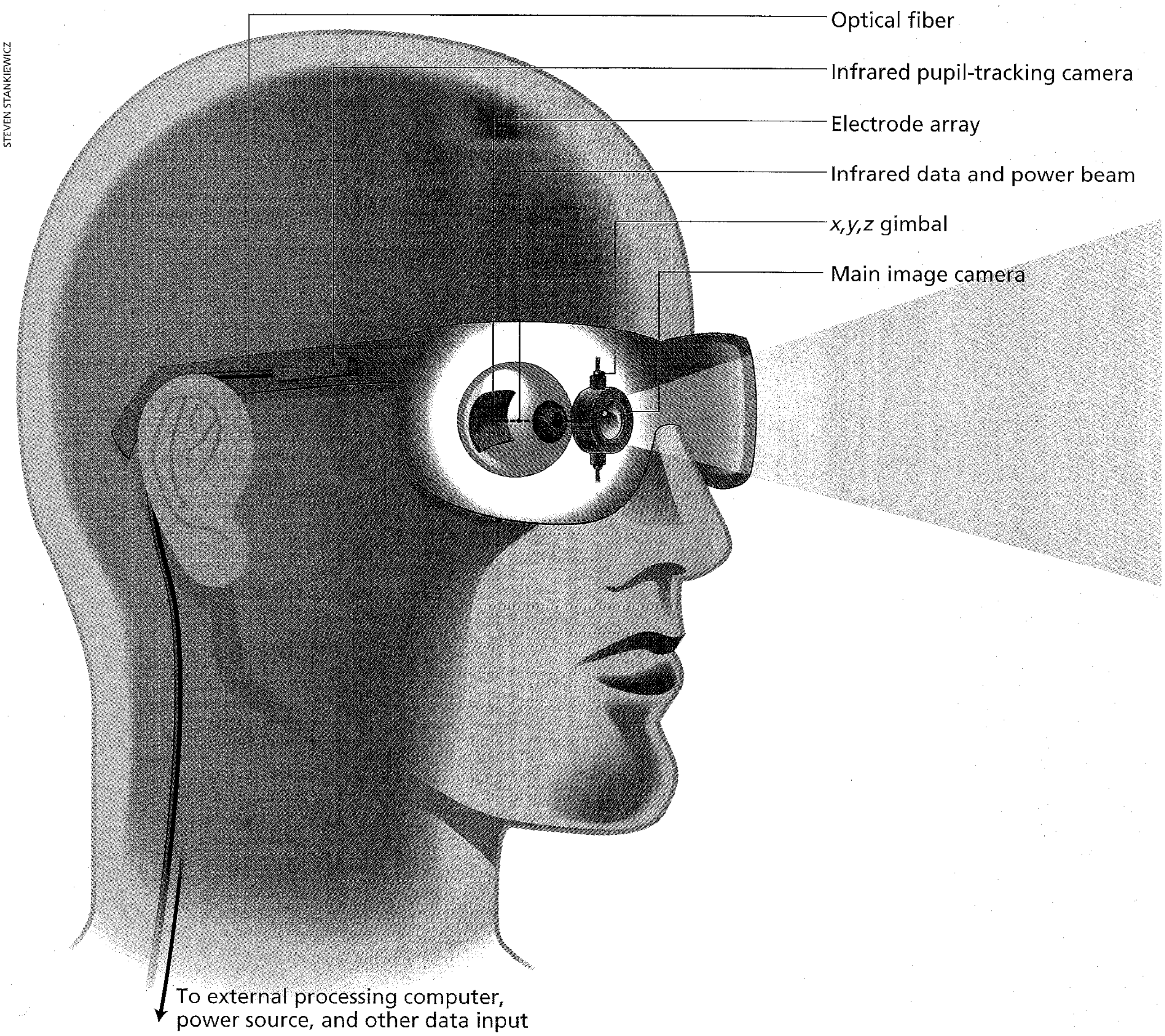
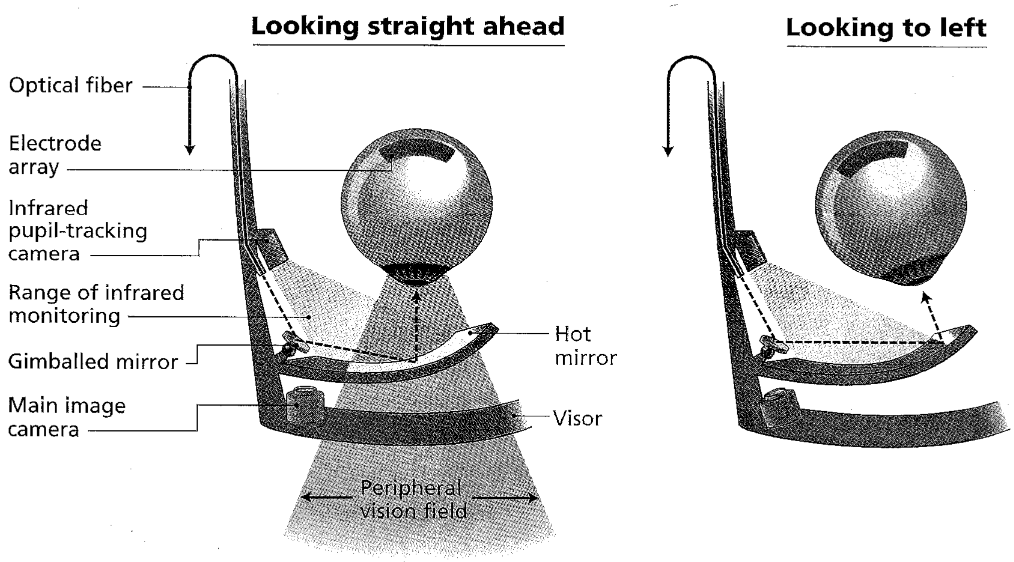
STEVEN STANKIEWICZ





[1] A biolectronic aid for those with no central vision, perhaps available around 2010, could rely on an electrode array in eye or brain to stimulate surviving neurons and elicit perceptions of light. Here, an IR laser beam carried on an optical fiber supplies an array implanted in the retina with both power and image data, derived from the main image camera. A pupil-tracking camera senses the eye's position from its reflection in the hot mirror (which reflects IR and passes visible light) and drives the gimballed mirror to deflect the IR beam onto the electrode array. The same drive signal also aligns the image camera with the eye position. Rough arrival dates for such prostheses are in the timetable [opposite].

