The Movies in Our Eyes

The retina processes information much more
than anyone has ever imagined,
sending a dozen different movies to the brain

By Frank Werblin and Botond Roska

We take our astonishing visual capabilities so much for granted that few of us
ever stop to consider how we actually see. For decades, scientists have likened
our visual-processing machinery to a television camera: the eye's lens focuses
incoming light onto an array of photoreceptors in the retina. These light detec-
tors magically convert those photons into electrical signals that are sent along
the optic nerve to the brain for processing. But recent experiments by the two
of us and others indicate that this analogy is inadequate. The retina actually
performs a significant amount of preprocessing right inside the eye and then
sends a series of partial representations to the brain for interpretation.

We came to this surprising conclusion
after investigating the retinas of rabbits,
which are remarkably similar to those in
humans. (Our work with salaman-
ders has led to similar results.)
The retina, it appears, is a
tiny crescent of brain matter
that has been brought out to
the periphery to gain more
direct access to the world.
How does the retina construct
the representations it sends? What
do they "look" like when they reach the
brain's visual centers? How do they con-
vey the vast richness of the real world? Do
they impart meaning, helping the brain to
analyze a scene? These are just some of
the compelling questions the work has be-
gun to answer.

Overall, we have found that specialized
nerve cells, or neurons, deep within the ret-
ina project what can be thought of as a doz-
en movie tracks—distinct abstractions of
the visual world. Each track embodies a
primitive representation of one aspect of
the scene that the retina continuously up-
dates and streams to the brain. One track,
for example, transmits a line-drawing-like
image that details only the edges of ob-
jects. Another responds to motion, often in
a specific direction. Some tracks carry in-
formation about shadows or highlights. The
representations of still other tracks are dif-
ficult to categorize.

Each track is transmitted by its own
population of fibers within the optic nerve
to higher visual centers in the brain, where
even more sophisticated processing takes

place. (The human auditory system has a similar architecture: each auditory nerve carries information about a very limited range of pitches, and the brain combines them.) Investigators studying the visual cortex have shown that features such as motion, color, depth and form are processed in various regions and that a lesion to a given region can cause a deficit in sensing one specific feature. But the brain's ability to even sense such features in the first place originates in the retinal movies.

The diagrams on the following pages convey our best explanations for how the retina creates the surreal electrical images that inform the brain. As we continue our research, we are beginning to shed some light on how each of the movies is constructed, but by no means are we ready to offer a full model. The 12 movies carry all the information the brain will ever receive to interpret the visual world, but we cannot yet say how their patterns are integrated. It could be that the movies serve simply as elementary clues, a kind of scaffolding upon which the brain imposes constructs. This notion is not dissimilar to the well-described "mind's eye" that knits the words of a novel into a meaningful narrative.

Although the retina's representations appear to fully capture the visual facts of a scene, such as a dinner table, waterfall or talking face, essential components also seem to be missing. Nothing about the feeling, attitude, texture or focus of the scene appears to be present. Perhaps these traits are somehow inherent in the movie tracks the brain interprets. Or perhaps, by using rabbit retinas, we may have failed to find all the representations that would be captured by a human retina— "high resolution" ones that might extract qualities such as feeling in ways we have yet to uncover.

Nevertheless, it is clear that the retina's representations form a natural visual language. Understanding that language has special significance today. Groups around the world are attempting to restore vision to the blind by introducing an artificial sensor right in front of the optic nerve that would take over for the retina. The work has advanced, but the results remain relatively crude, with transmissions limited to vague versions of basic patterns. Human trials have begun at the University of Southern California's Doheny Eye Institute and are about to get under way at Wayne State University Medical School. The final goal of these trials is probably far off, but their success ultimately lies in providing the brain with patterns of activity similar to those that are normally supplied by the retina, incorporating the natural language of vision. The subsequent challenge will be to discover how to "hook up" each abstraction to the appropriate fibers in the optic nerve.

