The Mind's Eye

Los Alamos researchers are at the forefront of a revolution in experimental brain science

Consider the human brain as a computer. It is an electrical signaling system capable of carrying out mathematical and logical operations. It has short-term and long-term memory. It exchanges inputs and outputs with external devices, like ears and arms. Estimates of human brain performance vary widely because no direct method of comparison to a computer is known, but based on the brain's hardware and architecture, some experts peg its computing power roughly on par with the world's fastest supercomputer.

Yet it is clear that the brain is not like a computer. For one thing, humans are notoriously bad at arithmetic. Even humans who excel at arithmetic are bad at arithmetic when compared to even the most limited calculators and computers. But while humans lose every "mathletic" contest hands-down, they utterly obliterate the electronic competition when it comes to more sophisticated tasks, such as recognizing other people—even when seen from different angles or illuminated by different light sources—and reading their emotional states from the subtle variations of their facial muscles. And while a supercomputer might store more bytes or achieve more operations per second, it takes up an entire room and consumes enough electricity to power thousands of homes. The human brain, on the other hand, fits neatly between the ears and runs on chicken and broccoli.

Researchers at Los Alamos National Laboratory and elsewhere have been striving to program a computer to accomplish complex tasks as well as a human being, with only limited success. One reason for the difficulty stems from the significant technical differences between a brain and a computer. For human beings, there is no rigid distinction between processors and memory chips (the same neurons are both), nor is there even a simple distinction between hardware and software. In addition, the basic processing unit of the brain, the synapse, is substantially more complex than a computer chip's transistor. The brain does amazing things, but it's not yet clear how its organization contributes to its success.

A real brainstorm: In order to examine a visual scene, the brain parses signals from the eyes into components, such as edges, shapes, colors, sizes, locations, recognized objects, and motion. It accomplishes this analysis, even for complex scenes like those shown on the following page, with far greater speed and accuracy than a computer. How exactly the brain does this is not yet known, but Los Alamos scientists are pioneering the experimental techniques that may provide the answers.

TWO BLUE COLLARS BY DICK MASON REPRINTED WITH PERMISSION FROM THE WINDSOR BETTS GALLERY

Teaching a computer to make sense of complex visual scenes like these may depend on first learning how humans do it.

One such amazing thing is the brain's ability to understand what the eyes see. In a fraction of a second, a person can recognize any of the tens of thousands of objects he or she frequently sees in the world, regardless of how each object appears in the scene. It could be a sycamore tree in the fog or a particular type of pen in a messy supply closet. "Humans doing object recognition are essentially flawless," says cognitive psychologist Amy Guthormsen, part of a team of researchers at Los Alamos led by John George and Garrett Kenyon trying to reveal how the human visual system works. The team's ingeniously programmed, state-of-the-art computer model for human vision, she says, "scores a B+ at best."

