



The New Science of Seeing Around Corners

Computer vision researchers have uncovered a world of visual signals hiding in our midst, including subtle motions that betray what's being said and faint images of what's around a corner.

By Natalie Wolchover

While vacationing on the coast of Spain in 2012, the computer vision scientist [Antonio Torralba](#) noticed stray shadows on the wall of his hotel room that didn't seem to have been cast by anything. Torralba eventually realized that the discolored patches of wall weren't shadows at all, but rather a faint, upside-down image of the patio outside his window. The window was acting as a pinhole camera — the simplest kind of camera, in which light rays pass through a small opening and form an inverted image on the other side. The resulting image was barely perceptible on the light-drenched wall. But it struck Torralba that the world is suffused with visual information that our eyes fail to see.

"These images are hidden to us," he said, "but they are all around us, all the time."

The experience alerted him and his colleague, [Bill Freeman](#), both professors at the Massachusetts Institute of Technology, to the ubiquity of "accidental cameras," as they call them: windows, corners, houseplants and other common objects that create subtle images of their surroundings. These images, as much as 1,000 times dimmer than everything else, are typically invisible to the naked eye. "We figured out ways to pull out those images and make them visible," Freeman explained.

The pair discovered just how much visual information is hiding in plain sight. In [their first paper](#), Freeman and Torralba showed that the changing light on the wall of a room, filmed with nothing fancier than an iPhone, can be processed to reveal the scene outside the window. Last fall, they and their collaborators [reported](#) that they can spot someone moving on the other side of a corner by filming the ground near the corner. This summer, they [demonstrated](#) that they can film a houseplant and then reconstruct a three-dimensional image of the rest of the room from the disparate shadows cast by the plant's leaves. Or they can turn the leaves into a "[visual microphone](#)," magnifying their vibrations to listen to what's being said.

"Mary had a little lamb..." a man says, in audio reconstructed from the motion of an empty chip bag that the scientists filmed through a soundproof window in 2014. (These were the first words Thomas Edison recorded with a phonograph in 1877.)

Research on seeing around corners and inferring information that's not directly visible, called "non-line-of-sight imaging," took off in 2012 with Torralba and Freeman's accidental-camera paper and [another watershed paper](#) by a separate group at MIT led by [Ramesh Raskar](#). In 2016, partly on the strength of those results, the Defense Advanced Research Projects Agency (DARPA) launched the \$27 million REVEAL program (for "Revolutionary Enhancement of Visibility by Exploiting Active

Light-fields”), providing funding to a number of nascent labs around the country. Since then, a stream of new insights and mathematical tricks has been making non-line-of-sight imaging ever more powerful and practical.



[Abe Davis](#)

<https://d2r55xnwy6nx47.cloudfront.net/uploads/2018/08/Hidden-Images-Audio1.mp3>

This recording of “Mary Had a Little Lamb” was reconstructed from the minuscule vibrations of an empty chip bag as viewed through a soundproof window.

Along with obvious military and spying applications, researchers rattle off possible uses in self-driving cars, robotic vision, medical imaging, astronomy, space exploration and search-and-rescue missions.

Torralba said he and Freeman didn’t have any particular application in mind when they started down this road. They were simply digging into the basics of how images form and what constitutes a camera, which naturally led to a fuller investigation of how light behaves and how it interacts with objects and surfaces in our environment. They began seeing things that no one had thought to look for. Psychological studies have shown, Torralba noted, “that humans are really terrible at interpreting shadows. Maybe one of the reasons is that many of the things that we see are not really shadows. And eventually the eye gave up on trying to make sense of them.”

Accidental Cameras

Light rays carrying images of the world outside our field of view constantly strike walls and other surfaces and reflect into our eyes. But why are these visual residues so weak? The answer is that there are too many of these light rays traveling in too many different directions. They wash out.

Forming an image requires greatly restricting the light rays that fall on a surface, allowing one particular set of them to be seen. This is what a pinhole camera does. Torralba and Freeman’s initial insight in 2012 was that there are many objects and features of our environment that naturally restrict light rays, forming faint images that are strong enough for computers to detect.