A detailed understanding of the natural language of vision formed within the retina is needed for effective prosthetic devices. At the same time, this understanding will help investigators learn much more about how the eye and brain together see clearly, are deceived by optical illusions, track fast-moving objects and fill in the missing pieces inherent in any rendering on a television, computer or drive-in movie screen. We hope our description of the retina's preprocessing power is a step toward that end. ●

Active Anatomy

The retina's surprising behavior arises from its complex architecture. Painstaking experiments by many specialists have added physiological detail to the classic model of retinal circuitry first delineated by the great Spanish anatomist Santiago Ramón y Cajal a century ago and repeated in textbooks ever since. The transparent retina ❶ consists of a beautifully organized layering of neurons ❷. The outer layer, farthest from the lens, contains the rod and cone cells, which absorb the incoming light and convert it to neuronal activity. These photoreceptors connect to 10 different kinds of neurons known as bipolar cells, which send long signal-carrying arms, or axons, into a central "inner plexiform" layer. This band looks like a series of 10 distinct parallel strata. The axon of each bipolar cell type delivers signals to just a few of the strata.

At the innermost side of the plexiform layer ❸ are 12 different types of ganglion cells (*purple*). Most types send fingers called dendrites into one distinct stratum, where they receive excitatory input from a limited number of the bipolar neurons (*green*). The ganglion cells output the movie streams that the optic nerve carries to different brain regions for interpretation. Some ganglion dendrites branch out widely, carrying diffuse information, whereas others branch more narrowly, carrying high-resolution information. Some respond to an increasing change in the rate at which bipolar cells release neurotransmitters (messenger molecules), some to a decreasing change in that rate.

The inputs sent by the bipolar cells to the ganglion output cells within each of the strata are not enough to create the dozen movie representations, however. The signals emitted by bipolar cells are modulated by a variety of small neurons called amacrine cells (*gray*). Some of these cells operate laterally within a stratum, inhibiting communication between distant ganglion cells in that stratum. Other amacrine neurons inhibit signals vertically between strata—and therefore between different movies—as if to in-

Overview/*Surreal Vision*

- The retina does much more than pass simple signals to the brain. Surprisingly, it extracts a dozen distinct representations of a visual scene—sophisticated, ghostlike movies formed by relatively few types of neurons.
- The brain uses these abstractions to build a visual world sharp with detail and rich in meaning.
- Understanding the "visual language" that these movies carry will aid researchers who are building artificial sensors that could help the blind see. Such insights should also bolster efforts to pin down how the eye and brain see clearly, as well as how they can be deceived.
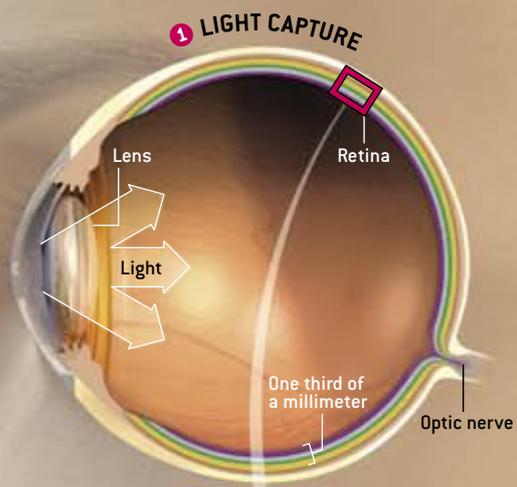
struct one stratum not to record what another stratum is recording. In this way, the amacrine cells pick up and emit signals to coordinate the movie tracks. Researchers such as Heinz Wassle of the Max Planck Institute for Brain Research in Frankfurt, Thomas Euler of the Max Planck Institute for Medical Research in Heidelberg and Richard Masland of Massachusetts General Hospital have identified at least 27 different amacrine cell types (as well as the 10 bipolar types and 12 ganglion types).

