Synthesis of Abstract Dynamic Quasiperiodic 3D Forms using SIRENs

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

This paper explores using SIRENs, neural networks with periodic activation functions, as a means for synthesising abstract three-dimensional dynamic forms. A SIREN is used to generate a field function for an implicit surface, with inputs for 3D position and time. A wide range of complex quasiperiodic forms can be created, with synthesis and rendering being achievable at interactive rates using modern graphics hardware.

Synthesis of 3D structures. Implicit neural representations. Digital art. Generative art
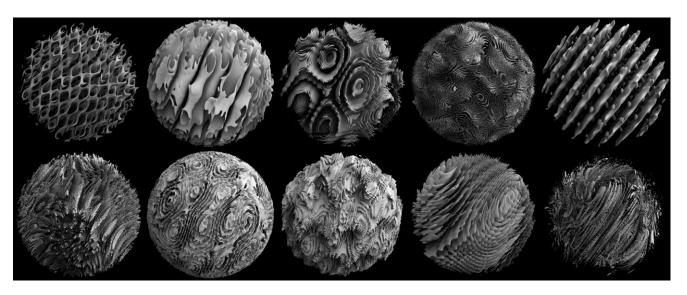
Figure 1: *Examples of 3D forms generated using SIREN networks*

1. INTRODUCTION

SIRENs, neural networks with periodic activation functions, were introduced by (Sitzmann et al. 2020). These networks are shown to have a number of useful properties compared with traditional neural networks that use activation functions such as sigmoid or rectified linear unit functions. In particular they are shown to be good at representing signals, such as audio or image data, including resolution of fine detail. One particularly useful property of SIRENs is that the derivative of a SIREN network is also a SIREN, which can be used to help solve systems with boundary conditions specified by partial differential equations. Sitzmann et al. also demonstrate the use of SIRENs to generate three-dimensional structures by calculating signed distance functions (Osher & Fedkiw 2003) from point cloud data, such as can be acquired from 3D scanning devices.

It is worth noting that there are similarities between SIRENS and synthesis of audio using phase modulation (Chowning 1973, Phase Modulation 2022). In phase modulation synthesis the output of one oscillator is fed into the input of another to

1

cause periodic phase offsets. One of the strengths of this form of synthesis is that a wide range of waveforms can be achieved with a relatively small number of operators. For instance: the Yamaha DX7, one of the most popular synthesisers from the 1980s, utilises only 6 periodic operators to create all the sounds that it generates. The similarity between SIRENs and phase modulation synthesis has been noted by (Janson 2021) and used to create a neural net based audio synthesiser.

Since SIRENs can be used to generate 3D structures as well as audio signals, this raises the idea of using SIRENs as a basis for synthesising novel three-dimensional forms rather than just to represent existing signals. Similar to the way that phase modulation synthesis can create a rich range of sounds with a small number of operators, a SIREN with a relatively small number of neurons may be capable of generating a wide range of different three-dimensional structures.

A common feature of music and certain types of art are motifs that recur but each time with variations. These appear to be forms that we respond to particularly strongly on an aesthetic level: signals that aren't completely regular but have patterns and structure that change over time and space. If the weights in a SIREN aren't simple rational multiples of each other, they have the potential to generate signals with these types of quasiperiodic patterns.

2. METHOD

To create time-varying 3D forms, a SIREN was implemented with the following structure:

- An input layer with 4 inputs: XYZ position and time.
- A number of fully connected hidden layers, each with the same number of neurons and using sinusoidal activation functions.
- An output layer with a linear activation function and a single output value.

The output value from the SIREN was used as a field function for an implicit surface, with the iso-surface at zero being used to create the 3D form.

The generation of the SIREN was parameterised using a number of hyper-parameters:

- Number of hidden layers (2 to 6).
- Hidden layer size (4 to 20).
- Hidden layer weight multiplier (1 to 100).
- Final layer weight multiplier (0.1 to 10).
- World scale multiplier (0.1 to 10).
- Time multiplier (0.1 to 1).
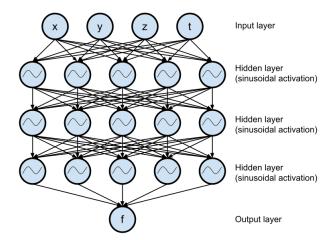- Power series ratio (0.5 to 1).

Figure 2: *Typical structure of SIREN used to generate 3D forms*

The weight and bias values for each neuron were generated using random values with ranges specified by the layer weight modifier values.

The Power Series Ratio parameter was used to specify a multiplication factor to be applied to the input weights of each successive neuron in each of the hidden layers. The intention was to ensure that the neurons in each layer would have a distribution of input weight values that would promote a range of frequency responses in the output.