The smaller the aperture of a pinhole camera, the sharper the resulting image, since each point on the imaged object will emit only a single light ray with the correct angle to make it through the hole. Torralba’s hotel-room window was too big to produce a sharp image, and he and Freeman knew that, in general, useful accidental pinhole cameras are rare. They realized, however, that “anti-pinhole” (or “pinspeck”) cameras, consisting of any small, light-blocking object, form images all over the place.

Imagine you’re filming the interior wall of a room through a crack in the window shade. You can’t see much. Suddenly, a person’s arm pops into your field of view. Comparing the intensity of light on the wall when the arm is and isn’t present reveals information about the scene. A set of light rays that strikes the wall in the first video frame is briefly blocked by the arm in the next. By subtracting the data in the second image from that of the first, Freeman said, “you can pull out what was blocked by the arm” — a set of light rays that represents an image of part of the room. “If you allow yourself to look at things that block light, as well as things that let in light,” he said, “then you can

expand the repertoire of places where you can find these pinhole-like images.”

Along with the accidental-camera work aimed at picking up on small intensity changes, Freeman and his colleagues also devised algorithms for detecting and amplifying subtle color changes, such as those in a human face as blood pumps in and out, as well as tiny motions — the trick behind talking chip bags. They can now easily spot motions as subtle as one-hundredth of a pixel, which would normally be buried in noise. Their method is to mathematically transform images into configurations of sine waves. Crucially, in the transformed space, the signal isn't dominated by noise, because the sine waves represent averages over many pixels and so the noise is spread out among them. The researchers can therefore detect shifts in the positions of sine waves from one frame of a video sequence to the next, amplify these shifts, and then transform the data back.

They've now started to combine these various tricks for accessing hidden visual information. In research reported last October led by Freeman's then graduate student, [Katie Bouman](#) (now at the Harvard-Smithsonian Center for Astrophysics), they showed that the corners of buildings act as cameras that create rough images of what's around the corner.

Like pinholes and pinspecks, edges and corners also restrict the passage of light rays. Using conventional recording equipment, even iPhones, in broad daylight, Bouman and company filmed a building corner's "penumbra": the shadowy area that is illuminated by a subset of the light rays coming from the hidden region around the corner. If there's a person in a red shirt walking there, for example, the shirt will project a tiny amount of red light into the penumbra, and this red light will sweep across the penumbra as the person walks, invisible to the unaided eye but clear as day after processing.

In groundbreaking work reported in June, Freeman and colleagues reconstructed the "light field" of a room — a picture of the intensity and direction of light rays throughout the room — from the shadows cast by a leafy plant near the wall. The leaves act as pinspeck cameras, each blocking out a different set of light rays. Contrasting each leaf's shadow with the rest reveals its missing set of rays and thus unlocks an image of part of the hidden scene. Accounting for parallax, the researchers can then piece these images together.

This light-field approach yields far crisper images than earlier accidental-camera work, because prior knowledge of the world is built into the algorithms. The known shape of the houseplant, the assumption that natural images tend to be smooth, and other "priors" allow the researchers to make inferences about noisy signals, which helps sharpen the resulting image. The light-field technique "requires knowing a lot about the environment to do the reconstruction, but it gives you a lot of information," Torralba said.

Scattered Light

While Freeman, Torralba and their protégés uncover images that have been there all along, elsewhere on the MIT campus, Ramesh Raskar, a TED-talking computer vision scientist who explicitly aims to "change the world," takes an approach called "active imaging": He uses expensive, specialized camera-laser systems to create high-resolution images of what's around corners.



Courtesy of ACM SIGGRAPH 2017

Ramesh Raskar, a computer vision scientist at MIT, pioneered a technique for active non-line-of-sight imaging.

In 2012, realizing an idea he had five years earlier, Raskar and his team pioneered a technique that involves shooting laser pulses at a wall so that a small fraction of the scattered light bounces around a barrier. Moments after each pulse, they use a “streak camera,” which records individual photons at billions of frames per second, to detect the photons that bounce back from the wall. By measuring the times-of-flight of the returning photons, the researchers can tell how far they traveled and thus reconstruct the detailed 3-D geometry of hidden objects the photons scattered off of behind the barrier. One complication is that you must raster-scan the wall with the laser to form a 3-D image. Say, for instance, there’s a hidden person around the corner. “Then light from a particular point on the head, a particular point on the shoulder, and a particular point on the knee might all arrive [at the camera] at the same exact time,” Raskar said. “But if I shine the laser at a slightly different spot, then the light from the three points will not arrive at the same exact time.” You have to combine all the signals and solve what’s known as the “inverse problem” to reconstruct the hidden 3-D geometry.