Everything we see in space is observed as time advances. Even the recording of a motionless black dot fixed in colorless three-dimensional space constitutes a movie, because the retina sees it continuously as time advances. Many cells of each ganglion type populate the retina, and the set of each type conveys a distinct movie. But unlike a box office film, which is generated frame by frame, the ganglion movies are continuous streams of signals.
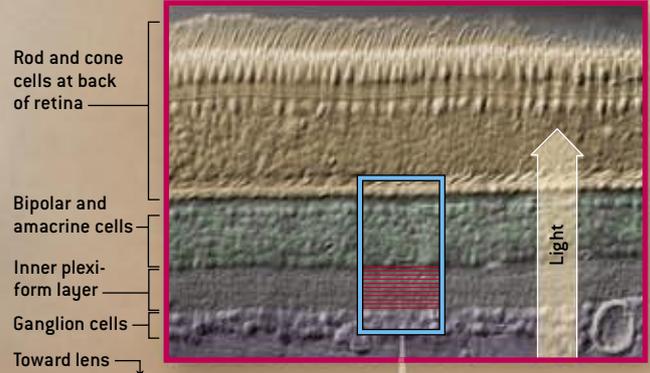
The interactions among bipolar and amacrine cells that are read out simultaneously by each set of ganglion cells make up the data we receive to interpret the visual world. As we read, grasp objects, recognize faces and walk about, various combinations of these movies are the only visual clues the brain receives. They form a fundamental "visual language," with its own phrasing and grammar that embodies the neural vocabulary of vision. ●

THE AUTHORS
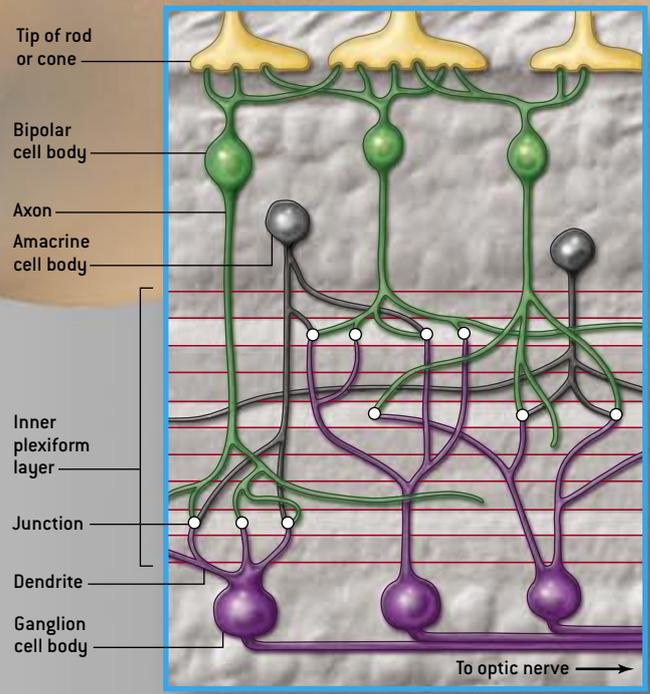
FRANK WERBLIN and BOTOND ROSKA together uncovered much about the retina's functional circuitry in the early 1990s at the University of California, Berkeley. Werblin continues there as a professor of neuroscience. In 1973 he published an article in *Scientific American* after discovering unique physiological characteristics of retinal neurons with John Dowling of Johns Hopkins University. Roska is a group leader at the Friedrich Miescher Institute for Biomedical Research in Basel, Switzerland, where he is developing genetic techniques for identifying visual pathways.

❶ LIGHT CAPTURE

Lens

Retina

Light

One third of a millimeter

Optic nerve

❷ RETINA LAYERS

Rod and cone cells at back of retina

Bipolar and amacrine cells

Inner plexiform layer

Ganglion cells

Toward lens

Light

❸ CONNECTIONS WITHIN PLEXIFORM LAYER

Tip of rod or cone

Bipolar cell body

Axon

Amacrine cell body

Inner plexiform layer

Junction

Dendrite

Ganglion cell body

To optic nerve →

Movies in a Flash

Our descriptions of the retina's complex activity are based on our own experiments. We record what is happening in individual ganglion cells with a tiny, hollow glass needle. This micropipette injects yellow dye that rapidly spreads through all the dendrites of a single ganglion cell, showing us the strata they reach. The pipette also functions as an electrode, measuring the electrical activity of the cell, which reflects the combination of excitatory signals from bipolar cells and inhibitory signals from amacrine cells.