Toward an Artificial Eye

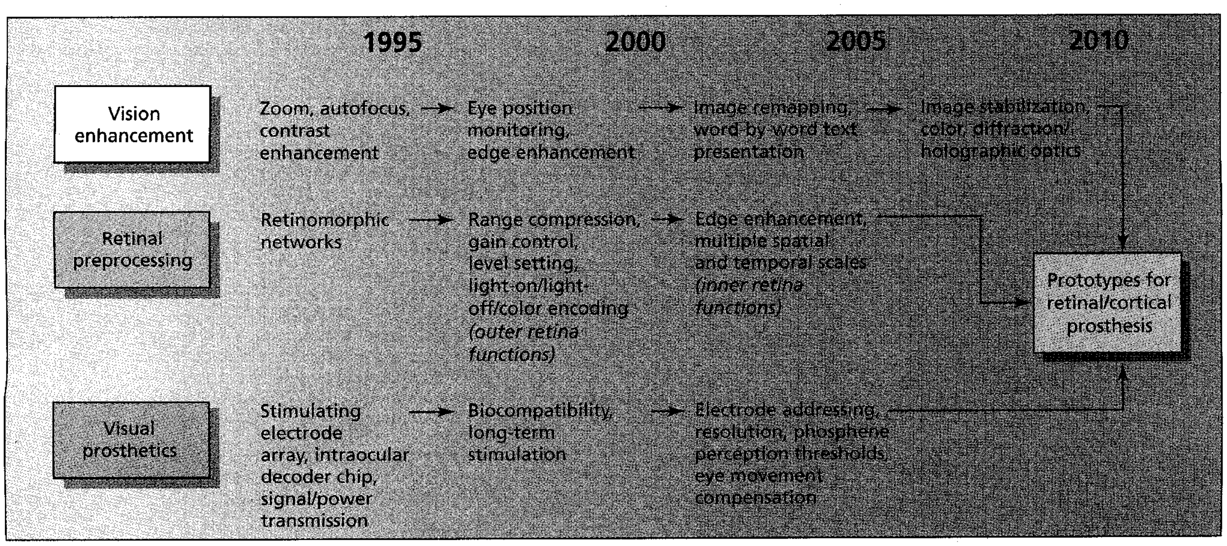
For millennia, restoring sight to the blind has counted as nothing less than miraculous. Today, computer engineering, ophthalmology, and biology are uniting in an effort to achieve just that. In this special report on bioelectronic vision, leading researchers describe how far the field has come and how far it has yet to go.

In the opening essay, Gislin Dagnelie and Robert W. Massof establish a perspective for ongoing research and estimate the prospects for the visually impaired [pp. 22–29]. Then comes a new look at the exquisite computation of the eye's neurons, by Frank Werblin and his co-workers [pp. 30–37]. The retina's molecular engineering is mimicked in silicon, in pathbreaking supercomputers known as neuromorphic vision chips. In this new breed of processor, say Christof Koch and Bimal Mathur, a dime-sized chip adapts in milliseconds to vast changes in illumination [pp. 38–46].

Surgeons naturally join engineers on the front line of biological implantation. Step by cautious step, both are investigating how points of light can be induced in the mind's eye of blind patients. John Wyatt and Joseph Rizzo are investigating an electrode array to be implanted on the retina: the electrodes respond to images from a camera by stimulating undamaged neurons [pp. 47–53]. An alternative approach by Richard Normann and his team, with a good start in human experimentation, has the electrodes implanted in the visually responsive areas of the brain [pp. 54–59].

Finally, Tomaso Poggio and David Beymer tell of the far frontier of computational and biological vision research: how all of us make sense of the visible world [pp. 60–69]. In this arcane realm, crucial to the success of any implanted device, individual brain cells and the largest brain structures perform a subtle physiological dance. Their steps in essence pattern the connections that make up the mind.

ROBERT BRAHAM
Special Report Editor



0018-9235/96/$5.00©1996 IEEE

After making a good living selling cellular videophones, Lucy Licht retired in 2010 at age 65. A few years later, an acute case of eye disease robbed her of her central vision. In the past she had played a mean round of golf, but now her only game was poker with large-print cards; any other reading was slow even with a portable vision enhancer. Then in the summer of 2020 she was fitted with a smart visual prosthesis, which not only brought back the joys of sight, but even improved on her original vision. With the new system—cameras slaved to her eye movements, combined with computer chips and stimulating electrodes, all interacting with her healthy nervous system—she also has some extra visual skills. For instance, she can zoom in on her golf ball and track it in flight, while comparing it with a freeze frame of her stroke tucked in a corner of her view. The prosthetic vision device, partly carried in her pocket, partly hidden under a visor, sends her the images of the outside world over an infrared beam—which conveniently enough can be used to transmit the latest Internet broadcasts, as well as private faxes and cellular communications.

SURELY NO ONE READING THIS special report could fail to be impressed by what has been achieved and what is in prospect for prosthetic vision devices rooted in biological functioning. The temptation is to conclude that the technology is at our doorstep, both for robotic systems on their own—in self-guided vehicles, say—and for robotic systems literally in direct communication with the nervous system of visually impaired people, as with our fictional Ms. Licht. In the long term, electronic sight should be available to those with no vision whatsoever, while for those with what is termed low vision, electronic aids should emerge rather earlier in the next millennium. As astounding as the technology is, it must be understood in terms of its still more astounding potential.

Three types of vision rehabilitation are being investigated, each having its own goals and probable paths of development [see timetable, p. 21]. The three types are:

- Enhanced vision, which refers to aids that process the image for maximum visibility and then present the information to still viable parts of an individual's seeing retina.
- Prosthetic vision, which presents processed visual information to the inner retina or visual pathways through electrical stimulation of the surviving neurons [Fig. 1].
- Artificial vision (not addressed in the time table), which processes and interprets visual information and presents the result to the individual through another sensory modality.

All three systems may reduce image information to symbolic or schematic form before presentation, but for enhanced and prosthetic vision, this is not necessary.

