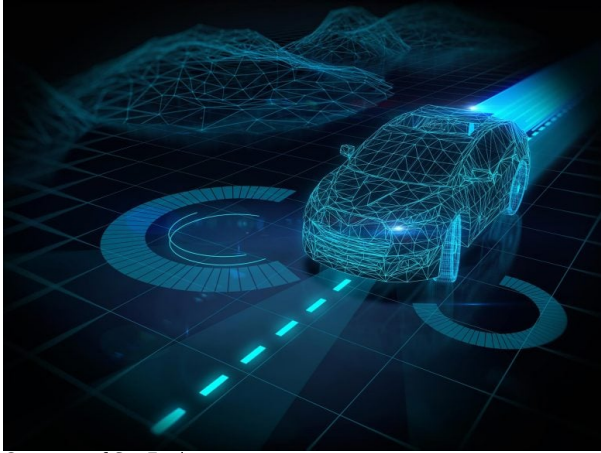


# New Technology Enables Automobiles to See Around Corners

Posted by [Okachinepa](#) 07/10/2024

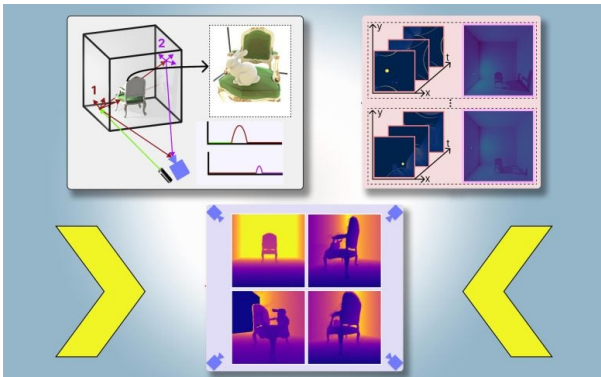


Courtesy of SynEvol  
Credit: MIT News

Envision yourself in an autonomous car going through a tunnel when you have no idea that there has been a crash that has stopped traffic ahead of you. Usually, you would have to depend on the vehicle ahead of you to signal when to apply the brakes. What if, however, your car could see around the car in front of it and apply the brakes even earlier?

One day, an autonomous car might be able to accomplish that thanks to a computer vision method created by researchers at MIT and Meta.

They have presented a technique that uses photos from a single camera position to generate physically accurate, 3D models of an entire environment, even parts obscured from view. Their method makes use of shadows to infer what is hidden from view in parts of the scene.



Courtesy of SynEvol  
Credit: MIT News

Based on a passage from the Greek philosopher Plato's "Republic," in which inmates shackled in a cave use shadows thrown on the cave wall to determine the reality of the outside world, they have named their method PlatoNeRF.

PlatoNeRF is able to produce 3D geometry reconstructions that are more accurate than those produced by certain other AI methods by fusing machine learning with lidar (light detection and ranging) technology. Furthermore, PlatoNeRF performs more fluidly while reconstructing scenes with high ambient light or dark backgrounds, when shadows are difficult to see.

PlatoNeRF has the potential to increase the efficiency of AR/VR headsets and improve the safety of autonomous vehicles by allowing users to model a room's geometry without having to wander around taking measurements. Additionally, it might make it easier for warehouse robots to locate objects in congested areas.

"Our main concept involved combining multibounce lidar and machine learning, two previously conducted projects in separate fields. Research assistant in the Camera Culture Group of the MIT Media Lab and lead author of a paper on PlatoNeRF, Tzofi Klinghoffer, an MIT graduate student in media arts and sciences, explains, "It turns out that when you bring these two together, that is when you find a lot of new opportunities to explore and get the best of both worlds."

Together with senior author Rakesh Ranjan, director of AI research at Meta Reality Labs, advisor Ramesh Raskar, associate professor of media arts and sciences at MIT, research assistant Siddharth Somasundaram in the Camera Culture Group, and Meta's Xiaoyu Xiang, Yuchen Fan, and Christian Richardt, Klinghoffer wrote the paper. The Conference on Computer Vision and Pattern Recognition will host a presentation of the findings.

It's a challenging task to reconstruct a whole 3D scene from a single camera viewpoint.

Certain machine learning techniques utilize generative artificial intelligence models to approximate the contents of obscured areas; nevertheless, these models have the potential to provide false positives for items. Other methods use shadows in a color image to try and deduce the forms of things that are hidden, but they have trouble working in situations when the shadows are difficult to see.

The MIT researchers expanded on these strategies for PlatoNeRF by utilizing single-photon lidar, a novel sensing technology. By generating light

pulses and timing how long it takes for the light to return to the sensor, lidars can map a three-dimensional scene. Single-photon lidars offer higher-resolution data since they are able to detect individual photons.

The researchers light a target point in the scene with a single-photon lidar. A portion of the light strikes that spot and returns straight to the sensor. However, before returning to the sensor, the majority of the light disperses and bounces off other things. PlatoNeRF depends on these light bounces.

PlatoNeRF measures the amount of time it takes for light to bounce twice before returning to the lidar sensor in order to obtain depth information about the scene. Information regarding shadows is also contained in the second bounce of light.

To identify which spots in the scene are under shadow (because there is no light), the system tracks the secondary light rays that leave the target point and go to other locations. PlatoNeRF can deduce the geometry of hidden objects based on the placement of these shadows.

The full 3D picture is reconstructed using several photos that are captured by the lidar as it sequentially lights 16 spots.

"We are casting fresh shadows in the scene each time we light a point. The area that is obscured and sits outside the range of the human sight is being defined by the numerous light sources that are available to us, according to Klinghoffer.

The combination of multibounce lidar and a unique kind of machine-learning model called a neural radiance field (NeRF) is essential to PlatoNeRF. A NeRF gives the model a strong ability to interpolate, or estimate, new views of a scene by encoding the geometry of the scene into the weights of a neural network.

When paired with multibounce lidar, this interpolation capability yields extremely precise picture reconstructions as well, according to Klinghoffer.

"The most difficult task was determining how to integrate these two elements. With multibounce lidar, we really have to consider the physics of light propagation and how to simulate that using machine learning," he explains.

Two popular alternatives, one that just employs lidar and the other that only uses a NeRF with a color image, were compared to PlatoNeRF.

They discovered that their approach may perform better than both, particularly in cases when the resolution of the lidar sensor was lower. Given that lower resolution sensors are frequently used in commercial devices, their method would be more feasible to implement in the real world.

Our team created the first camera that could "see" around corners about 15 years ago. It does this by taking advantage of multiple light bounces, or "echoes of light." These methods made use of three light bounces as well as specialized lasers and sensors. Since then, lidar technology has proliferated, which is what motivated us to do our research on fog-seeking cameras. Because this new technique only uses two light bounces, the 3D reconstruction quality is remarkable and the signal to noise ratio is quite high, according to Raskar.

The researchers hope to track more than two light bounces in the future to determine if that helps with scene reconstruction. To obtain texture information, they are also interested in utilizing more deep learning approaches and combining PlatoNeRF with measurements from color images.

Although 3D reconstruction from shadow camera images has long been researched, this work revisits the issue in the context of lidar and shows notable advances in the accuracy of hidden geometry reconstruction. The work demonstrates how innovative algorithms can enable extraordinary capabilities when coupled with common sensors, such as the lidar systems that many of us carry around in our pockets, according to David Lindell, an assistant professor in the University of Toronto's Department of Computer Science who was not involved in this study.