To gain a feel for the movies that ganglion cells stream to the optic nerve, we started fairly simply: by first recording how a linear array of ganglion cells represented a square flash of light shone directly onto the retina of a rabbit ❶. The flash lasted one second and was confined to a square measuring 600 microns on each side. Thus, the flash fell on a small, well-defined region of the retina for a specific length of time.

We recorded the excitation and inhibition signals received by one type of ganglion cell over this period, repeating the procedure for each of the dozen cell types. Each type had a unique response, and the range of responses was remarkably diverse. In the plot below ❷, one box represents one second, and color indicates the magnitude of the signal current in one cell type.

Interestingly, for the ganglion cell type illustrated here, cells across the width of the flash responded, but they were not active for the whole time the light was shining. And oddly, some of them outside the 600-micron span became active after the flash had ended—behavior that appears on the plot as two lobes (*blue*) that arise after the one-second interval. A third area, within the flash region, also activates slightly, near the two-second mark.

How are we to interpret this pattern? If all the cells were sending outputs for the full second, the pattern would be "lit" across the entire span for the entire second, filling the corresponding square on our grid ❸. In reality, the output is filtered; it is as wide as the flash but is truncated in time, lasting perhaps one tenth of a second and starting about one tenth of a second after the flash began. Not only was there a slight delay before the ganglion cells responded, but they apparently responded only long enough to note how incoming light had changed—from dark to bright. Perhaps this ganglion type represents the onset of illumination but not its sustained presence. The slight activation of the cells represented in the two outlying lobes might convey some kind of "off" signals. The third blue spot at two seconds is a signal component we do not yet understand.