Taking In the View

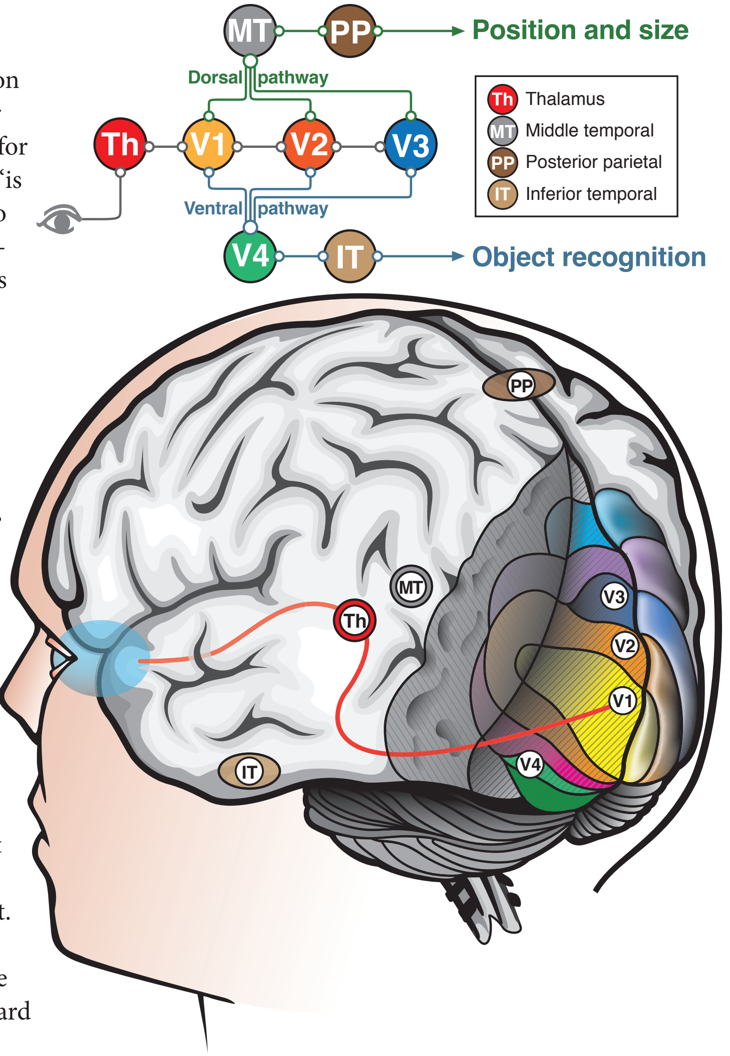
What little is currently understood about human vision goes like this: The rods and cones in the retina respond to the intensities and colors of light entering the eye from each direction within the field of view. This information is then transmitted, like a video stream, to the thalamus near the center of the brain. The retina and the thalamus (in that order) each perform some minimal processing of the visual data; for example, the retina identifies regions of contrast and relative amounts of color within the scene. After acquiring such preliminary information, the data stream is relayed to the visual cortex at the rear of the brain.

The visual cortex is organized into several component regions believed to process visual data in a hierarchical fashion—with more complex information extracted at higher levels of the visual cortex (see upper figure on page 5). The lowest level, called V1, seems to extract some basic information about edges, orientations, and motion. Up a level to V2, you get some other edge feature and color information. At V3 and V5 you get more insight about motion, and at V4 you get simple shape recognition. It is clear that the human visual system shows compartmentalization: different regions do somewhat different tasks, generally in order of increasing complexity.

Interestingly, what seems to be missing entirely from the brain's image processing system is an actual image. "There's no JPEG file in the brain," explains Michael Ham, a Los Alamos physicist who studies computer vision. "It's not as though the brain forms an image and shows it to some kind of mental processor for analysis; the brain extracts different pieces of information from the visual data without ever assembling it into an image." Indeed, this may be why it has been so difficult to design an artificial (robotic) visual system to mimic the human one: With computer image processing, you start with a still image and try to identify its components. But with human visual processing, you begin dissecting and reorganizing the data stream before it even leaves the eye.

Zhengping Ji, of the Laboratory's applied mathematics group, also works on computer modeling of human vision. Using results from human-subject experiments run by Guthormsen and others, Ji structures a computer model to process information in the complicated manner seemingly employed by the human visual cortex. This has allowed the model to outperform earlier models, but it still can't compete with an actual human being. The problem is partly a lack of knowledge about exactly how (and how often) the various regions like V1 and V2 communicate with one another and within themselves. For example, some activity in V1 has been observed to occur after other activity in V2, implying feedback. But it's not clear what exactly is accomplished by this feedback or how each region contributes.

Another poorly understood aspect of human vision is the higher-level processing that takes place further up the hierarchy of the visual cortex. Starting after V1 at the back of the brain, visual signal processing splits along two main pathways through the brain. The dorsal pathway runs forward along the top of the brain's surface (broadly called the cortex), while the ventral pathway runs forward along the bottom of the cortex. The upper pathway terminates at the posterior parietal (PP) lobe, while the lower pathway terminates at the inferior temporal (IT) lobe. Each pathway appears to serve a different purpose. The PP seems to determine the position and size of objects in the field of view. The IT, on the other hand, identifies the objects themselves. The PP might say, for example, "On top of the lamppost," and the IT would say, "is a black crow." (PP: "A big one, too.") How the PP and IT do their jobs is not yet known, nor is it known how they communicate with the lower levels of the visual cortex. But Los Alamos scientists might know how to find out.