Note that low vision is a qualitative term. A person with low vision cannot perform ordinary tasks involving sight because of a visual condition impossible to correct medically, surgically, or with the aid of lenses. In practice, this includes anyone with a corrected visual acuity of 20/40 or worse in the better eye, with restricted visual fields, or with impaired contrast sensitivity.

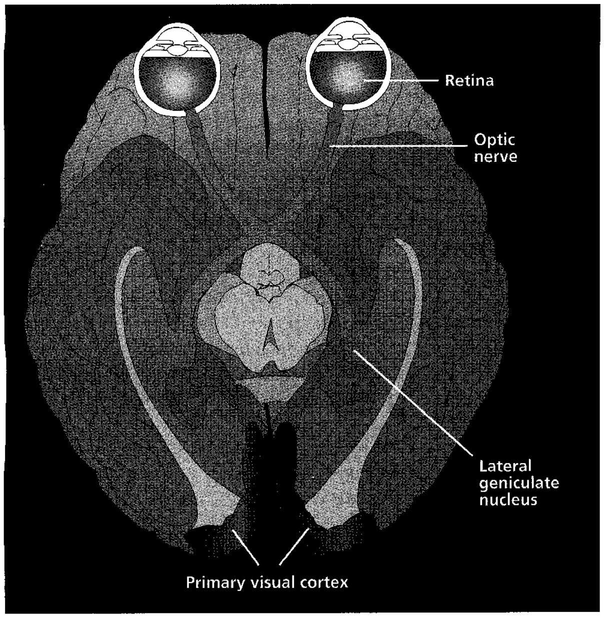
In 1990, the latest date for which data are available, some four million people in the United States fell into this category, the majority of whom were over 65 years old [see "Who benefits," pp. 26–27]. With the increase in the proportion of the elderly in the population at large, this number is expected to double in the next three decades.

For an estimated 1.5 million people in the United States, some everyday visual tasks are impossible, even with the help of simple aids, also known as assistive devices. Over half of these visually impaired meet the criteria for legal blindness (corrected visual acuity of 20/200 or less in the better eye, or a combined visual field diameter in both eyes of less than 20 degrees), and approximately 100 000 people are truly blind,

GISLIN DAGNELIE & ROBERT W. MASSOF
Lions Vision Research and Rehabilitation Center,
Wilmer Eye Institute, Johns Hopkins University School of Medicine

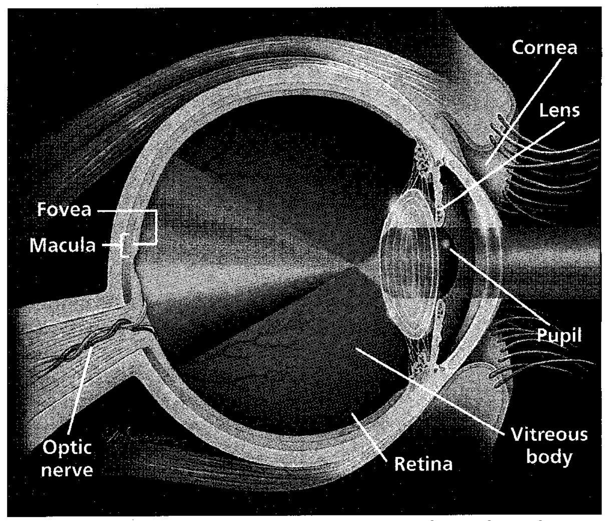
Defining terms

Action potential: also known as a nerve impulse or *spike*.
Adaptation: see *plasticity*.
Age-related macular degeneration (AMD): a vision disease that strikes mainly the elderly, its most usual effect being the destruction of cells in the central area of the retina.
Amacrine cells: two types of retinal cells that perform a broad field-erase function. The *narrow-field* amacrine cells have a narrow spread of processes and primarily feed back to *bipolar* cells to cut short their activity in time. The *widely ramifying* amacrines broadcast activity related to change and movement over wide spaces on the retina.
Artificial vision: electronic circuitry for processing and interpreting visual information, presenting the result to the individual through a sensory modality other than sight.
Axon: output fiber from a nerve cell. Each retinal ganglion cell has a single axon, typically about 1 μm wide and 5 cm long.
Bipolar cells: in the retina, the third sheet of cells, which accept the information formed by the interactions of the *horizontal* cells and *photoreceptors* and convey it to the inner retina, where further interactions occur. There are two, complementary types: ON bipolar cells respond to increases in intensity; OFF bipolar cells respond to decreases in intensity.
Block matching: a popular machine vision algorithm for computing optical flow. At a particular point (pixel) in the image, a block of, say, 8-by-8 pixels is chosen from the current frame and compared with all 8-by-8 blocks near the pixel in the previous frame. The best match for the subframe yields an estimate of the point's movement.
Convolution: in the case of vision, a linear, shift-invariant operation by which an image is filtered by a particular function.
Dendrites: the regions of the neuron that receive synaptic input.
Early vision: processes that recover the physical properties of the visible surfaces of solid objects from their two-dimensional intensity arrays.
Enhanced vision: assistive devices that present visual information to whatever remains of an individual's sighted retina after processing the image for maximum visibility.
Fovea: the small central portion of the retina, corresponding to about 1 degree of visual field, where a very high receptor density provides the finest resolution for observation of fine spatial detail.
Ganglion cells: the output cells of the retina. Their *dendrites* reach up into the inner retina and read out activity formed by interactions of *bipolar* and *amacrine* cells. The *axons* of the ganglion cells form the *optic nerve*.
Horizontal cells: the network of interconnected neurons that form a sheet just beneath the *photoreceptors*. The horizontal cell network is responsible for averaging visual activity over space and time and for controlling the gain and offset of the *photoreceptor* signal.
Hot mirror: a mirror that reflects infrared light but is transparent to visible light.
Line width (minimum feature size): the smallest line width used in a given semiconductor design. While most university-designed analog chips use 2.0- or 1.2-μm design rules for convenience, high-density memory chips can be built with 0.35-μm technology.
Macula: a larger central region of the retina, including the *fovea* at its center, that corresponds to about 30 degrees of visual field.
Narrow-field cell: see *amacrine cells*.
Neural activity: activity in neurons, expressed and measured in millivolts, that is a measure of signal intensity. Graded-potential cells, like most of those in the retina, polarize in the range from about –30 to –50 mV. In *spiking* cells (like the *ganglions*), spike frequency is the measure of intensity. A typical neuron is at rest near –70 mV and action potentials polarize neurons to near +10 mV for about 1 ms. Typical sensory system spike rates are up to about 50–60 per second.
Optic nerve: a bundle of about one million nerve fibers that carries the visual message from the eyeball to higher visual centers in the brain.
Optical flow field: the output of a vision algorithm or circuit that assigns every location in an image a vector indicating the direction in which the image intensities moved during the preceding interval.
Parameter space: a set of coordinated parameters describing an object. From the computer graphics perspective, these are the "control knobs" for rendering an image of the object. From the computer vision perspective, the parameters are estimated from an image of the object.
Percept: loosely, the smallest component of visual sensation.
Phosphene: the sensation of light produced by electrical or mechanical