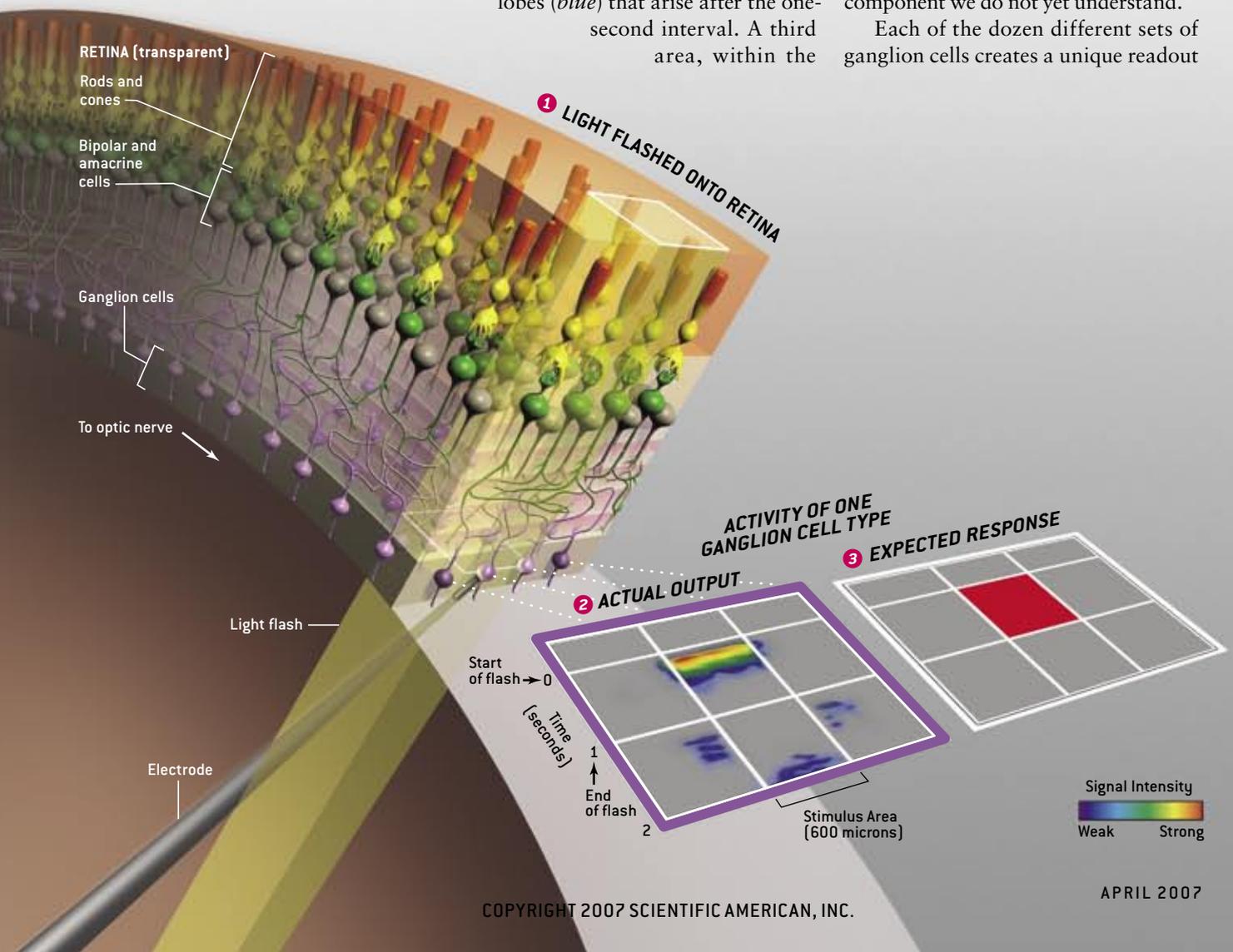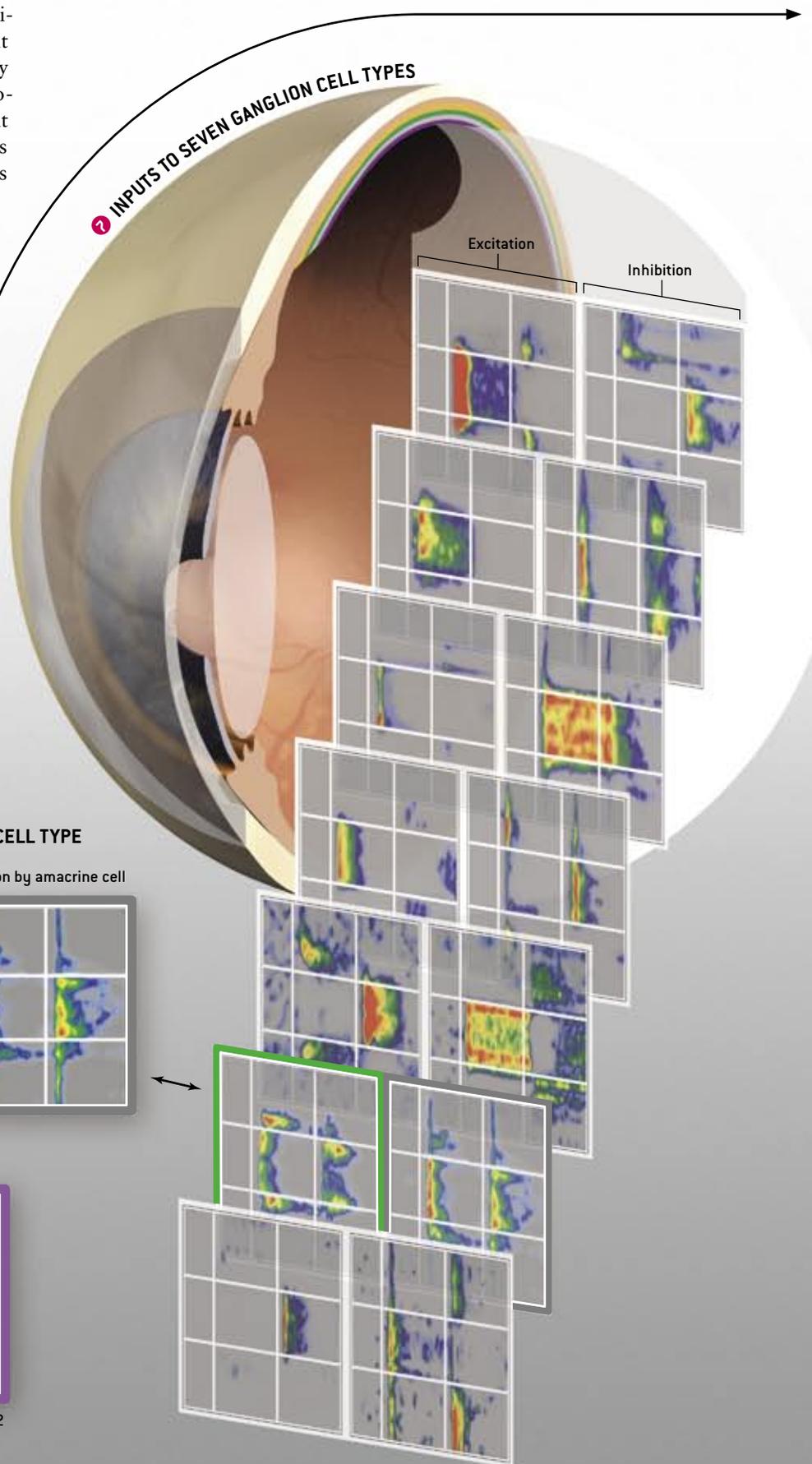Each of the dozen different sets of ganglion cells creates a unique readout

