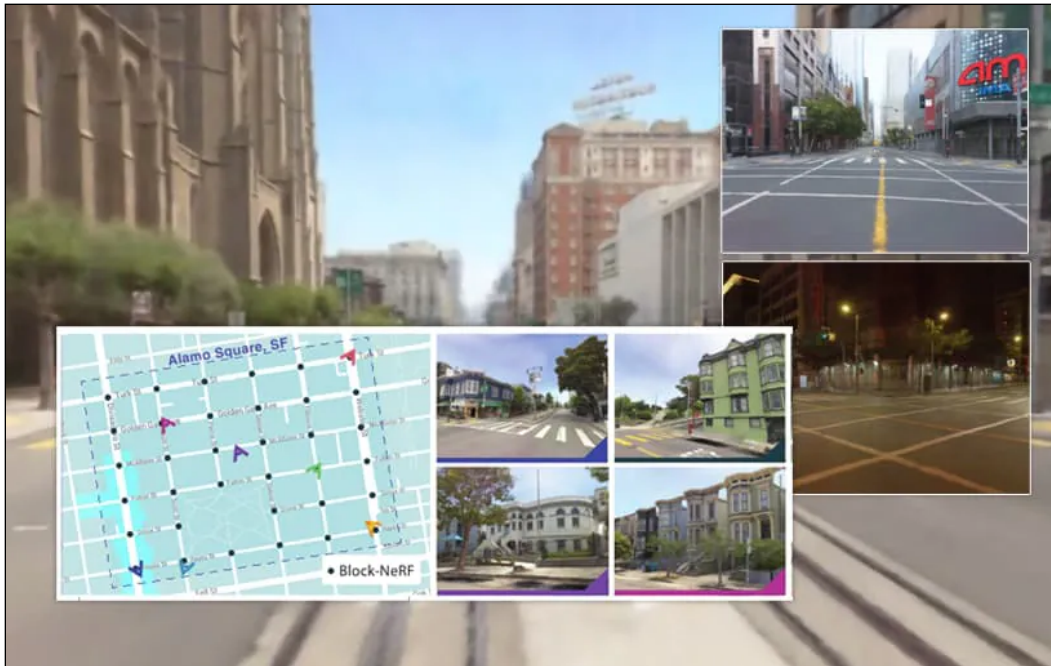


The New CGI: Creating Neural Neighborhoods With Block-NeRF

Originally published at <https://www.unite.ai/the-new-cgi-creating-neural-neighborhoods-with-block-nerf/>



Neural Radiance Fields ([NeRF](#)) allow objects to be recreated and explored inside neural networks using only multiple viewpoint photographs as input, without the complexity and expense of traditional CGI methods.

However, the process is computationally expensive, which initially limited NeRF environments to [tabletop model](#) scenarios. Nonetheless, NeRF has been adopted by a dedicated, even frantic research community, which has over the last year [enabled exterior reconstructions](#) as well as [editable neural humans](#), besides many other innovations.

Now a new research initiative, which includes the participation of Google Research, recognizes the possible hard limits on optimizing NeRF, and concentrates instead on stitching together NeRF environments to create on-demand neighborhoods comprising multiple coordinated NeRF instances.



CLICK TO VIEW ANIMATED GIF

Viewpoint from a Block-NeRF network of linked NeRFs. See embedded video at end of article, and also source link for high-resolution full-length supplementary videos. Source: <https://waymo.com/research/block-nerf/>

Navigating the network of linked NeRFs effectively makes NeRF scalable and modular, providing navigable environments which load extra parts of the neighborhood as they're needed, in a manner similar to the resource optimization methods of videogames, where what's around the corner is rarely loaded until it becomes clear that the environment is going to be needed.

In a major drive to [disentangle](#) separate facets such as weather and hour, Block-NeRF also introduces ‘appearance codes’, making it possible to dynamically change the time of day:



CLICK TO VIEW ANIMATED GIF



CLICK TO VIEW ANIMATED GIF

Changing the time of day with Block-NeRF. See embedded video at end of article, and also source link for high-resolution full-length supplementary videos. Source: <https://waymo.com/research/block-nerf/>

The new paper suggests that NeRF optimization is approaching its own thermal limit, and that future deployments of neural radiance environments in virtual reality, other types of interactive spheres, and VFX work, are likely to depend on parallel operations, similar to the way that Moore’s Law eventually gave way to multi-core architectures, parallel optimizations and new approaches to caching.