BARCAN/LYONS DESIGN INC.

[2] The main visual pathways extend from the eyes to structures deep within the brain. Neither in the retina nor in the brain do electrical signals travel a straight path.



KARAPELOU/PHOTOTAKE

[3] This horizontal cross section of the human eye shows the main areas affected by blinding disorders. The macula and fovea contain the highest concentration of photoreceptors, and disease of these areas is particularly debilitating. The retinal surface is interrupted at the exit point of the optic nerve, creating a normal physiological blind spot unnoticed by sighted individuals.

unable to make any practical use of their vision. Because of this variability, it is helpful to group the technological goals for all types of vision assistance.

Normal sight

The human visual system is a remarkable instrument, consisting of two mobile image-acquisition units, each with formidable preprocessing circuitry, placed at a remote location from the central processing system. It is faced with transmitting images within a viewing angle of at least 140 degrees and resolution of 1 arcmin over a limited-capacity carrier, the million or so fibers in each optic nerve, through which the signals are passed to the so-called higher visual centers, the lateral geniculate nucleus, and the visual cortex [Fig. 2].

The nervous system solved the problem by confining such capability to just part of the retinal surface: whereas the center of the retina has a 1:1 ratio between the photoreceptors (the specialized neurons that transduce the incoming light) and transmitting elements, the far periphery (our field of peripheral vision) has a ratio of 300:1, resulting in a gradual shift in resolution and other system parameters [Fig. 3].

The retina also faces ambient light that over a single day can change by a factor of 100 million in intensity. The solution here is to create two photoreceptor systems, each equipped with a gain control mechanism, and to follow them with further range compression in subsequent nerve cell layers. There is even a crude form of spectroscopy: three differently pigmented types of photoreceptors are alternated in the high-intensity receptor matrix and differences between adjacent receptors are encoded. There are also classes of cells that transmit information about rapid changes and edges.

At the brain's highest level, the visual cortex, an impressive array of feature extraction mechanisms can rapidly adjust the eye's position to sudden movements in the peripheral field of objects too small to be perceived when stationary. The visual system can resolve spatial depth differences by combining signals from both eyes (among other clues), with a precision less than one-tenth of the size of a single photoreceptor.

Eye movement compensation

When the eyes of a normally sighted person move, the image on the retina undergoes a translation. The same is true for a visually impaired person using an enhanced vision system: what is observed is a transformed image on a "projection screen," and an eye movement across the screen results in a shift of the

stimulation of the visual pathway of the nervous system.
Photoreceptors: two types of photosensitive cells that convert light to neuronal signals. Rods are the more sensitive and are used in moonlight, say. Cones are less sensitive, operate over a broad range of intensities, and convey color information.
Plasticity: changeability in the performance of a neuronal system, usually as the result of experience and usually manifest as changes in synaptic efficacy. Long-term plasticity may involve physical changes such as the formation of new synapses. Short-term plasticity may involve the enhancement of existing synaptic connections.
Prosthetic vision: processed visual information sent to the inner retina or other visual pathways through electrical stimulation of the surviving neurons.
Retina: the neuronal image-processing system that lines the back of the eye. Most processing done here is performed by spikeless, analog neuronal interactions.
Retinitis pigmentosa (RP): a disease of the photoreceptors, frequently inherited, which at its most advanced stage allows only a tiny island of central vision.
Retinotopy: the notion that receptor cells in the retina are mapped to points on the surface of the visual cortex of the brain.
Synapse: the functional contact between two neurons. Signal propagation can occur by physical contact with a bidirectional flow of ions, as in the *ganglions* and most neurons, or by chemicals known as neurotransmitters.
Spike: also known as a nerve impulse or action potential, the means by which excitation is propagated along an *axon*.
Supervised learning: a way of setting the weights of coefficients in a neural network (weights are roughly analogous to the strength of synaptic junctions in biological organisms). In machine learning, it usually indicates training a neural network through a set of examples consisting of a "correct" input and output.
Tuning: essentially, the sensitizing of neurons to specific optimal stimuli, such as a line of a certain orientation moving at a certain speed. A unit (biological or computational) is tuned when its response to specific input stimuli is at its maximum.
Widely ramifying cell: see *amacrine cells.*

enhanced retinal image. For a person wearing a more advanced device, a tracking retinal or cortical visual prosthesis, however, the situation may be different. If the camera is built into the natural or artificial eyeball, the shift occurs just as in normal vision; but if the camera is mounted on an external visor or tripod, an eye movement does not result in an image shift.

Shifting the image under the control of eye movement is important, however. Whereas a still image stabilized on the retina will quickly fade, a moving image will not. Flicker and other forms of temporal modulation will also prevent fading, as will head or body movements when cameras are carried around by the user. It appears that for early versions of prosthetic vision, a flickering but stabilized image will be acceptable. In a more elaborate "artificial eye," monitoring and feedback of eye position become indispensable. This is true for enhanced-vision systems as well as for prosthetic systems, albeit for different reasons.

First, for a visually impaired person to derive the greatest benefit from an enhanced-vision system, the image must be adapted to his or her particular blind areas and areas of poor acuity or contrast sensitivity, and then shifted around those areas in terms of the instantaneous information arriving at the eye. Now the thrust of all prosthetic vision devices is to elicit in the user via an electrode array perceptions of points of light (phosphenes) that are correlated with the outside world [Fig. 1].

