319-326).
- NVIDIA (2020). *GeForce RTX 2080 Ti*. <https://www.nvidia.com/en-gb/geforce/graphics-cards/rtx-2080-ti> (retrieved 11 March 2020).
- Parker, S. G., et al. (2010). OptiX: a general purpose ray tracing engine. *ACM transactions on graphics*, 29(4), 1-13.
- Todd, S. and Latham, W. (1992). *Evolutionary Art and Computers*. Academic Press, London.
- Turing, A. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 237: 37–72.
- Wikipedia Contributors (2019). *Depth Perception*. https://en.wikipedia.org/wiki/Depth_perception (retrieved 11 March 2020).
- Wikipedia Contributors (2020). *Stereoscope: Wheatstone Stereoscope*. https://en.wikipedia.org/wiki/Stereoscope#Wheatstone_stereoscope (retrieved 13 March 2020)