NEWS

Vision's Grand Theorist

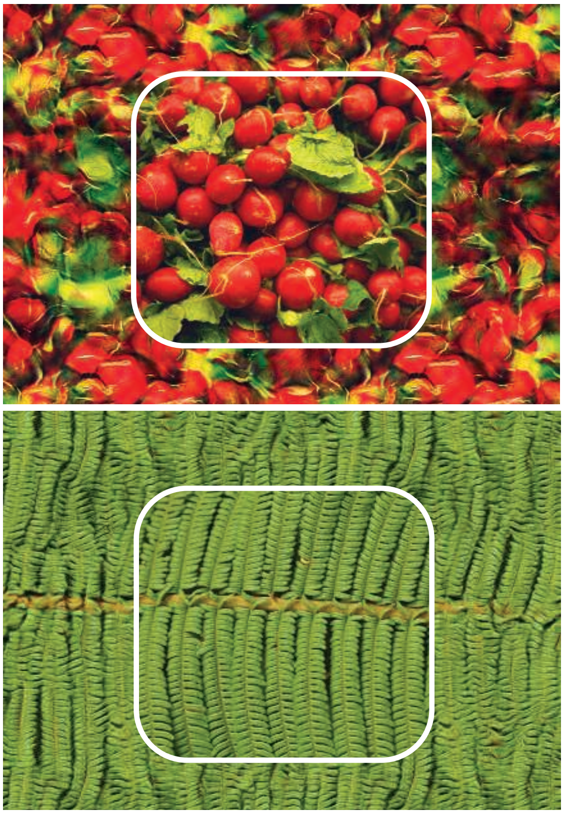
Eero Simoncelli has an eye for mathematical truths that explain human vision—and he's adept at translating that knowledge into practical tools such as image-compression techniques

A great divide traditionally separates theory from experiment in neuroscience. Theorists typically deal in idealized mathematical abstractions far removed from nitty-gritty physiological data. Experimental neuroscientists often view such musings with disdain, considering them irrelevant or too mathematically dense to be of any use.

Eero Simoncelli, a Howard Hughes Medical Institute vision researcher at New York University (NYU), is one of a small but growing cadre of computational neuroscientists bridging this divide. Forty years after researchers revealed the cellular fundamentals of vision, how the electrical signals delivered by the eye's rods and cones assemble into full-scale visual perceptions remains largely an enigma. To sharpen the picture, Simoncelli is working to make neuroscience more like physics, a field in which theory and experiment more easily blend. Just as physicists replaced loose, qualitative descriptions of the physical world with mathematically precise language, Simoncelli aims to devise fundamental equations of vision. "I'm working to encapsulate the conceptual principles used by the brain in precise mathematical terms," he says.

Simoncelli's analyses have already solved several longstanding mysteries in visual science: for example, how the brain assembles a moving picture of the world and why humans drive too quickly in the fog. He's also helped explain how evolution may have sculpted the brain to respond ideally to the visual environment on Earth. On a more practical side, Simoncelli has developed novel methods for image compression and for cleaning up visual noise, such as TV snow. "Eero can hang out with the people who make JPEGs look better or compress info onto DVD," says NYU neuroscientist Anthony Movshon, who collaborates with Simoncelli. "But to make this fit to the biology is a unique skill."

Simoncelli even hopes that his work will lead to insights into consciousness. His peers say that's not arrogance but quiet confidence. "Eero's work is … both powerful and simple," says Matteo Carandini, a neuroscientist at the Smith-Kettlewell Eye Research Institute in San Francisco, California. "His group is the best thing around." Bruno Olshausen, a computational neuroscientist at the University of California (UC), Berkeley, adds that Simoncelli's work "has been very inspirational to lots of people, including me."



Perception problem. By understanding mathematically how the brain perceives texture, Eero Simoncelli has developed software that can synthesize the textures in an image. It works best when the object has a regular pattern.

CREDIT: J. PORTILLA AND E.P. SIMONCELLI, *INT'L JOURNAL OF COMPUTER VISION* **40** 49–71 (2000)

Brain as machine

Simoncelli has wanted to study the brain since childhood. But he could not relate to—or remember—the piles of facts he was asked to learn in his introductory biology course at Harvard University. So he decided to major in physics instead of biology and later got a Ph.D. in electrical engineering while working in Edward "Ted" Adelson's visual science laboratory at the Massachusetts Institute of Technology. In his Ph.D. thesis, Simoncelli mathematically described a network of neurons that processes visual motion. His simulated brain cells performed computations mimicking the responses that neurophysiologists had recorded from cells in their laboratories. "He has brilliant intuitions about images and vision," Adelson says. "Combining engineering principles and biological insights, he's developed models of visual processing that are among the best in the world."

Simoncelli's Ph.D. analysis of visual motion captured a vexing oddity that other researchers had glossed over: the nonlinearity of vision-processing neurons. Engineers favor linear systems because they behave according to a simple law: If two stimuli are combined, the system's response to the combination is equal to the sum of its responses to each separate stimulus. By contrast, nonlinear systems generate more complex responses. "One of the reasons we have so much trouble trying to understand the brain is that it doesn't behave according to the rules of our standard engineering toolbox," Simoncelli says.

In the mid-1990s, as a computer science professor at the University of Pennsylvania, Simoncelli again embraced nonlinearity, producing a novel solution to a classic image-analysis problem: He identified a new set of mathematical regularities in the relations between the pixels that make up photographic scenes. His pixel analysis led to a state-of-the-art technique for compressing images and a method for eradicating visual noise that remains the best in the world as judged by experimental tests. Such a noise-removal technique might eventually be used to make crisper, filmlike image sensors in digital cameras or clear up pictures received from TV satellite dishes.