In (Sitzmann et al. 2020) when generating 3D surfaces they impose a constraint using a partial differential equation to ensure that the field function generated by the SIREN is a signed distance function. This introduces an additional step where the weights in the SIREN are fitted to the boundary conditions using gradient descent. For this work a signed distance function wasn't considered necessary since all that was required was an iso-surface at zero which was rendered using marching cubes to create a triangle mesh. Relaxing this constraint allowed the SIRENs to be generated without fitting weights, allowing significantly faster generation of the SIREN from the hyper-parameters.

To make use of massively parallel processing, such as is available with modern graphics hardware, the SIREN network was implemented in CUDA. The extraction of the iso-surface from the SIREN was also implemented in CUDA, with the output of the SIREN evaluated over a regularly spaced 3D grid of voxels and using marching cubes (Lorensen 1987) to generate a triangle mesh. This data was then used in NVIDIA's OptiX ray-tracing library for rendering (Parker et al. 2010).

Additional final layer outputs could be optionally added to the SIREN to generate an RGB colour for each position in 3D space. This was used as a procedural 3D texture to change the diffuse colour

2

of the iso-surface during rendering. By being generated using the same network, the intention was that these colours would coherently change with the shape of the surface.
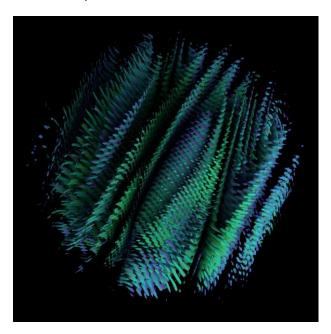

Figure 3: SIREN form with RGB colour

Additional hyper-parameters could be included to add an offset to the field function based on the radial distance from the world centre. This allowed forms to be created that have a limited bounding volume, since beyond a certain radius the field functions will either be strictly positive or strictly negative.
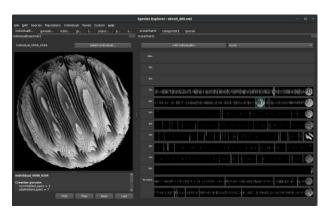

Figure 4: Selecting SIREN hyper-parameters using Species Explorer

To explore the range of possible outputs, the author used his Species Explorer software (Lomas 2016) to vary the hyper-parameters and sample results. The output was automatically classed as being a failure and allocated a score of zero if no triangles were generated when extracting the iso surface. All other cases were given a score value by the author on a scale from 1-10, which was used as a fitness value to generate new individuals using

interactive genetic algorithms and machine learning methods implemented in Species Explorer.

To evaluate the dynamic behaviour of the forms over time, extensions were added to Species Explorer to allow the viewing of animated .gif files for each form, as well as to launch a live simulation running the SIRENs in an interactive viewer.
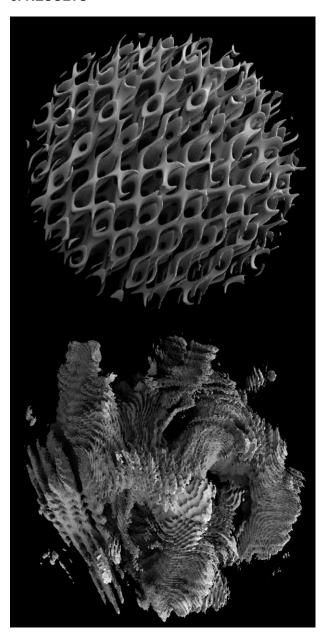
3. RESULTS

Figure 5: Generated forms showing varying amounts of regularity

The described system created a wide range of generated forms: from structures that appeared to have significant amounts of regularity through to forms that appeared to be far more noisy, irregular structures. In between these extremes the system

generated many forms with repeating motifs but where each repetition shows different patterns.

The computation of the forms (evaluating the SIREN in a voxel grid, generating triangles using marching cubes and rendering using OptiX) is sufficiently fast to allow animated structures to be generated in real-time if the resolution of the voxel grid is sufficiently small. A voxel grid with 128 cubes on each side (2,097,152 sample points) would generally allow updates at 30fps or faster using an RTX 2080 Ti GPU and rendering at 1024x1024 pixels. For higher quality renders, voxel grids of up to 512 cubes on each side (134,217,728 sample points) could be calculated while staying within the GPU's memory limitations with render times of around 1 second per frame.

The output of the system was tested in a variety of contexts including:

- A simple OpenGL viewer that allows the user to interactively move the form in 3D space using a virtual trackball interface while it updates dynamically.
- Anaglyph stereoscopic rendering to view the results in 3D.
- Stereoscopic viewing in a VR environment using Unreal Engine and an HTC Vive head mounted display.
- Rendering for presentation in a Fulldome environment, including with anaglyph stereoscopic 3D.

