Visual Indeterminacy in Generative Neural Art

Aaron Hertzmann
Adobe Research
601 Townsend St.
San Francisco, CA 94103
hertzman@dgp.toronto.edu

1 Introduction

Generative Adversarial Networks (GANs) have become fertile tools for artistic exploration. Artists such as Refik Anadol, Robbie Barrat, Sofia Crespo, Mario Klingemann, Jason Salavon, Helena Sarin, and Mike Tyka generate fascinating imagery with models learned from natural imagery. These artists work in very different ways, but each has produced some work that shares a common GAN aesthetic: realistic, but unrecognizable.

Why are GANs such powerful tools for making art? This essay argues that GAN art often exhibits *visual indeterminacy*, a term coined by Pepperell [16]. GANs cause visual indeterminacy by creating plausible compositions and textures that nonetheless defy coherent explanation, and these are the GAN images often used in recent artworks. Because visual indeterminacy can be understood as a perceptual process [12], GANs provide a potential tool for both art and for neuroscience experiments based on perceptual uncertainty modeling.

2 Visual Indeterminacy and Aesthetic Experience

Often, the initial appearance of an image invites the viewer to investigate further, but the image confounds explanation. For some images, this investigation leads to an "Aha!" moment, where the viewer understands the structure of an image [13], e.g., they see a vivid 3D object where there had been abstract 2D shapes. This moment is pleasurable because the posterior distribution collapses: some understanding has been gained. But the image may also become less interesting as result. "Visual indeterminacy" describes images where the "Aha!" moment never happens, and the image continues to invite investigation.

In short, visual indeterminacy occurs in a "seemingly meaningful visual stimulus that denies easy or immediate identification" [17]; it is the "lack but promise of semantic stability" [12]. Ambiguity has been present in art since cave painting, but became particularly valued in the Modern art era [7]. Many recent studies connect these ideas to perceptual theory and neuroscience; see [12, 21] for reviews, and [9] for a discussion of visual ambiguity in terms of perceptual posterior probability.

3 Generating Visual Indeterminacy

Some GAN images are naturalistic; some look like unusual but realistic scenes, e.g., portions of Refik Anadol's "Machine Hallucinations," Mike Tyka "EONS", and Figure 1(left). Some images of humans or animals seem real and grotesque, like portions of Mario Klingemann's "Memories of Passersby I." But a dominant mode of GAN art is visual indeterminacy. It looks like there is a real scene being depicted in photorealistic detail, but what is it?

GANs seem predisposed to indeterminate, intriguing imagery. This can be seen by experimenting with Artbreeder [19] (Figure 1), formerly Ganbreeder. Images drawn from Ganbreeder have been exhibited as art [1].

2019 NeurIPS Workshop on Creativity for Learning and Design

Figure 1: Images from BigGAN [3] created with Ganbreeder (now called Artbreeder) [19]. The leftmost image is, basically, naturalistic. The other images are visually indeterminate: they appear realistic at first glance, and suggest various associations, but they do not yield coherent realistic interpretations on closer study. Image credits in the Appendix.

Why do GANs create indeterminate images so often? Recent results by Bau et al. [2] suggest that early GAN layers correspond to large-scale objects, and later layers capture fine-scale details and textures. This suggests that they construct scenes in pictorial space, first arranging "objects" into compositions, and then placing appropriate textures and details for those objects. Sometimes GANs compose textures in unexpected arrangements, composing realistic parts from familiar types of objects, in unfamiliar combinations. This produces visual indeterminacy. The images are most intriguing when they place familiar-looking elements into evocative but indeterminate configurations that, seemingly, no one would have created without these tools.

As a neural explanation for visual indeterminacy, Muth and Carbon [12] suggest that multiple local neural predictions fail to converge to a coherent interpretation. They quote Gombrich, who wrote about Cubism, "each hypothesis we assume will be knocked out by a contradiction elsewhere" [8].

Intriguingly, some of the paintings that Pepperell created years ago in search of visual indeterminacy [17] appear quite similar to GAN images. As a very unscientific experiment, I showed two of these paintings (Figure 2) to some colleagues, and asked "Can you guess how these images were made?" Four of the five responses hypothesized some kind of neural networks.

Eventually, generative networks may get so good that they rarely, if ever, produce unrealistic images. Perhaps there is an "Uncanny Ridge," along which generators are only just good enough to produce a diverse set of intriguingly indeterminate images, and past which the outputs are less and less surreal. Once generators pass this Uncanny Ridge, artists will need to find new ways to "break" the models, to coerce them into making new "errors."