RETINA (transparent)
Rods and cones
Bipolar and amacrine cells
Ganglion cells
To optic nerve
Light flash
Electrode

❶ LIGHT FLASHED ONTO RETINA

ACTIVITY OF ONE GANGLION CELL TYPE

❷ ACTUAL OUTPUT

❸ EXPECTED RESPONSE

Start of flash → 0
Time (seconds)
1
End of flash
2
Stimulus Area (600 microns)

Signal Intensity
Weak   Strong

that accentuates some aspect of the visual world. But recall that this output results from the excitation produced by bipolar cells and the inhibition produced by amacrine cells. The net result is a pared-down final pattern. The plots below ❹, ❺, ❻ show the two inputs and final output for a ganglion cell type different from the one illustrated earlier.

In this way, each ganglion cell type sends a final spacetime representation along the optic nerve to the brain. Each representation is a unique product that arises from a pair of excitation and inhibition patterns ❼. The 12 ganglion cell types continually send 12 of these movie streams to the brain as time advances. (We recorded only seven to make the experiment manageable.) An incredible diversity of activity occurs in response to a simple flashed square.

❼ INPUTS TO SEVEN GANGLION CELL TYPES

Excitation

Inhibition

ACTIVITY OF A SECOND GANGLION CELL TYPE

❹ Excitation by bipolar cell

❺ Inhibition by amacrine cell

❻ Final output to brain

Stimulus Area
(600 microns)

Time (seconds)

