Do Silhouettes Dream?

Yannis Kalantidis
skamalas.com

Cheng Xu
1x1.design

Clayton Mellina
sigil.io

ABSTRACT
This paper complements our interactive art installation consisting of a back-projected screen and a depth sensor. As participants stand in front of the projection, their silhouettes create an area of hallucination that changes with their movement. Hallucinations may vary; some are "dreams", transporting the area covered by the silhouette to past or future times. Others escape reality and either alter the visual artistic style or form "deep dreams", i.e. hallucinations created through a feedback loop on artificial deep neural networks. Hallucinations vary in intensity and "depth" as participants move, increasingly distorting reality with proximity to the screen. To our knowledge, this installation is the first of its kind, placing fascinating results of cutting-edge artificial intelligence in a public embodied experience.

INTRODUCTION
Recent advances in computer vision have produced impressive visualizations of the internal representations of deep convolutional neural networks trained on images. In one such visualization method called *Deep Dreaming* [3], a trained network is induced to "hallucinate" by iteratively modifying the pixels of an input image to maximize internal neural activations. Different layers of neurons can be selectively targeted by the deep dreaming process. Indeed, by studying how different choice of layer produces different visual modification of the input image, we can gain insight and intuition for the network's representations. For instance, typically the visual patterns imposed on the image by maximizing layers nearer the input of the network resemble simple geometric strokes, while deeper layers produce recognizable parts of semantic classes, such as parts of buildings or animals.

C&C '17 June 27-30, 2017, Singapore, Singapore

ACM ISBN 978-1-4503-4403-6/17/06.

DOI: http://dx.doi.org/10.1145/3059454.3059499

How might a viewer understand this abstract continuum and build a mental model of it? Could an embodied experience help make these abstract concepts more approachable? This question motivated us to create our installation. In our work, we map the viewer's physical distance from the image to an abstract quality, enabling an embodied mode of exploration. We experiment with a variety of such mappings—or *hallucination modes*—including mapping distance to the 'depth' of dreams or the time dimension of a timelapse. The participants' silhouettes become visual vessels transported through time or hallucination intensity, and their bodies become windows into abstract spaces.

To our knowledge, this installation is the first of its kind, placing the recent, fascinating results of artificial intelligence in a public embodied experience. By exhibiting in a gallery setting, we further want to raise the question: in what contexts can visuals driven by learning algorithms be perceived as art?

INSTALLATION
Physically our installation consists of a large back-projected screen equipped with a Microsoft Kinect depth sensor. With no audience present, it would appear as a static image. As a participant's body is captured by the sensor in the area in front of the screen, hallucination effects emerge and become more prominent as they approach the projection screen. The depth at each point in the sensor space is translated to a quantity that depends on the current hallucination mode, which in turn determines the visual appearance at the corresponding point in the projection. For instance, depth maps to hallucination intensity in the deep dream mode and time in the timelapse mode. The images and hallucination modes may rotate during the exhibition, or multiple modes may be simultaneously present for different participants' silhouettes.

At first glance, the installation offers a visually complex interactive experience. Upon closer examination and exploration, it allows the audience to build a visual or even spatial mental model of how an image evolves in a hallucinating neural network, or how time sculpts a landscape. A short explanatory note that explains in greater detail how deep dream visualization works invites participants to learn about the structure of deep neural networks and to engage critically with the piece.

Figure 1. Style transfer example: Proximity controls artistic style. ***Creative Commons image courtesy of Mike Cartmell on Flickr*.**

Figure 2. With a day-to-night timelapse of Singapore as input, proximity controls the time of day for the areas covered by the silhouettes.

HALLUCINATION MODES

We experiment with multiple spatio-temporal inputs or *hallucination modes*. Some are dreams transporting the visual to past and future times, others escape reality and either distort the visual artistic style or visualize neural network deep dreaming.

Dreaming with neural networks. We experiment with different applications of artificial neural networks. An example for deep dreaming [3] is shown in the title Figure[1]. Physical depth can be mapped to different parameters of the deep dreaming process. We have experimented with mapping depth to both the number of iterations of the deep dreaming process and the depth of the neural layer directing the deep dreaming process. We are further experimenting with visualizing filter outputs, as well as class specific visualizations, *i.e.* starting from random noise images and forcing the network to hallucinate a specific semantic class.

Style transfer with neural networks. Another direction is artistic style transfer, depicted in Figure 1. DeepStyle [2] is a recent technique that allows *style transfer* between a source image (e.g. the landscape in Figure 1) and a style image (e.g. Van Gogh's "Starry Night" in the leftmost frame of Figure 1). Recent papers, *e.g.* [1], present extensions for not only jointly learning multiple styles, but also making it easy to interpolate between them.

Timelapse imagery. Controlling time through movement can be very empowering; we are experimenting with different inputs extracted from time-lapse videos. An example from a day-to-night panorama of Singapore [6][2] can be seen in Figure 2. Any time-lapse video can be used as input.

Beyond silhouettes. One may use the *average* distance/depth from the depth sensor to define a single depth value for the entire projection. In this setting, participants no longer appear as silhouettes, but are able to use coordinated movement to scroll the entire projection through time or neural network depth. A compelling application of this interaction is the visualization of timelapsed geographic imagery. For example, Google Earth recently published timelapses of different locations over 32 years [4], highlighting both climate and population changes over time. Similarly, visualizing Global Forest Cover Change data [5] in this way forces participants to confront metaphors between proximity and deforestation.

FUTURE WORK

We envision that the interactive entanglement of depth, space, and time can be used for more than visual exploration. As embodied actors, movement changes our physical perspective on the world and is deeply intuitive. Mapping movement to new perceptual transformations offers a powerful embodied interaction that may be used for narrative purposes or as a building block in more expansive experiences. We intend to pursue these directions in the future.

The python code used in this project is publicly available at `https://github.com/skamalas/depth2time`

REFERENCES

1. D. Chen *et al.* 2017. Stylebank: An explicit representation for neural image style transfer. *arXiv preprint arXiv:1703.09210* (2017).
2. L. Gatys *et al.* 2015. A Neural Algorithm of Artistic Style. *CoRR* abs/1508.06576 (2015).
3. Google. 2015. DeepDream blog post. (2015). `https://research.googleblog.com/2015/07/deepdream-code-example-for-visualizing.html`.
4. Google. 2016. Timelapse - Google Earth Engine. (2016). `https://earthengine.google.com/timelapse/`.
5. M. C. Hansen *et al.* 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342, 6160 (2013), 850–853.
6. Harrison Wood. 2015. Singapore City Day-Night Time lapse. YouTube. (2015). `https://www.youtube.com/watch?v=2UFUfj1Eb3c`.

[1] Creative Commons image, courtesy of Mike Cartmell on Flickr.

[2] Used with permission by Harrison Wood.