The authors of the [paper](#) (entitled *Block-NeRF: Scalable Large Scene Neural View Synthesis*) used 2.8 million images to create the largest neural scene ever attempted – a series of neighborhoods in San Francisco.



CLICK TO VIEW ANIMATED GIF

Block-NeRF navigates San Francisco’s Grace Cathedral. See embedded video at end of article, and also source link for high-resolution full-length supplementary videos. Source: <https://waymo.com/research/block-nerf/>

The lead author on the paper, representing UC Berkley, is Matthew Tancik, the [co-inventor of Neural Radiance Fields](#), who undertook the work while an intern at autonomous driving technology development company Waymo, host of the [project page](#). The initiative also offers a video overview at YouTube, embedded at the end of this article, besides many supporting and supplementary video examples at the project page.

The paper is co-authored by several other NeRF originators, including Ben Mildenhall (Google Research), Pratul P. Srinivasan (Google Research), and Jonathan T. Barron (Google Research). The other contributors are Vincent Casser, Xinchun Yan, Sabeek Pradhan, Henrik Kretzschmar and Vincent Casser, all from Waymo.

Block-NeRF was developed primarily as research into virtual environments for autonomous vehicle systems, including self-driving cars and drones.



CLICK TO VIEW ANIMATED GIF

The Embarcadero roadway from a 180-degree view stance in Block-NeRF. See embedded video at end of article, and also source link for high-resolution full-length supplementary videos. Source: <https://waymo.com/research/block-nerf/>

Other factors that can be dynamically changed in Block-NeRF are lens aperture (see image above), weather and seasons.

However, changing season can cause related changes in the environment, such as trees without leaves, which requires an even more extensive input dataset than was constructed for Block-NeRF. The paper states:

'[Foliage] changes seasonally and moves in the wind; this results in blurred representations of trees and plants. Similarly, temporal inconsistencies in the training data, such as construction work, are not automatically handled and require the manual retraining of the affected blocks.'

Apocalyptic Rendering

If you take a look at the video embedded at the end, you'll notice a *Walking Dead*-style sparseness to the networked Block-NeRF environment. For various reasons, not least to provide a simulated starter environment for robotic systems, cars, pedestrians, and other transient objects were deliberately matted out from source material, but this has left some artifacts behind, such as the shadows of 'erased' parked vehicles:



The phantom shadow of an erased car. Source: <https://waymo.com/research/block-nerf/>

To accommodate a range of lighting environments such as day or night, the networks have been trained to incorporate disentangled streams of data relating to each desired condition. In the image below, we see the contributing streams for Block-NeRF footage of a highway by day and by night:



The on-demand facets behind an apparently 'baked' Block-NeRF render, allowing a user to switch on the night as required. Source: <https://waymo.com/research/block-nerf/>

Environmental and Ethical Considerations

Over the last few years, research submissions have begun to include caveats and disclaimers regarding possible ethical and environmental ramifications of the proposed work. In the case of Block-NeRF, the authors note that the energy requirements are high, and that accounting for short-term and long-term transient objects (such as leaves on trees and construction work, respectively) would require regular re-scanning of the source data, leading to increased 'surveillance' in urban areas whose neural models need to be kept updated.

The authors state:

'Depending on the scale this work is being applied at, its compute demands can lead to or worsen environmental damage if the energy used for compute leads to increased carbon emissions. As mentioned in the paper, we foresee further work, such as caching methods, that could reduce the compute demands and thus mitigate the environmental damage.'