Once a visual signal leaves the eyes and is relayed by the thalamus to the visual cortex, its progression divides into two major pathways. Along the dorsal (upper) pathway, successively higher-level processing leads to the posterior parietal (PP) lobe, from which emerges detailed size and position information about objects in the visual scene. Along the ventral (lower) pathway, the inferior temporal (IT) lobe ultimately identifies what the objects in the scene are. A variety of intermediate steps also help dissect the scene—its edges, colors, and motion, for example.

EEG, MEG, MRI, Oh My!

The most tried-and-true method for observing the brain in action is functional magnetic resonance imaging (MRI). This type of brain scan uses strong magnetic fields to probe for oxygenated hemoglobin. The logic goes, when part of the brain is in use, it requires more oxygenated blood, causing the MRI scan to take notice. The scan can resolve the location of the extra oxygenated blood in the brain within a few millimeters in any direction.

But while the MRI's spatial resolution ("where") is excellent, its temporal resolution ("when") is poor. Because it takes 2–5 seconds for the body to supply the extra oxygenated blood once some part of the brain has "requested" it, the MRI can only tell what parts of the brain were in use 2–5 seconds ago, and therefore it is most useful for studying brain tasks that last at least that long. Visual processing, however, happens much faster than that. Complex object recognition takes less than half a second, and more "primitive" tasks, such as triggering the response to duck when something not-yet-identified is coming toward your head, are virtually instantaneous.

Fortunately, there are other types of brain scans with better temporal resolution than the MRI: electroencephalography (EEG) and magnetoencephalography (MEG). When a current flows in a circuit—or charged ions flow in a neuron—an electromagnetic signal is produced, and EEG and MEG scans pick up different parts of that signal. The strength of the signal depends on the angle between the detector and the actual "wire" (neurons are long and thin, like wires). Due to the geometry of the cortex, neurons on the smooth part of the brain's surface show up better with EEG, while the neurons located within the brain's folds show up better with MEG.

Both EEG and MEG have the advantage of capturing actual electrical activity in the brain, rather than using oxygenated blood as a proxy for it. Both employ an arrangement of sensors on the head, and both have excellent temporal resolution, allowing them to determine the timing of field changes down to about a millisecond. A drawback of EEG is

Los Alamos's computer model for visual object recognition has as much processing power as the human visual cortex and knows how to recognize certain objects. But the model underperforms a human being in both speed and accuracy. Boxes in this image indicate where the model successfully identified the features of a vehicle.

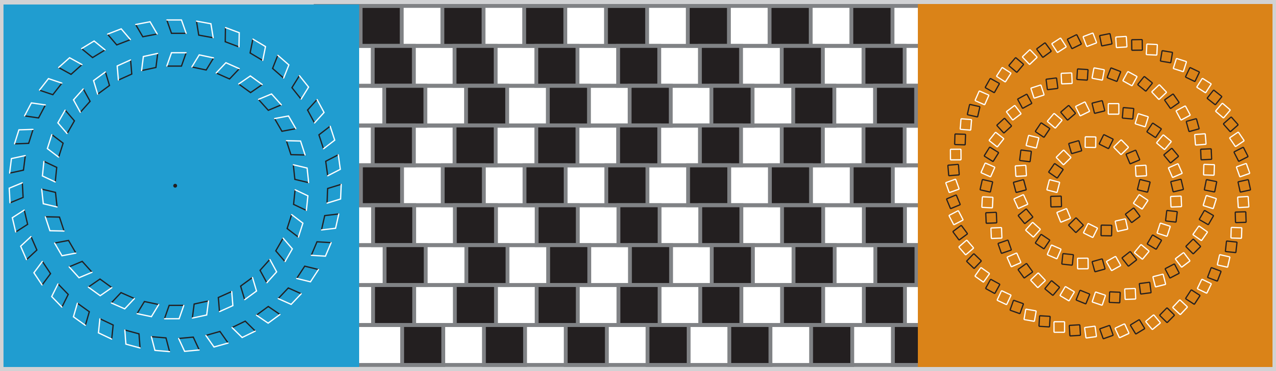
that materials in the human head block electric fields to varying degrees, and any effort to compensate requires an approximation. The same is not true for magnetic fields, making MEG signals less ambiguous, but the equipment necessary for performing MEG is sufficiently specialized and expensive that its use is restricted to larger laboratories, including Los Alamos.