4 Unifying Perception and Aesthetics

Natural image models, vision neuroscience, and image synthesis have long been tightly-coupled fields. Discoveries about the visual cortex [11] led to natural image statistics analysis [15, 20], which led to texture synthesis algorithms [5, 18] which led to style transfer algorithms [4, 10]. At the same time, cortical modeling also led to deep convolution networks [6], which led to GANs and trained discriminative networks, which, in turn, have led to improved neuroscience models [22]. Can aspects of aesthetic experience be understood with the same models?

Ideally, a generator would accurately represent a distribution over natural images; a recognition model would be the inverse, a providing a posterior distribution over interpretations that would approximate human perceptual uncertainty. Optimization against this model would give artists more precise control over the type of perceptual uncertainty present in images, for example, to produce images with specific types of visual indeterminacy.

This would provide artists with higher-level controls to explore artistic creation. It could also provide a richer testbed to develop perceptual theories of aesthetic experience, rather than using hand-crafted artworks as in [13, 17].

A more fine-grained neural model of indeterminacy is needed, to account for the temporal evolution from first impressions [14], to the movement of attention, to the possibility of the "Aha!" moment. Moreover, the categorization of perceptual ambiguity in art is very preliminary and much work remains to be done to expand upon and refine it.

References

[1] J. Bailey. Why is AI art copyright so complicated?, Mar. 2019. `https://www.artnome.com/news/2019/3/27/why-is-ai-art-copyright-so-complicated`.

[2] D. Bau, J.-Y. Zhu, H. Strobelt, B. Zhou, J. B. Tenenbaum, W. T. Freeman, and A. Torralba. GAN dissection: Visualizing and understanding generative adversarial networks. In *Proc. ICLR*, 2019.

[3] A. Brock, J. Donahue, and K. Simonyan. Large scale gan training for high fidelity natural image synthesis. In *Proc. ICLR*, 2019.

[4] A. A. Efros and W. T. Freeman. Image Quilting for Texture Synthesis and Transfer. In *Proc. SIGGRAPH*, 2001.

[5] A. A. Efros and T. Leung. Texture synthesis by non-parametric sampling. In *Proc. International Conference on Computer Vision*, 1999.

[6] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4), Apr 1980.

[7] D. Gamboni. *Potential Images: Ambiguity and Indeterminacy in Modern Art*. Reaktion Books, 2004.

[8] E. H. Gombrich. *Art and Illusion: A Study in the Psychology of Pictorial Representation*. Phaidon Press, 5th edition, 1960/2002.

[9] A. Hertzmann. Non-Photorealistic Rendering and the Science of Art. In *Proc. NPAR*, 2010.

[10] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin. Image Analogies. In *Proc. SIGGRAPH*, 2001.

[11] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1), 1968.

[12] C. Muth and C.-C. Carbon. Seins: Semantic instability in art. *Art & Perception*, 4(1–2):145–184, 2016.

[13] C. Muth, M. H. Raab, and C.-C. Carbon. Semantic stability is more pleasurable in unstable episodic contexts. on the relevance of perceptual challenge in art appreciation. *Frontiers in Human Neuroscience*, 10, 2016.

[14] A. Oliva and A. Torralba. Building the gist of a scene: the role of global image features in recognition. In *Visual Perception: Fundamentals of Awareness: Multi-Sensory Integration and High-Order Perception*, chapter 2. Elsevier, 2006.

[15] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381, 1996.

[16] R. Pepperell. Seeing without objects: Visual indeterminacy and art. *Leonardo*, 39, 2006.

[17] R. Pepperell. Connecting art and the brain: An artist's perspective on visual indeterminacy. *Frontiers in Human Neuroscience*, 5, 2011.

[18] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1), Oct 2000.

[19] J. Simon. Artbreeder (website), 2019. `artbreeder.com`.

[20] E. P. Simoncelli and B. Olshausen. Natural image statistic and neural representation. *Annu. Rev. Neurosci.*, 24, 2001.

[21] S. Van de Cruys and J. Wagemans. Putting reward in art: A tentative prediction error account of visual art. *i-Perception*, 2(9), 2011.

[22] D. L. K. Yamins and J. J. DiCarlo. Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, 19, 2016.