Thus, to achieve the expected shift of the image across the stimulating electrode array, the camera capturing the image has to follow the wearer's eye or pupil movements by observing the front of the eye under infrared (IR) illumination. The eye-position monitor controls the image camera's orientation [Fig. 1, bottom]. If the main image acquisition camera is not mounted on the head, compensation for head movements also will be needed.

Finally, if a retinal prosthesis is to receive power and signal input from outside the eye via an IR beam entering the pupil, the transmitter has to be aligned with the intraocular chip. The beam has two roles: it sends power, and it is pulse- or amplitude-modulated to transmit the image data. The main imaging camera for each eye can swivel in any direction under control of eye movement, and each is located just outside the user's field of view so as not to block her remaining peripheral vision. This camera captures the image of the outside world and transmits the information through an optical fiber to a signal-processing computer worn on the body.

An even better solution is to slave the beam path to the eye position. This is achieved by two mirrors, one being large, curved, and fixed, while the other is small and gimballed. The large mirror is a special kind, known as a hot mirror, which lets outside light pass through to the peripheral vision but reflects infrared light. The hot mirror reflects two light sources: the infrared image of the front of the eye to the pupil-tracking camera, and the narrow data/power beam emanating from the optical fiber into the pupil. The data/power beam is correctly deflected into the hot mirror by the small gimballed one. When the pupil-tracking camera/computer unit detects a change in eye position, it sends signals to servomotor controllers of the main camera and the deflecting mirror. Thus the data/power beam and imaging camera are always aligned with the eye position and the implanted prosthetic chip.

Essential functions

Enhanced, prosthetic, and artificial vision systems may be useful even if they perform less well than the intact human eye. A short list of functions for any kind of advanced system in the future would include:

- Eye position monitoring. Accurate real-time information about the direction of gaze can be used to control the camera, remap the image, and transduce the image to a retinal prosthesis.
- Dynamic range compression. The system requires an adaptive mechanism to enable it to operate at illumination levels ranging from bright daylight to at least dim artificial lighting. The automatic gain control systems in today's video cameras are already close to meeting this requirement. Even better, by means of currently available analog and digital techniques, an image's luminance or color contrast can be automatically compressed or expanded, so that the full brightness range can be used (see "Neuromorphic vision chips," pp. 38–46).
- Local contrast enhancement. Contrast must be enhanced only over short ranges; otherwise, making local contrast visible might blank out areas of stronger contrast by driving them into saturation. What is needed is a boost of high spatial frequencies, visible as edge enhancement. While real-time spatial filtering of video-rate images on a single chip is not yet possible, it may very well be available in a decade. Moreover, new developments in optical computing may allow this and other types of filtering in true real time.
- Image remapping. This feature is crucial if a person with field loss in both eyes, or in an eye with much better vision than its fellow, is to view a scene using whatever area of the retina(s) remains functional [Fig. 4]. The blind spot(s) could obscure vital information, such as the approach of a vehicle, so the equipment must be able to distort (warp) the image in such a way that information from the blind spot is mapped onto the seeing retina. Obviously, the image around the blind area will look squashed, but users should become accustomed to this distortion rather quickly [Fig. 5].

The system needs also an accurate map of the wearer's visual fields and moment-by-moment information on eye position. Image remapping may also be necessary for individuals who experience distortions, for which the algorithm could be used to compensate. In that case, the system would need an accurate map of the distortions, which can be obtained only by means of special tests (for example, aligning dots on a screen into a regular pattern under strict eye movement control).

- Customization. Whatever the capabilities and functions of the system, it must be adaptable to the user's needs and remaining visual capabilities. Moreover, the extent to which individual features are used may vary from one person to the next, according to the task at hand. Ideally, this tailoring should require only a few adjustments by the manufacturer and only minimal training and effort on the part of the user. This is not the same as having the system adapt to the long-term learning of the patient's visual system, which is discussed below.

Desirable functions

Depending on configuration of the system, several desirable functions can be addressed at a later stage of development. The first set of functions, for prosthetic and artificial vision, begins with servo control of the cameras and the infrared beam sending data and power (if the camera is external) in step with eye position. Also, high resolution of the image-capture chip should be sought, be it an array of "dumb" CCDs currently on the market or the neuromorphic vision chips that can take over much retinal processing. A high-resolution system would eliminate the need for a zoom lens, and would allow a minutely detailed central portion of the image surrounded by a less detailed area, just as in the normal visual system.

For enhanced vision, the needs are more familiar, involving displays and optical projection. Current electronic display systems for the visually impaired rely on cathode-ray tubes (CRTs) because of their better resolution in visible lines per image, larger screen area, and higher contrast; however, their size and weight are inconvenient for both portable and stationary systems. Research in flat-panel display technology over the next decade is likely to produce high-resolution, high-contrast screens from a centimeter to a meter in size, at a price that can compete in the consumer market.

For lightweight, head-mounted enhanced vision systems, a projected image has to be presented to the wearer. Here, further developments in the areas of holographic optical elements and diffractive optics are likely to produce the components necessary for new forms of image projection.

The final issue in displays is color. The main reasons monochrome systems are preferred over color systems for the visually impaired are their superior resolution and contrast and lower cost. These considerations should disappear as components become cheaper and better.

Advanced features

Some features we believe are desirable and achievable for enhanced vision, let alone prosthetic and artificial vision, may at first seem unnecessary. Familiar by now in video and television, they are all concerned with controlling the input of the image—freeze frame, image storage and instant replay, automatic tracking, zooming, and image stabilization—but have rarely been considered for augmenting human vision.