Although both EEG and MEG have sufficient temporal resolution to study what happens when, they both suffer from the same serious flaw: they can't tell where the signals originate. The same EEG or MEG signal can be produced at a particular sensor location by many different combinations of neurons firing all over the brain, making it impossible to uniquely identify the region or regions of the brain responsible for the combined signal. What's needed is a way to obtain the spatial resolution of the MRI with the temporal resolution of an EEG or MEG (or both).

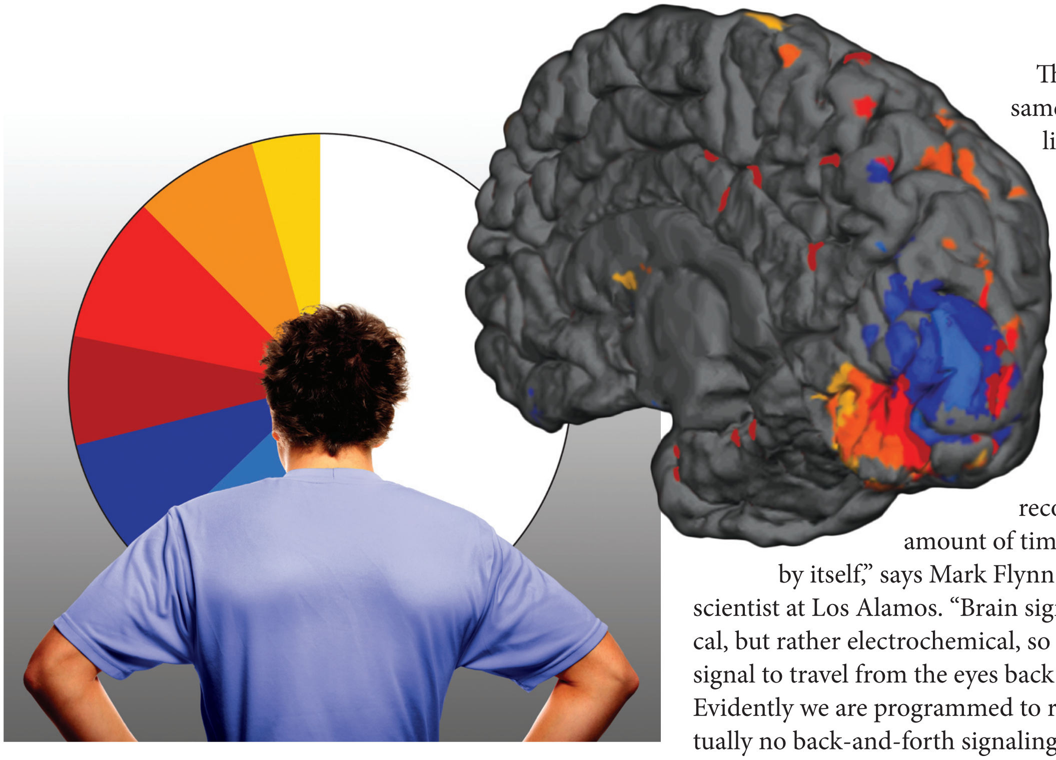
The brute force approach would be to measure both at the same time. Unfortunately, the MRI uses such powerful magnetic fields that it completely drowns out the subtle signals measured by EEG or MEG. Nonetheless, a team led by Michelle Espy of the Lab's Applied Modern Physics group is close to taking simultaneous readings from an MEG during an MRI scan. If Espy succeeds, it should be possible to sample what the different parts of the brain are doing in real time, in the visual cortex or anywhere else. Then it will be necessary to develop a technique for blending the two different types of data into a coherent picture of how signals are shuffled around in the brain.