0 1 2

1 GANGLION CELL REPRESENTATIONS OF A TALKING FACE

Movie from cell type 1

Cell type 2

3

4
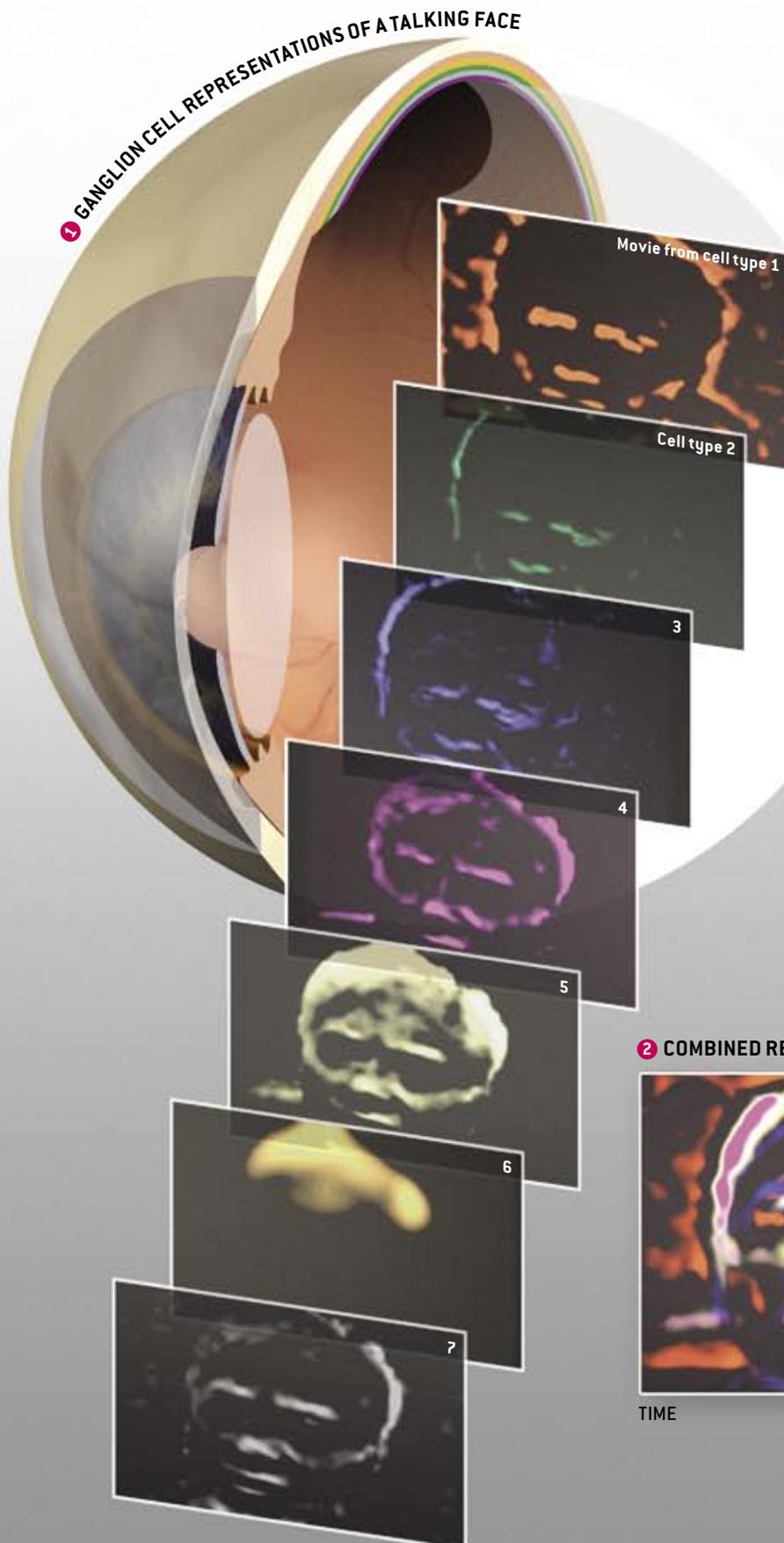
5

6

7
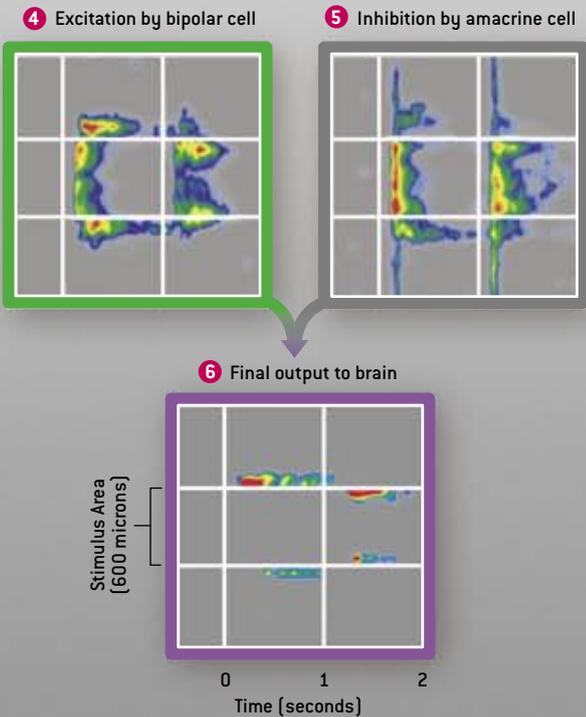
Our goal, of course, is to learn how each set of ganglion cells extracts meaning from the visual world. Because the retina is designed to handle information more interesting than a flash of light, we wondered what would happen when the retina witnessed a natural scene, such as a person talking. What would each of the 12 representations show? Would some feature be extracted by one movie but ignored by others?