Bridging the gap

Next, Simoncelli wanted to link his image analysis to the human visual system. He hypothesized that evolution may have forced the brain to encode the visual world in the most efficient, mathematically optimal way. Using that concept, Simoncelli and his colleagues reported in 2001 that the nonlinear responses of neurons, such as those in the primary visual cortex at the back of the brain, are

Published by AAAS

Visual insights. Simoncelli has explained why drivers speed in the fog and how the brain makes sense of moving objects.

well-matched to the statistical properties of the visual environment on Earth, that is, the mathematical patterns of lightness and darkness that recur in visual scenes. The result may help explain how evolution nudged certain visual neurons to be acutely sensitive to object edges and contours, for example.

Last year, Simoncelli and his colleagues reported building an image-compression tool based on his nonlinear model of cortical neurons. Simoncelli reasoned that if the brain's visual cortex is optimally efficient at processing images, it should also do a superior job of compressing them. What's more, any distortions introduced by his compression process should be tolerable. "If the cortical representation is like what's in our brains, we won't notice the difference," he says. Indeed, the new compression technique's performance far outstripped that of the JPEG standard.

Working with postdoc Javier Portilla, Simoncelli has similarly devised a novel mathematical description of how the brain achieves visual texture perception. That's led to a better way of synthesizing pictures—say, an image of a patch of a certain type of grass or cloth—that maintain a material's distinctive appearance. "The model provides a good description of what a person sees when looking at texture," Simoncelli says, adding that he and Portilla have tested it on an extensive number of texture images.

"It does something almost artistic," says UC Berkeley's Olshausen of Simoncelli's texture model. The model, Olshausen adds, not only points vision scientists to the essential properties of texture, but it also could be useful to filmmakers who would like to paint textures onto computer-generated images.

Despite the practical relevance of his work, Simoncelli has largely stayed within the ivory tower. Although he has filed for patents in the past, earning three, Simoncelli hasn't applied for any on his new texture work, or for his most recent noise-reduction and image-compression techniques. One reason, he says, is that patenting delays publication of his ideas. Moreover, applying for a patent on software, versus an actual device, "feels like playing the lottery because the chances are low that it's going to hold up. I don't care enough about money to make it a priority."

In motion

Recently, Simoncelli has helped solve several riddles of motion perception. In the April issue of *Nature Neuroscience*, Simoncelli and his postdoc Alan Stocker explained the Thompson effect, in which motion seems to slow down when the visual landscape lacks contrast. This illusion, first described 25 years ago by psychologist Peter Thompson, helps account for why people drive too quickly in the fog. Simoncelli and Stocker asked five people to judge which of two computer-generated gratings looked like it was moving faster. The researchers varied the gratings' speed and contrast, and each volunteer was asked to make about 6000 separate judgments. Stocker and Simoncelli then analyzed the data using Bayesian statistics, a branch of mathematics that combines expectations with new information, and deduced each person's expectations from his or her speed perceptions. It turns out that people expect slow movement over fast, and that those expectations trump actual perceptions when the perceptual data are sketchy, as occurs in low-contrast situations (*Science*NOW, 21 March, sciencenow.sciencemag.org/cgi/content/full/2006/321/2)

Another 25-year-old motion mystery is also about to succumb to Simoncelli. Scientists have long known that cells in the primary visual cortex process pieces of a visual scene and that those pieces are then assembled into a greater whole by cells in other brain areas. But when an object is moving, it was not at all clear how a brain put the pieces together. Ever since his Ph.D. thesis, Simoncelli has worked on the calculations a computer should perform to mimic a system that can combine pieces of a moving image and spit out a coherent response. Again, he used Bayesian mathematics to try to make sense of people's perceptions of motion and the physiological data from visual neurons. He then mapped all of his computations onto a simulation of neuronal responses that starts in the retina and ends in the visual motion-processing region known as area MT.

In a paper to appear in *Nature Neuroscience* this fall, Simoncelli and Movshon along with postdocs Nicole Rust and Valerio Mante offer the first precise mathematical description of how cells in MT translate pieces of a moving scene into the movement of the whole. They vetted their model against new recordings from individual MT neurons in monkeys exposed to a specific set of stimuli: wiggling lines that look like the ripples on the surface of water. From the model, the researchers could extract biological information about MT cells, including which visual cortex cells feed into them. MT neurons are "profoundly nonlinear," Simoncelli says. "The model explains how that profound nonlinearity can arise from a cascade of very simple nonlinear steps."

Movshon, who did the experimental work buttressing the new model, describes Simoncelli's solution as "simple and elegant," and says the work also gives the field more sophisticated techniques for analyzing and extracting information from recordings of neuronal responses. Moreover, Simoncelli and his colleagues are putting the finishing touches on a set of algorithms that should help neuroscientists better interpret the flood of information that comes from recording large groups of neurons simultaneously in the retina, instead of one at a time as is traditionally done.

Ultimately, Simoncelli aims to put many of his individual findings, and those of his collaborators, into nothing less than a grand unified theory of visual motion perception. "In 10 years, I think we will have a clean computational model of motion," he predicts.

And if that wasn't ambitious enough, Simoncelli is digging for deeper truths. "As we build better descriptions of the brain and test them experimentally, we hope to arrive at fundamental principles that can explain all brain activity, from sensation to consciousness," he says. "That's going to help us understand who we are." Now that's a grand vision.

–INGRID WICKELGREN

CREDIT: NICOLAS BONNIER