

Computer vision technique leverages reflections to image the world

May 10 2023, by Adam Zewe



Researchers from MIT and Rice University have created a computer vision technique that leverages reflections to image the world by using them to turn glossy objects into "cameras," enabling a user to see the world as if they were looking through the "lenses" of everyday objects like a ceramic coffee mug or a metallic paper weight. Credit: Massachusetts Institute of Technology

As a car travels along a narrow city street, reflections off the glossy paint or side mirrors of parked vehicles can help the driver glimpse things that would otherwise be hidden from view, like a child playing on the sidewalk behind the parked cars.

Drawing on this idea, researchers from MIT and Rice University have created a computer vision technique that leverages reflections to image the world. Their method uses reflections to turn glossy objects into "cameras," enabling a user to see the world as if they were looking through the "lenses" of everyday objects like a ceramic coffee mug or a metallic paper weight.

Using images of an object taken from different angles, the technique converts the surface of that object into a virtual sensor that captures reflections. The AI system maps these reflections in a way that enables it to estimate depth in the [scene](#) and capture novel views that would only be visible from the object's perspective. One could use this technique to see around corners or beyond objects that block the observer's view.

This method could be especially useful in autonomous vehicles. For instance, it could enable a self-driving car to use reflections from objects it passes, like lamp posts or buildings, to see around a parked truck.

"We have shown that any surface can be converted into a sensor with this formulation that converts objects into virtual pixels and virtual sensors. This can be applied in many different areas," says Kushagra Tiwary, a graduate student in the Camera Culture Group at the Media Lab and co-lead author of a paper on this research.

Tiwary is joined on the paper by co-lead author Akshat Dave, a graduate student at Rice University; Nikhil Behari, an MIT research support associate; Tzofi Klinghoffer, an MIT graduate student; Ashok Veeraraghavan, professor of electrical and computer engineering at Rice

University; and senior author Ramesh Raskar, associate professor of media arts and sciences and leader of the Camera Culture Group at MIT. The research will be presented at the [Computer Vision and Pattern Recognition Conference](#) held in Vancouver, June 18–22. A pre-print version is available on the *arXiv* server.

Reflecting on reflections

The heroes in crime television shows often "zoom and enhance" surveillance footage to capture reflections—perhaps those caught in a suspect's sunglasses—that help them solve a crime.

"In real life, exploiting these reflections is not as easy as just pushing an enhance button. Getting useful information out of these reflections is pretty hard because reflections give us a distorted view of the world," says Dave.

This distortion depends on the shape of the object and the world that object is reflecting, both of which researchers may have incomplete information about. In addition, the glossy object may have its own color and texture that mixes with reflections. Plus, reflections are two-dimensional projections of a three-dimensional world, which makes it hard to judge depth in reflected scenes.

The researchers found a way to overcome these challenges. Their technique, known as ORCa (which stands for Objects as Radiance-Field Cameras), works in three steps. First, they take pictures of an object from many vantage points, capturing multiple reflections on the glossy object.

Then, for each image from the real camera, ORCa uses machine learning to convert the surface of the object into a virtual sensor that captures light and reflections that strike each virtual pixel on the object's surface.

Finally, the system uses virtual pixels on the object's surface to model the 3D environment from the point of view of the object.

Catching rays

Imaging the object from many angles enables ORCa to capture multiview reflections, which the system uses to estimate depth between the glossy object and other objects in the scene, in addition to estimating the shape of the glossy object. ORCa models the scene as a 5D radiance field, which captures additional information about the intensity and direction of light rays that emanate from and strike each point in the scene.

The additional information contained in this 5D radiance field also helps ORCa accurately estimate depth. And because the scene is represented as a 5D radiance field, rather than a 2D image, the user can see hidden features that would otherwise be blocked by corners or obstructions.

In fact, once ORCa has captured this 5D radiance field, the user can put a virtual camera anywhere in the scene and synthesize what that camera would see, Dave explains. The user could also insert virtual objects into the environment or change the appearance of an object, such as from ceramic to metallic.

"It was especially challenging to go from a 2D image to a 5D environment. You have to make sure that mapping works and is physically accurate, so it is based on how light travels in space and how light interacts with the environment. We spent a lot of time thinking about how we can model a surface," Tiwary says.

Accurate estimations

The researchers evaluated their technique by comparing it with other

methods that model reflections, which is a slightly different task than ORCa performs. Their method performed well at separating out the true color of an object from the reflections, and it outperformed the baselines by extracting more accurate object geometry and textures.

They compared the system's depth estimations with simulated ground truth data on the actual distance between objects in the scene and found ORCa's predictions to be reliable.

"Consistently, with ORCa, it not only estimates the environment accurately as a 5D image, but to achieve that, in the intermediate steps, it also does a good job estimating the shape of the object and separating the reflections from the object texture," Dave says.

Building off of this proof-of-concept, the researchers want to apply this technique to drone imaging. ORCa could use faint reflections from objects a drone flies over to reconstruct a scene from the ground. They also want to enhance ORCa so it can utilize other cues, such as shadows, to reconstruct hidden information, or combine reflections from two objects to image new parts of a scene.

"Estimating specular reflections is really important for seeing around corners, and this is the next natural step to see around corners using faint reflections in the scene," says Raskar.

"Ordinarily, shiny objects are difficult for vision systems to handle. This paper is very creative because it turns the longstanding weakness of object shininess into an advantage. By exploiting environment reflections off a shiny object, the paper is not only able to see hidden parts of the scene, but also understand how the scene is lit. This enables applications in 3D perception that include, but are not limited to, an ability to composite virtual objects into real scenes in ways that appear seamless, even in challenging lighting conditions," says Achuta Kadambi,

assistant professor of electrical engineering and computer science at the University of California at Los Angeles, who was not involved with this work.

"One reason that others have not been able to use shiny objects in this fashion is that most prior works require surfaces with known geometry or texture. The authors have derived an intriguing, new formulation that does not require such knowledge."

More information: Kushagra Tiwary et al, ORCa: Glossy Objects as Radiance Field Cameras, *arXiv* (2022). [DOI: 10.48550/arxiv.2212.04531](https://doi.org/10.48550/arxiv.2212.04531)

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: Computer vision technique leverages reflections to image the world (2023, May 10) retrieved 3 May 2024 from <https://techxplore.com/news/2023-05-vision-technique-leverages-image-world.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--