"We don't have to wait for simultaneous location and timing data," Guthormsen points out. "As long as we know in detail what parts of the brain carry out some particular mental operation, we can blend MEG and MRI data today." That's where the choice to study vision comes in. The retina sends specific parts of a visual scene to specific parts of the visual cortex. Thus, if Guthormsen shows a test subject a series of lights appearing in different parts of the subject's field of view during an MRI examination (allowing 2–5 seconds each time), the MRI will determine where in the brain those lights are processed. This allows the construction a retinotopic map, which shows the parts of the field of view that are directed from the retina to specific parts of the V1 cortex (see figure at right). And while one can't work backward from an EEG or MEG signal to locate the neurons involved in producing that signal, one can use the combination of neurons identified in a retinotopic map to project the EEG or MEG signal that should result. Guthormsen needs only show her test subjects visual stimuli in particular parts of the visual field.

This methodology solves a core physics problem associated with localizing MEG or EEG data: that it is not possible to isolate the unique set of firing neurons responsible for generating a measured MEG or EEG signal. However, it is possible to attribute that signal, arising in a combination of sensors around the head, to the visual processing that produces it if the neurons involved have already been identified—in this case, by a retinotopic map.



The human visual system builds a model of the world—a best guess as to what real-world scene could have given rise to an observed pattern of data. This is an interpretive rather than algorithmic process. Each of these optical illusions plays upon this distinction by inducing the brain to falsely identify visual elements, such as those associated with 3-D perspective or motion. Left: Watch the central dot as you move your head toward and away from the page. You will see the circular patterns appear to spin. Middle: The horizontal lines appear to alternately converge and diverge, but they are in fact parallel. Right: What appears to be a set of spirals is really just a set of concentric circles.



This retinotopic map shows where in the visual cortex at the rear of the brain different parts of a test subject's field of view are processed. The pinwheel pattern shows angular position on the left side of a subject's view. (Radial position, which would appear as concentric circles, is also mapped but is not shown here.) From both eyes, signals pertaining to this left-side view are relayed to the right hemisphere of the brain, shown here. For example, the dark red color indicates that objects seen at eye-level on the subject's left side are being analyzed, in part, at the center of the V1 area of the visual cortex on the right hemisphere.

The Truth about Cats and Dogs

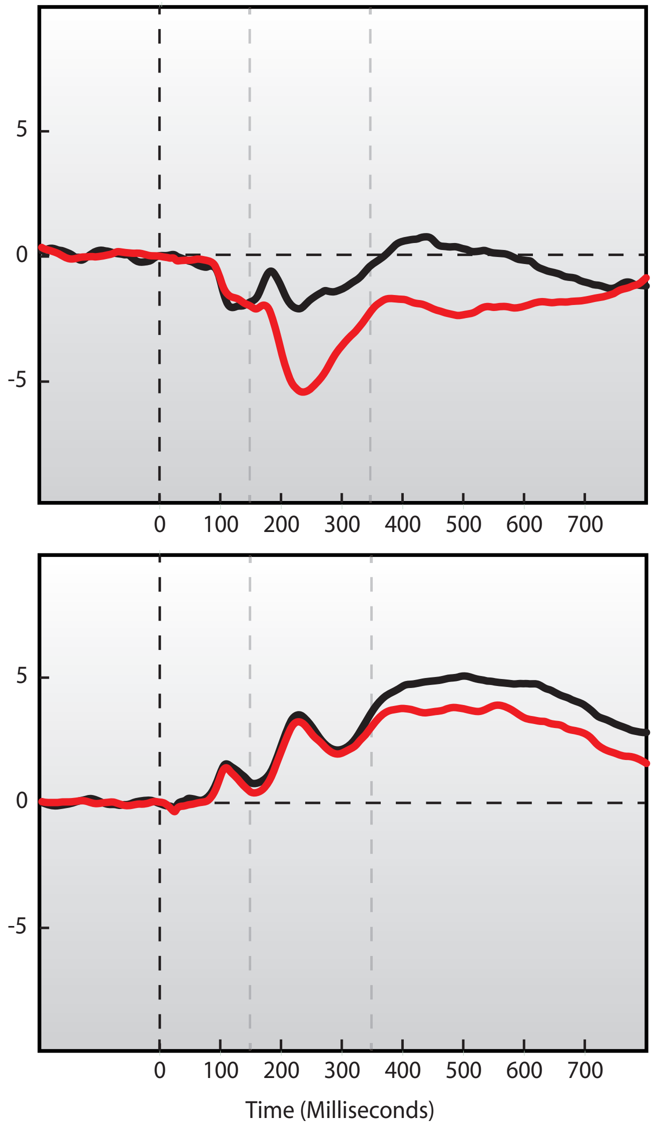
Los Alamos scientists are working to pioneer the merging of MRI (retinotopic map) and electromagnetic (EEG or MEG) data from the visual cortex. If successful, they hope to identify how the brain coordinates information between various parts of the visual cortex to comprehend a scene. But they have already glimpsed the kind of surprising results such an approach can provide in a related experiment.