The use of field functions facilitates creating transitions between different forms by simply interpolating between the field functions before extracting the iso-surface. This works even if the surfaces generated by two SIRENs have different topologies. A video showing a series of such transitions can be seen at (Lomas 2022a), together with an anaglyph stereoscopic 3D version of the same transitions for viewing with red blue/cyan glasses at (Lomas 2022b).
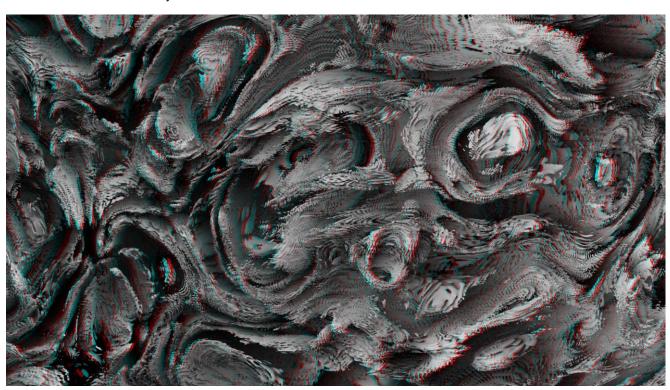

Figure 6: 3D anaglyph rendering of a SIREN form (for viewing with red blue/cyan glasses

4. DISCUSSSION

The ability to generate forms in real-time allows the potential to interactively manipulate parameters, including some of the hyper-parameters used to generate the SIREN. The author has tested this using a MIDI controller to update parameter values in real-time while viewing the results.

In some of the generated forms there are noticeable sampling artefacts. These are to be expected since the field function is being evaluated at a discrete set of regularly spaced points. Simply increasing the resolution of the voxel grid should improve this quality, but it should be noted that phase modulation techniques often generate high frequency sidebands that could exceed any discrete sampling level.

There are alternative rendering algorithms for iso-surfaces that could yield higher quality results but potentially at the cost of increased rendering time. For this work the author explicitly didn't use the constraint from (Sitzmann et al. 2020) to generate signed distance functions. This was done to make generation of the SIREN networks significantly faster. However, signed distance functions have potentially useful properties, including enabling use of rendering algorithms such as sphere tracing that can generate very high quality anti-aliased results (Hart 1996).

5. CONCLUSION

This study shows that SIRENs with a small number of neurons can be used to synthesise a wide range of different 3D forms with potentially interesting quasiperiodic structures. This can be seen as a technique for synthesising form in a similar manner to phase modulation synthesis of audio, creating rich complex forms from a small number of periodic operators. These are often visually interesting, creating dynamic structures with repeating but varying motifs.

Using current GPU technology, the results can be calculated in real-time, including manipulating hyper-parameters used to generate the SIRENs. This opens possibilities for live interaction, creating 3D forms that respond to user interactions or other input such as live music.

There are a number of directions that future work could take, including:

- Experimenting with different topologies of the SIREN network and different distributions of weights on the neurons.
- Audio synthesis using the same SIREN used to generate 3D forms to create audio directly connected to visual material.
- Implementation of a tool to generate forms using SIRENs for AV displays or for live mixing by VJs.

6. ACKNOLEDGEMENTS

7. REFERENCES

Chowning, J. M. (1973). The synthesis of complex audio spectra by means of frequency modulation. *Journal of the audio engineering society, 21*(7), 526-534.

Hart, J. C. (1996). Sphere tracing: A geometric method for the antialiased ray tracing of implicit surfaces. *The Visual Computer, 12*(10), 527-545.

Janson, A. (2021). *Implicit neural differentiable FM synthesizer*. Available at: https://pythonawesome.com/implicit-neural-differentiable-fm-synthesizer/ (retrieved 7 January 2022)

Lomas, A. (2016). Species Explorer: An interface for artistic exploration of multi-dimensional parameter spaces. *Electronic Visualisation and the Arts*, pp.95-102.

Lomas, A. (2022a). *Siren Transitions*. [online video] Available at: https://www.youtube.com/watch?v=_IbN_9t8ixk

(retrieved 8 January 2022)

Lomas, A. (2022b). *Siren Transitions (Anaglyph 3D)*. [online video] Available at: https://www.youtube.com/watch?v=fjnMWOvDG6o

(retrieved 8 January 2022)

Lorensen, W.E. and Cline, H.E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. *ACM SIGGRAPH computer graphics, 21(4)*, pp.163-169.

Osher, S. and Fedkiw, R. (2003). Signed distance functions. In *Level set methods and dynamic implicit surfaces* (pp. 17-22). Springer, New York, NY.

Parker, S.G., Bigler, J., Dietrich, A., Friedrich, H., Hoberock, J., Luebke, D., McAllister, D., McGuire, M., Morley, K., Robison, A. and Stich, M. (2010). OptiX: a general purpose ray tracing engine. *ACM transactions on graphics (tog)*, 29(4), pp.1-13.

Phase Modulation (2022). In *Wikipedia*. https://en.wikipedia.org/wiki/Phase_modulation (retrieved 21 February 2022)

Sitzmann, V., Martel, J.N., Bergman, A.W., Lindell, D.B. and Wetzstein, G. (2020). Implicit neural representations with periodic activation functions. *arXiv preprint arXiv:2006.09661*.