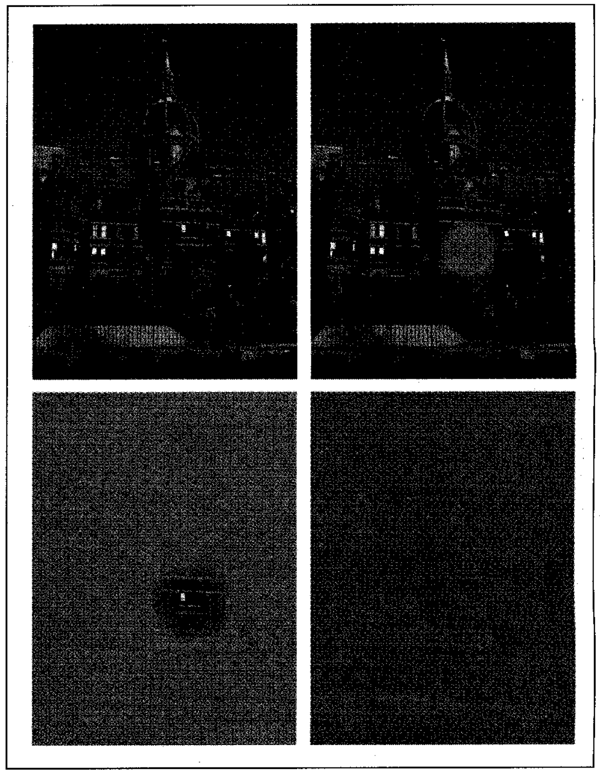
These image control functions are particularly important to those with low vision. The components of the visual field should be easy to track with algorithms for feature extraction used in conjunction with the hardware/software arrangement that lets the camera roam in any direction. Such tracking makes it possible to visually enhance depth, parallax, and other motion information (stairs, a curb, an approaching vehicle), or announce a warning through an auditory or tactile signal. Systems capable of interpreting signs and landmarks would help wearers orient themselves in unfamiliar surroundings, complementing rather than replacing the long cane.

Image stabilization could be done in two ways: by using accelerometers or gyros mounted on the cameras, or by responding to low-pass temporal filtering of the image as it shifts across the array of picture elements (pixels). Image stabilization would also exploit the roaming camera, and could greatly facilitate system use at high magnification, especially by tremulous elderly people. Take the case of someone who finds it difficult to read print but who wishes to prepare a dish from a recipe. He may have to refer to the recipe repeatedly, and would have to pick up a magnifier, relocate the place in the recipe where an ingredient is mentioned, read the instruction, put down the recipe and magnifier, and divert his gaze again to cook. Storing the recipe in video memory and then displaying it in a corner of the visual field could simplify this task.

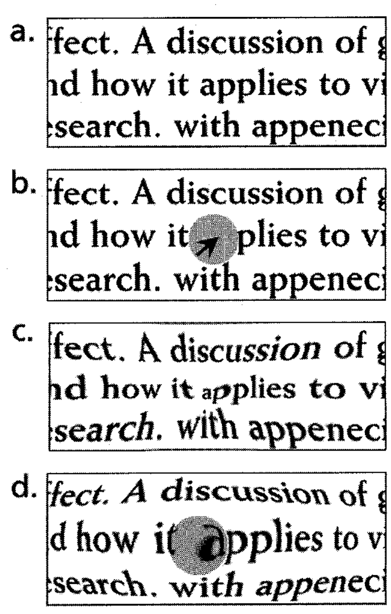
The picture-in-picture function is all the more desirable if the user is reading by means of another computer aid for enhanced vision: word-by-word (scrolled) presentation of highly enlarged text or images. The scrolling (sometimes called serial presentation) presents a person incapable of sustained reading with continuous text in a more readily navigable form. For instance, excess eye movements would be eliminated if a sequence of words was shown in a single location, as a scrolling string without line or page breaks, or conveyed by a different sensory modality as spoken text or Braille.

Two final advanced features are the ability to interpret color or depth, and the display of synthetic computer graphics or video from sources other than the local visual field. For the first task, any visual prosthesis using electroneuronal transduction may have great difficulty conveying those image attributes. But this aspect of the visual world can be mapped (temporarily) onto another sense or component of the image, and after some training, the user should be able to interpret the information.

Having the device incorporate the view (and sounds via a headphone) from any external data source—Internet feeds, virtual reality scenes, and video communications, for example—into the local visual field might seem like a gimmick to normally sighted individuals. But to a low-vision patient, they could



[4] An undistorted image [top left] may appear with a washed-out central area [top right], a washed-out periphery [bottom left], or "foggy," from low contrast sensitivity [bottom right].



[5] Given an accurate map of the patient's blind spots, plus a set of warping algorithms optimized for specific viewing tasks, information may be mapped away from the blind retinal area, held in a frame buffer, and warped over the blind area.

mean the difference between disability and employment, between isolation and integration.

For truly specialized needs (and as often fancied by writers of science fiction), no intrinsic reason prevents limiting the device to visible light or to a single vantage point. The main camera imaging could be in ultraviolet or infrared spectral bands. Features of this kind would be used most likely in specialized applications such as surveillance and interpreting X-rays using false color.

Who benefits?

Blindness disables millions of people throughout the world. In developing countries, it is mostly caused by cataract, malnutrition, or infection. Then there are vision problems and blindness due to chronic systemic diseases, such as tuberculosis, pneumonia, and sexually transmitted diseases. Among preventable eye diseases, the three most common are vitamin A deficiency; trachoma, an inflammatory disease endemic in Asia and Africa that is caused by a micro-organism transmitted by flies; and onchocerciasis, also called river blindness, a parasitic infestation transmitted by the blackfly.

In developed countries, ready access to medical and surgical care mostly eliminates these problems. Instead, almost all blindness and low vision (functionally disabling poor sight that is uncorrectable by lenses) is associated with diseases that destroy tissue in the retina, optic nerve, or visual cortex. Further, because people live much longer, age-related degeneration of the macula (AMD), a key portion of the retina, is the leading cause of blindness.

The visual system can fall prey to other attacks. In glaucoma, pressure in the eye or weakness of support tissue damages the fibers of the optic nerve.

The group of diseases known as retinal dystrophies affect the photoreceptors and may be inherited. One of the best-known is retinitis pigmentosa, commonly first made manifest by night blindness and loss of the mid periphery of the visual field; this area of loss then expands to form a ring hole in the field of view that spreads outward and inward until only a narrow island of central vision is left.

Other conditions affecting the retina derive from systemic diseases, such as diabetes; a clot blocking blood flow to the retina; or retinal detachments. Blindness may also follow a stroke that damages the visual cortex.

