

Can Apple's HDR Augmented Reality Environments Solve Reflections for Neural Rendering?

Originally published at <https://www.unite.ai/can-apples-hdr-augmented-reality-environments-solve-reflections-for-neural-rendering/>



Apple's vigorous, long-term investment in Augmented Reality technologies is accelerating this year, with a new slate of developer tools to [capture and convert](#) real world objects into AR facets, and a [growing industry conviction](#) that dedicated AR eyewear is coming to support the immersive experiences that this blizzard of R&D can enable.

Among a tranche of new information on Apple's efforts in Augmented Reality, a [new paper](#) from the company's computer vision research division reveals a method for using 360-degree panoramic high dynamic range (HDR) images to provide scene-specific reflections and lighting for objects that are superimposed into augmented reality scenes.

Entitled *HDR Environment Map Estimation for Real-Time Augmented Reality*, the paper, by Apple Computer Vision Research Engineer Gowri Somanath and Senior Machine Learning Manager Daniel Kurz, proposes the dynamic creation of real-time HDR environments via a convolutional neural network (CNN) running in a mobile processing environment. The result is that reflective objects can literally mirror novel, unseen environments on demand:



CLICK TO VIEW ANIMATED GIF

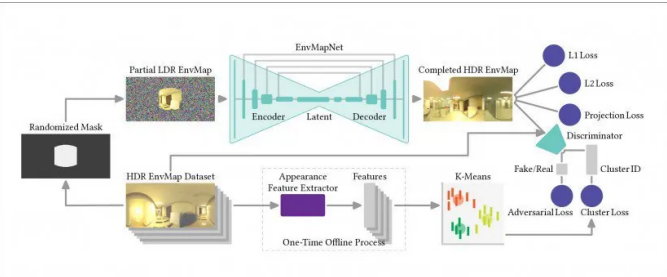
In Apple's new AR object generation workflow, a pressure cooker is instantiated by photogrammetry complete with its ambient environment, leading to convincing reflections that aren't 'baked' into the texture. Source: <https://docs-assets.developer.apple.com/>

The method, debuted at CVPR 2021, takes a snapshot of the entire scene and uses the [EnvMapNet](#) CNN to estimate a visually complete panoramic HDR image, also known as a ‘light probe’.



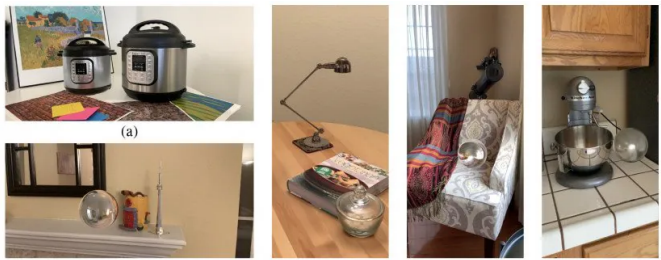
CLICK TO VIEW ANIMATED GIF

The resulting map identifies strong light sources (outlined at the end in the above animation) and accounts for them in rendering the virtual objects.



The architecture of EnvMapNet, which processes limited imagery into full-scene HDR light probes. Source: <https://arxiv.org/pdf/2011.10687.pdf>

The algorithm can run in under 9ms on an iPhone XS, and is capable of rendering reflection-aware objects in real time, with reduced directional error of 50% compared to previous and different approaches to the problem.



Light Probes

HDR lighting environments have been a factor in visual effects since high dynamic range images (invented in 1986) became a notable force through advances in computer technology in the 1990s. Anyone watching behind-the-scenes footage may have noticed the surreal on-set presence of technicians holding up mirrored balls on sticks – reference images to be incorporated as environmental factors when reconstructing CGI elements for the scene.

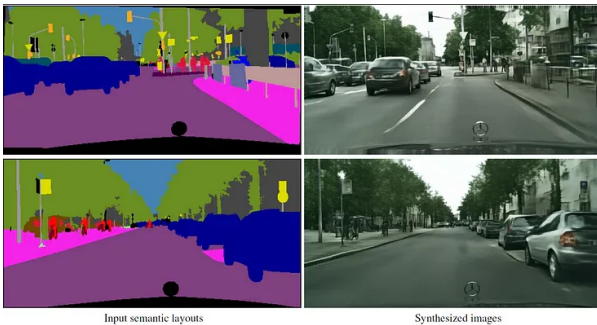


Source: <https://beforesandafters.com/>

However, using chrome balls for [reflection mapping](#) textures predates the 1990s, going back to the 1983 SIGGRAPH paper *Pyramidal Parametrics*, which featured still images of a reflective CGI robot in a style that would become famous nearly a decade later through via the 'liquid metal' effects of James Cameron's *Terminator 2: Judgement Day*.