Despite the seemingly straightforward explanations of how we captured the processing of a square of light, it is incredibly difficult to actually tap a living rabbit retina with enough electrodes during a simple one-second flash, much less a natural scene lasting a minute. For the latter exercise, we programmed the data from the flash experiment into a computer that simulates a famous artificial retina chip—the Cellular Neural Network—developed by Leon Chua of the University of California, Berkeley, and Tamás Roska of the Hungarian Academy of Sciences

2 COMBINED REPRESENTATIONS SENT TO BRAIN

TIME   1 second

ILLUSTRATION BY DON FOLEY; IMAGE DATA FROM FRANK WERBLIN

in Budapest. [*Tamás is the father of author Botond Roska.*] The system transformed the flashed square into a dozen spacetime patterns of excitation and inhibition that very closely resembled the patterns generated by the living retina.

Encouraged, we presented the programmed retina chip with a natural scene: one of us (Werblin) sat in front of a camera and talked for just over one minute. The simulator, which was programmed for this exercise by David Balya of the Budapest University of Technology and Economics, generated movie data for seven of the different ganglion cell representations ❶.

To confirm that the chip simulation was accurate, we measured the reactions of several neurons in the living rabbit retina to the talking face. It soon became evident that each population of ganglion cells acts as a filter, extracting a unique spacetime representation of the world that is sent in a unique movie to the brain. We imposed a color on each of the computer-generated representations to distinguish one from another.

For example, one filter (*opposite page, orange*) seemed to extract only the edges of features on the moving face, showing the world essentially in line-drawing form. Another filter (*purple*) accentuated the shadows underneath the eyes and nose. A third filter (*beige*)

produced highlights rather than shadows or edges.

Of course, our conclusions about the information that each of the 12 filters gleaned might not be correct. Unfortunately, it is impossible to represent accurately the patterns we recorded on the printed page because they run continually as movies, but it should be noted that they contain many blank intervals. Each movie only bursts into activity for a few milliseconds at a time and is otherwise dark. Nevertheless, our method shows that each filter is sensitive to a particular quality of the face's physical appearance and movement; each type of ganglion cell has its own unique way of depicting the world.

Coloring the representations also allows us to track the contributions of each set of ganglion cells to the final, combined representation produced when the movies are superimposed. We

combined the seven streams into one master movie. Four frames from different instances during Werblin's one-minute oration ❷ give a sense of how his face shifts to and fro as his lips open and close, with certain representations surging and waning, making him look like a ghostly apparition. This is what the retina produces. This is what the brain receives.

Our movies are only approximations. Still, they make it clear that, remarkably, the paper-thin neural tissue at the back of the eye is already parsing the visual world into a dozen discrete components. These components travel, intact and separately, to distinct visual brain regions—some conscious, some not. The challenge to neuroscience now is to understand how the brain interprets these packets of information to generate a magnificent, seamless view of reality. ⬛

MORE TO EXPLORE

Directional Selectivity Is Formed at Multiple Levels by Laterally Offset Inhibition in the Rabbit Retina. Shelley I. Fried, Thomas A. Münch and Frank S. Werblin in *Neuron,* Vol. 46, No. 1, pages 117–127; 2005.

Parallel Processing in Retinal Ganglion Cells: How Integration of Space-time Patterns of Excitation and Inhibition Form the Spiking Output. Botond Roska, Alyosha Molnar and Frank S. Werblin in *Journal of Neurophysiology,* Vol. 95, pages 3810–3822; 2006.

sa Movie created by the retina of the talking face can be seen at *www.sciam.com/ontheweb*

2 seconds

3 seconds

4 seconds