Raskar’s original algorithm for solving the inverse problem was computationally demanding, and his apparatus cost half a million dollars. But significant progress has been made on simplifying the math and cutting costs. In March, a [paper published in Nature](#) set a new standard for efficient, cost-effective 3-D imaging of an object — specifically, a bunny figurine — around a corner. The authors, [Matthew O’Toole](#), [David Lindell](#) and [Gordon Wetzstein](#) at Stanford University, devised a powerful new algorithm for solving the inverse problem and used a relatively affordable SPAD camera — a semiconductor device with a lower frame rate than streak cameras. Raskar, who supervised two of

the authors earlier in their careers, called the work “very clever” and “one of my favorite papers.”

Previous algorithms got bogged down by a procedural detail: Researchers usually opted to detect returning photons from a different spot on the wall than the laser pointed to, so that their camera could avoid the laser’s backscattered light. But by pointing their laser and camera at almost the same point, the Stanford researchers could make outgoing and incoming photons map out the same “light cone.” Whenever light scatters off a surface, it forms an expanding sphere of photons, and this sphere traces out a cone as it extends through time. O’Toole (who has since moved from Stanford to Carnegie Mellon University) translated the physics of light cones — developed by Albert Einstein’s teacher, Hermann Minkowski, in the early 20th century — into a concise expression relating photon times-of-flight to the locations of scattering surfaces. He dubbed this translation the “light-cone transform.”

Self-driving cars already have LIDAR systems for direct imaging and could conceivably someday also be equipped with SPADs for seeing around corners. “In the near future these [laser-SPAD] sensors will be available in a format that could be handheld,” predicted [Andreas Velten](#), the first author of Raskar’s seminal 2012 paper, who now runs an active-imaging group at the University of Wisconsin, Madison. The task now is to “go into more complex scenes” and realistic scenarios, Velten said, “rather than having to take great care in setting up a scene with a white object and black space around it. We want a point-and-shoot.”

Where Things Are

Researchers in Freeman’s group have started integrating passive and active approaches. [A paper](#) led by the postdoctoral researcher Christos Thrampoulidis showed that in active imaging with a laser, the presence of a pinspeck camera of known shape around a corner can be used to reconstruct the hidden scene, without requiring photon time-of-flight information at all. “We should be able to do it with a regular CCD camera,” Thrampoulidis said.

Non-line-of-sight imaging could someday aid rescue teams, firefighters and autonomous robots. Velten is collaborating with NASA’s Jet Propulsion Laboratory on a project aimed at remote-imaging the insides of caves on the moon. Meanwhile, Raskar and company have used their approach to [read the first few pages of a closed book](#) and to [see a short distance through fog](#).

Besides audio reconstruction, Freeman’s motion magnification algorithm might come in handy for health and safety devices, or for detecting tiny astronomical motions. The algorithm “is a very good idea,” said [David Hogg](#), an astronomer and data scientist at New York University and the Flatiron Institute (which, like *Quanta*, is funded by the Simons Foundation). “I’m like, we’ve got to use this in astronomy.”

When asked about the privacy concerns raised by the recent discoveries, Freeman was introspective. “That’s an issue that over my career I’ve thought about lots and lots and lots,” he said. A bespectacled camera-tinkerer who has been developing photographs since he was a child, Freeman said that when he started his career, he didn’t want to work on anything with potential military or spying applications. But over time, he came to think that “technology is a tool that can be used in lots of different ways. If you try to avoid anything that could ever have a military use, then you’ll never do anything useful.” He added that even in military situations, “it’s a very rich spectrum of how things can be used. It could help someone avoid being killed by an attacker. In general, knowing where things are is an overall good thing.”

What thrills him, though, is not the technological possibilities, but simply to have found phenomena hidden in plain view. “I think the world is rich with lots of things yet to be discovered,” he said.