There is no single standard of "good vision." The principal measures used are acuity, sensitivity to contrast, and defects in the visual field. Other tests measure hyperacuity (for instance, the ability to align two line segments), depth perception, color discrimination, dark adaptation, and the ability to assess facial expressions, not to mention reading speed and visually guided manipulation of objects.

Defects in the visual field are variously mapped. In hyperacuity and depth-adjustment tests of the fovea, the most discriminatory part of the retina, normal subjects can achieve thresholds corresponding to resolutions of 5 arc seconds or less. If the same tests are performed for the periphery, thresholds rise rapidly—at least three times as fast as the size increase in the smallest recognizable letter on the acuity chart.

Low-vision patients fall into several broad classes, depending on the types of defects brought on by eye disease.

• Lack of foveal development from extreme premature birth; albinism; an absence of cone receptors, causing aversion to light; and involuntary oscillating eye movements (nystagmus)—any of which degrades performance on any acuity-related task but which can be somewhat helped by magnification. Many patients also have problems with contrast sensitivity and illumination. Visual acuity may be low, in the 20/100 to 20/200 range. (The

Eye or brain interface?

Once an image has been captured and preprocessing has been performed, a transducer will present the information to the user. To disregard for the moment the presenting of visual information in auditory or tactile form, only two methods are available: enhanced vision utilizing surviving photoreceptors and prosthetic vision, in which retinal or higher neurons are electrically stimulated. If both are options, which should be chosen?

If a person has enough sight left for tasks such as orientation and mobility, simply inputting light to the retina might be adequate for most visual tasks, providing the input was suitably packaged. At present, training in special fixation techniques, such as off-center (eccentric) viewing, scanning, and tracking can increase proficiency at tasks that normally require the now-useless visual field area.

In the future, such training could be simplified with eye-movement monitoring and techniques such as biofeedback. Moreover, preprocessing features in the enhanced vision system could lessen the need for training, and allow even patients with severe vision loss to function adequately using visual information.

Higher-order neurons may be stimulated electrically by the methods developed in retinal and cortical stimulation research [see "Ocular implants for the blind," pp. 47–53, and "Cortical implants for the blind," pp. 54–59]. As to whether retinal or cortical stimulation is preferable—that is, visual processing at early or late levels in the human visual system—much depends on the results of future physiological experiments.

However, as explained in "The computational eye" [pp. 30–37], multiple levels of processing already take place in the retina. The same is true for the lateral geniculate nucleus and the visual cortex, which probably also contain elaborate feedback mechanisms. At those levels, a simple feed-forward processing scheme may not suffice under any circumstances. As the interface is placed later and later in the visual pathway, the complexity of preprocessing required to approximate "lifelike" input signals, and hence presumably lifelike vision, increases.

More important, cells become ever more specialized as the signal travels up the visual pathway and as multiple representations of different aspects of the visual world are being processed in parallel. Therefore, the likelihood decreases that an interface element can provide its target cell(s) with the type of information to which it once learned to respond. A typical cell in the primary visual cortex can be extraordinarily "learned," being tuned, for example, to respond with maximum output to the presence of a red-green border 18 degrees from fixation at 9:30 hours azimuth, running under a 45-degree angle, and moving toward the eyes' fixation point at 50 degrees per second. To make matters even more difficult, neighboring cells may code for the same edge, but at a different resolution, or for a different stimulus aspect altogether.

True, many of these cells, especially in the cortex, may be able to relearn [see "Learning to see," pp. 60–69]. Still, it is far from certain how flexible this capacity is, or to what extent it survives, especially late in life.

On the other hand, it may be necessary to resort to a higher-level interface if early levels of the system are too severely damaged to react to stimulation, let alone reliably transmit information. But intervention at one level may preclude future intervention at a lower level, should a better method of treatment at that location become available. In the meantime, while prosthetic systems are still under development, the wisest course seems to be to choose the lowest viable level of neurons for electroneural transduction.

What can the visual system learn?

In a person with blind areas in both eyes, enhanced vision systems use remapping and other methods to concentrate information on the functioning parts of the retina. This strategy forces the visual system to perform more tasks in certain parts of the visual pathway's connections to the viable retina, while leaving other parts unused. In prosthetic vision, on the other hand, the visual system is suddenly forced to process information in areas underlying convention is to assign 20/20 to the letter size a normally sighted person can just recognize at 6 meters.)

• Central field loss—generally due to macular degeneration but also linked with some retinal dystrophies and optic nerve disorders. In age-related macular degeneration, photoreceptors in the area are destroyed. In the disease's "dry" form, an area of atrophy slowly expands into a horseshoe-shaped blind region around the fovea, closes into a ring, and invades the fovea. In a "wet" form of the disease, new blood vessels grow under the retina. Acuity suffers badly, dropping to 20/70 to 20/200 or 20/400 in the wet form if treated, as low as 20/1000 if scarring continues. Most patients have only a milder form of the disease and less loss of vision. Even so, it forces them to rely on their peripheral vision, which, even with proper magnification, has such poor text resolution that adjacent letters look jumbled together.

• Peripheral field loss—due to advanced glaucoma (pressure around the optic nerve fibers), optic nerve disorders, or diseases such as retinitis pigmentosa. A special case is the loss of much of the left or right field of vision after a stroke has affected the visual cortex. These patients often find it hard to orient themselves and avoid obstacles. In advanced stages (tunnel vision), the restricted field hampers reading, and acuity and sensitivity to contrast may also be reduced.

• Localized gaps in vision—known as scotomas, which may be the results of diabetic damage to the retina or blocking of retinal blood vessels. Nowadays they are often an unfortunate side-effect of laser surgery intended to prevent further damage from diabetic retinopathy or wet age-related macular degeneration.

• Reduced contrast sensitivity—often a secondary effect of a disease process that has been left undiagnosed. It has much the same effect on vision as a dense fog: the luminance in images vary and color looks washed out. Slight changes in brightness and hue across a scene become very difficult to detect. Not surprisingly, patients can find it hard to recognize faces.

• Illumination and adaptation problems—often the lot of patients with diseases of the photoreceptors. Many patients with age-related macula degeneration report a need for bright illumination, but this need can often be addressed by contrast enhancement.

• Distortions. Any distortion could point to a membrane growing under the retina. The standard explanation alludes to wrinkling of the retina, but it seems the cortical map of the visual field may reorganize itself around the projection of a small scotoma, such as a laser scar.