The team showed a test subject a series of photographs; for each, the subject was instructed to indicate quickly whether he or she saw a "target" animal in the photograph. In some experiments the target was a cat and in others it was a dog, but the photographs included pictures of cats, dogs, and various inanimate objects. An EEG apparatus recorded electrical activity over time at various locations, allowing the construction of waveforms (see figure on page 8) that can be compared to identify at what point the brain's electrical activity begins to differ when processing different images.

The waveforms look the same for the first 150 or so milliseconds (ms); at that point, a waveform associated with observing an animal begins to diverge from one associated with observing an inanimate object. While it's not clear what exactly the brain is doing differently in the two cases, it is clear that some degree of recognition must occur in that amount of time. "That's quite amazing all by itself," says Mark Flynn, a biologist and computer scientist at Los Alamos. "Brain signals are not purely electrical, but rather electrochemical, so it takes about 150 ms for a signal to travel from the eyes back through the visual cortex. Evidently we are programmed to recognize animals with virtually no back-and-forth signaling within the cortex because 150 ms just isn't enough time for it."

On the other hand, the waveform for a cat image doesn't begin to differ from that for a dog image until about 350 ms have elapsed. This suggests that there may be substantial crosstalk needed within the visual cortex to recognize the difference between similar objects (animals in this case). The researchers believe that these kinds of results can help them discriminate between competing theories of how the brain understands vision. Prevailing theories, for instance, have held that the visual cortex processes signals upward along the hierarchy only—from V1 to V2 and straight up each pathway to the PP and IT. That may be adequate, and perhaps necessary, when comparing puppies with inanimate objects. But for more sophisticated comparisons, the extra 200 ms may imply the need for signaling back and forth across and within levels. For example, to distinguish a cat from a dog, the IT may (somehow) request more detail from V2, say, to see if there are any whiskers and from V4, perhaps, to determine the shape of the eyes. The IT may then coordinate the results. For example, finding both whiskers and vertical-sliver eyes, the IT concludes the animal is a cat.

"At this point, we can only speculate about how the visual cortex actually functions," Guthormsen admits. In order to rigorously demonstrate how the cortex communicates across levels and within them, the team needs to combine temporally sensitive waveforms with spatially sensitive MRI data. When MEG sensors obtain signals consistent with neuron activity in the regions prescribed by retinotopic