HDR Environments In Neural Rendering?

Neural rendering offers the possibility to generate photorealistic video from very sparse input, including crude segmentation maps.



Intel ISL's *segmentation>image neural rendering* (2017). Source: <https://awesomeopensource.com/project/CQFIO/PhotographicImageSynthesis>

In May, Intel researchers [revealed](#) a new initiative in neural image synthesis where footage from Grand Theft Auto V was used to generate photorealistic output based on datasets of German street imagery.



CLICK TO VIEW ANIMATED GIF

Source: <https://www.youtube.com/watch?v=0fhUJT21-bs>

The challenge in developing neural rendering environments that can be adapted to various lighting conditions is to separate out the object content from the environmental factors that affect it.

As it stands, reflections and anisotropic effects remain functions either of the original dataset footage (which makes them inflexible), or requires the same type of schema that the Intel researchers employed, which generates semi-photorealistic output from a crude (game) engine, performs segmentation on it and then applies style transfer from a 'baked' dataset (such as the German Mapillary street view set used in the recent research).

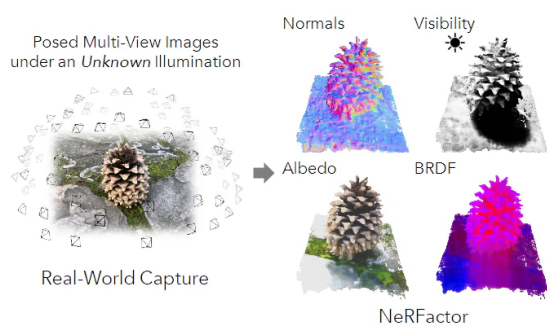


CLICK TO VIEW ANIMATED GIF

In this neural rendering derived from GTA V footage (left), the vehicle in front demonstrates convincing glare and even saturates the sensor of the fictitious virtual camera with reflections from the sun. But this lighting aspect is derived from the lighting engine of the original game footage, since the neural facets in the scene have no autonomous and self-referring lighting structures that can be changed.

Reflectance In NeRF

Imagery derived from [Neural Radiance Fields](#) (NeRF) is similarly challenged. Though recent research into NeRF has made strides in separating out the elements that go to make a neural scene (for example, the MIT/Google [collaboration on NeRFactor](#)), reflections have remained an obstacle.



The MIT and Google NeRFactor approach separates out normals, visibility (shadows), texture and local albedo, but it does not reflect a broader (or moving) environment, because it essentially exists in a vacuum. Source: <https://arxiv.org/pdf/2106.01970.pdf>

NeRF can solve this problem with the same kind of HDR mapping that Apple is using. Each pixel in a neural radiance field is calculated on a trajectory from a virtual camera up to the point where the 'ray' can travel no further, similar to ray-tracing in traditional CGI. Adding HDR input to the calculation of that ray is a potential method to achieve genuine environmental reflectance, and is in effect an analogue to CGI's 'global illumination' or radiosity rendering methods, wherein a scene or object is partially lit by perceived reflections of its own environment.

Though it's guaranteed that an HDR matrix won't do anything to ease NeRF's notable computational burdens, a great deal of [research](#) in this field at the moment is concentrating on addressing this aspect of the processing pipeline. Inevitably, reflectance is one of the many factors waiting in the wings to re-fill and challenge that newly-optimized architecture. However, NeRF can't achieve its full potential as a discrete neural image and video synthesis methodology without adopting a way to account for a surrounding environment.

Reflectance In Neural Rendering Pipelines

In a putative HDR-enabled version of the Intel GTA V neural rendering scenario, a single HDR could not accommodate the dynamic reflections that need to be expressed in moving objects. For instance, in order to see one's own vehicle reflected in the vehicle in front as it pulls up to the lights, the front vehicle entity could have its own animated HDR light probe, the resolution of which would degrade incrementally as it recedes from the end user's point of view, to become low-res and merely representative as it pulls away into the distance – a proximity-based LOD similar to 'draw distance' delimiters in video games.

The real potential of Apple's work in HDR lighting and reflection maps is not that it is particularly innovative, since it builds on previous work in general image synthesis and in [AR scene development](#). Rather, the possible breakthrough is represented by the way that severe local computing restraints have combined with Apple's M-series machine learning hardware innovations to produce lightweight, low-latency HDR mapping that's designed to operate under constrained resources.

If this problem can be solved economically, the advent of semantic segmentation>photorealistic video synthesis may come a significant step closer.



Source: <https://docs-assets.developer.apple.com/>

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