Regarding surveillance, they continue:

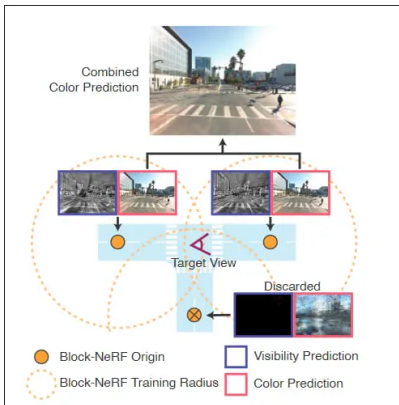
'Future applications of this work might entail even larger data collection efforts, which raises further privacy concerns. While detailed imagery of public roads can already be found on services like Google Street View, our methodology could promote repeated and more regular scans of the environment. Several companies in the autonomous vehicle space are also known to perform regular area scans using their fleet of vehicles; however some might only utilize LiDAR scans which can be less sensitive than collecting camera imagery.'

Methods and Solutions

The individual NeRF environments can be scaled down, in theory, to any size before being assembled into a Block-NeRF array. This opens the way to the granular inclusion of content that is definitely subject to change, such as trees, and to the identification and management of construction works, which may persist in time over even years of re-capture, but are likely to evolve and eventually become consistent entities.

However in this initial research outing, discrete NeRF blocks are limited to the actual city blocks of each depicted environment, stitched together, with a 50% overlap ensuring consistent transition from one block to the next as the user navigates the network.

Each block is constrained by a geographical filter. The authors note that this part of the framework is open to automation, and, surprisingly, that their implementation relies on OpenStreetMap rather than Google Maps.



The intersection radius for a Block-NeRF 'active' render space. Source: [Waymo](#)

Blocks are trained in parallel, with needed blocks rendered on demand. The innovative appearance codes are also orchestrated among the block-set, ensuring that one does not travel unexpectedly into different weather, time of day, or even a different season.



Block-NeRF segments are conditioned on exposure in a manner analogous to High Dynamic Range (HDR) in photographic source material. Source: [Waymo](#)

The ability to switch lighting and other environmental variables is derived from the Generative Latent Optimizations introduced in [NeRF in the Wild](#) (NeRF-W), which itself derived the method from the 2019 Facebook AI [research paper](#) *Optimizing the Latent Space of Generative Networks*.

A semantic segmentation model originated for [Panoptic-DeepLab](#) in 2020 is used to block out undesired elements (such as people and vehicles)

Data

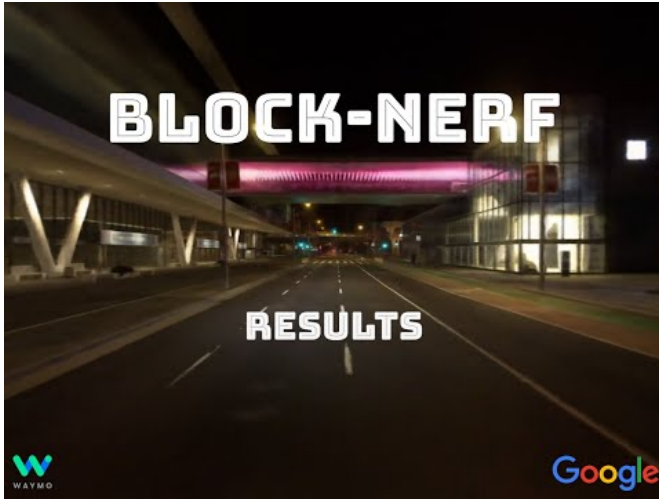
Finding that common urban datasets such as [CityScapes](#) were not suitable for such intensive detail-work as Block-NeRF entails, the researchers originated their own dataset. Image data was captured from 12 cameras encompassing a 360-degree view, with footage taken at 10 Hz with a scalar exposure value.



The San Francisco neighborhoods covered were Alamo Square and Mission Bay. For the Alamo Square captures, an area approximating 960m x 570m was covered, divided into 35 Block-NeRF instances, each trained on data from 38 to 48 different data collection runs, with a total drive time of 18-28 minutes.

The number of contributing images for each Block-NeRF ran between 64,575 to 108,216, and the overall driving time represented for this area was 13.4 hours across 1,330 different data collection runs. This resulted in 2,818,745 training images just for Alamo Square. See the paper for additional details on the data collection for Mission Bay.

Block-NeRF



□

First published 11th February 2022.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More