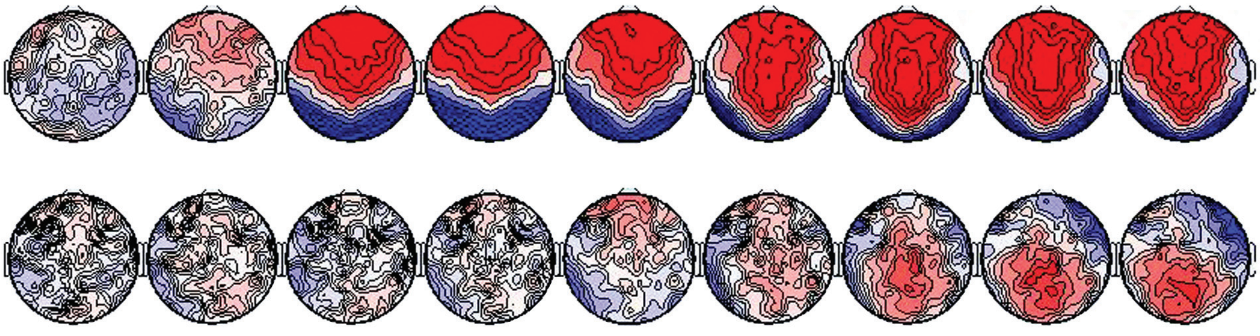


These EEG waveforms show differences in brain activity when people recognize different objects. In each case, the test subject is instructed to distinguish a specified target object (e.g., a dog) from a nontarget object. In the upper frame, the black line resulted from showing a test subject a picture of the target (a dog), and the red line resulted from showing an inanimate object. The two waveforms track one another for about 150 milliseconds—approximately the time needed for a brain signal to travel directly from the eyes into the visual cortex—at which point some recognition that the target and nontarget differ evidently kicks in. In the lower frame, the red line shows the waveform obtained when the test subject is shown a cat as the nontarget object. Due to the similarities between dogs and cats, it takes longer for the brain to recognize the difference and the waveforms track together for about 350 milliseconds, allowing enough time for different parts of the visual cortex to signal back and forth to "compare notes."

mapping, they will have the data they need. "Of course," she says, "obtaining the data is one thing; making sense of it will take some time. But at least now we've shown it can be done."

Doing What Comes Naturally

Presumably, the complicated manner in which the human brain processes visual information is an evolutionary optimization. Some other animals, for example, are particularly good at identifying the shape and motion patterns of their predators and prey; their brains may have organized to maximize these abilities. Because human evolution followed a particular path—walking upright and using arms and hands to manipulate objects, socializing for cooperative benefit and protection, choosing mates based on various visual clues to their genetic quality—it stands to reason that human brains are organized to succeed at these tasks. The visual cortex needs to understand objects and people seen from different angles and in different contexts. Evolution, therefore, needed



These "scalp plots" show a top view of a test subject's head and indicate the variation in EEG data at different locations (ears and noses are shown for orientation). Left to right, the top row shows the progression of EEG signal differences every 50 milliseconds after showing an image of a target animal versus an inanimate object. Red and blue both indicate brain locations where there are different EEG results for the two cases, while white indicates identical EEG data. The bottom row is the same for target animals versus nontarget animals (dogs vs. cats). Greater spatial detail will be needed to constrain theories of what the brain is actually doing in each case.

Left to right: Michael Ham, Amy Guthormsen, and Mark Flynn pose to demonstrate just how capable the visual cortex is.

to find every possible trick to enhance these abilities because it couldn't house (or supply adequate power to) an entire supercomputer inside a primate's head. The result is the clever, capable, and very complicated brain.

So far, scientists and engineers have been unable to construct an artificial intelligence to match the capability of the brain, so they study the brain in the hope of duplicating its methods. One advantage evolution had over today's researchers, however, is time. Humans spend years in early childhood accumulating information and learning how to understand what their eyes see, while artificial systems are generally expected to function right out of the gate. It may be more fruitful to invent a robot-computer-camera system that can acquire visual sense over years of experience, just as human children do. In the meantime, the Los Alamos team and others in the field think it's wise to try to understand the brain as designed by nature.

If they succeed, the results could be world-changing, allowing robotic systems to attain human-level object recognition capability. This could allow automation of many tasks currently carried out by human labor, and it could lead to new technologies for assisting people with vision disabilities—perhaps eventually including computer and camera elements that link to the brain. And if Los Alamos succeeds in blending MRI and MEG systems to obtain data simultaneously, the benefits need not be limited to artificial vision. Without the need for a retinotopic map to provide the spatial detail, brain researchers could uncover the tricks behind human processing of language, emotion, humor, and so on.

But far from trying to speed the world along toward a cyber-science-fiction future, Guthormsen and her colleagues obtain their daily thrill in the pure science of studying how the brain works. "We find challenge and reward enough," she says, "just trying to uncover how people do the incredibly difficult things they do everyday without appearing to make the slightest effort." ❖ LDRD

—Craig Tyler