• Unexplained problems in sustained reading. These may be said to occur if a patient is unable to read running text at a print size a few times her acuity limit, but can "spot-read," say, a single word or number with the same print size. Here, the sufferer may have a small island of relatively good acuity surrounded by a scotoma, as in mid-stage "dry" age-related macular degeneration. Alternatively, a scotoma (blind area) to the right of the fixation point (sometimes called a leading scotoma) may deprive the patient of the ability to glance ahead, which presumably enables readers to choose the next fixation target.

—G.D. and R.W.M.

that may have been unused for many years, possibly from birth.

This raises three questions. Can a "busy" cortical cell population usurp "idle" cell populations? Can cells in the busy population learn new tasks? Can cells that have never processed visual information learn to do so?

Several things are known about the learning ability of cells in the visual system. First, if the brain is not visually stimulated by the time its owner reaches a certain age, the person may remain blind no matter how well the eyes and retina are later made to function. Thus in a case where the high-resolution area of the retina called the fovea is not properly developed, the cortical area that works with the fovea may never achieve proper resolution, even with high-resolution transducers and extensive training. The extremely elaborate neuronal connections that would be required are unlikely to form after the first months or years of life. The same is true for binocularity and depth perception and most other aspects of vision. However, it is unknown just how much cortical reorganization is possible, and up to what age.

Conversely, there is some evidence that cortical cells, when deprived of input from the retinal area they once served, form new connections and accept input from neighboring retinal areas: this is one explanation for distortions around blind areas. New input to these cells from the original area of the visual field could lead to conflicts, perhaps even double vision, but these would probably be eventually resolved through relearning.

The position today

The vast amount of research being conducted in visual perception and image processing, the marked increase in research and clinical work aimed at understanding low vision over the last decade, and the diverse equipment being developed for the blind make it all the odder that so little effort is going into developing broadly applicable technology for visually impaired users. Most devices now available are either task-specific (closed-circuit television readers, screen enlargement software) or range-specific (such as magnifiers with fixed working distance and telescopes and microscopes with fixed magnification).

There have been a few research projects in low-vision enhancement in recent years, notably the European Community–funded portable optical vision enhancement system, which is still in development. To date, the only successful and commercially available system that integrates video technology, image processing, and low vision research is the low vision enhancement system (LVES), developed at our research center [see box, p. 28].

Among the advantages of LVES over conventional optical low-vision aids are its variable magnification and working distance, autofocus operation under control of a digital signal-processing chip, and (optionally) enhancement and reversal of contrast. Its advantages over current optoelectronic low-vision aids such as video magnifiers are portability (compared to most closed-circuit TVs) and range of application.

Steps toward the first visual prosthesis

As indicated in the timetable on p. 21, work toward a visual prosthesis is developing along three parallel tracks over the coming decade: one each for vision enhancement, retinal preprocessing, and vision prosthetics using retinal or cortical stimulation.

• **Track 1: Vision enhancement.**

If this track is subdivided into capture, processing, and display of image information, a comparison with our list of needs discussed earlier shows areas close at hand and others still in intensive research. Engineers worldwide are already investigating high-resolution image capture over large dynamic ranges, high-resolution flat-panel displays with high contrast, and advanced optics.

Areas specific to low-vision applications, as detailed above, require more concentrated effort: these include eye position monitoring, camera movement control, local contrast enhancement, image remapping, automatic track and zoom, image stabilization, feature extraction and recoding, frame storage, symbol translation, and recoding a lost dimension (color, say).

Research in enhanced vision serves a dual purpose: to improve

Where electronic vision enhancement stands today

Over three million people in the United States are affected by low vision—defined as any chronic visual condition, uncorrectable by lenses, that impairs the ability to perform everyday functions. So the reports of the mid-1980s on virtual-reality helmets and on the image-processing technology being developed by NASA researchers proved an inspiration to The Johns Hopkins University's Wilmer Eye Institute in Baltimore, Maryland. The ophthalmologists and researchers there proposed a head-mounted video system for low-vision patients and in late 1985 applied formally to NASA for a technology transfer project.

Early studies indicated that components from commercially available camcorders might do for a first-generation system for acquiring and displaying images, so that NASA's contribution to the project would be in the area of image processing. For the wide-field viewing system required, 19-mm monochrome cathode-ray tube (CRT) screens were the only displays on the market with enough resolution and contrast. So until better options emerge, the low-vision enhancement system must be monochromatic.

A contract to develop the system's projection optics went to Polaroid Corp., Cambridge, Mass. Prototypes were manufactured early in 1991, but it was not until the middle of 1994 that the first production model was completed by Visionics Corp., a Golden Valley, Minn.-based company founded specifically to manufacture and market the enhancement system.

The enhancement system has two main components: the headset and a control unit, which doubles as a battery pack and is worn around the waist. On the headset are three cameras based on charge-coupled devices (CCDs): two for orientation and a third for refined viewing. The orientation cameras located in front of the wearer's eyes have a fixed-focus objective lens, a 60-degree field of view, and near-unity magnification. The two provide separate images with near-natural binocular disparity.

The third camera, with 1.5–9X variable magnification and variable focus, is placed

above the center axis of the unit and provides both eyes with the same cyclopean view. Its focal range at low magnification is about 1 cm (macro) to infinity; as magnification is increased, this range is progressively limited. Permanently attached to the zoom objective lens is a low-power (2–4-diopter) lens. Its job is to reduce the near-point distance at maximal zoom: with a 2-diopter lens, the focus range at 9X is 25 cm to infinity. When the wearer is working close up, the camera housing can be tilted downward.

Signals from the cameras are sent to the control unit, where a digital-signal-processing chip maximizes edge sharpness in a previously selected region of interest. Switches let the user select the camera, motorized zoom, autofocus lock (in which case the zoom lever can double as a manual focus control), and contrast. With contrast selection, the user may choose normal or inverted polarity, regular or enhanced contrast, and, if enhanced contrast is selected, the midpoint of the brightness scale. (The belt pack also has a nickel-cadmium battery providing 1.5–2 hours of operation; the unit can also be powered with 12 V dc using an ac or car-battery adapter.)

After all the selecting and conditioning, the video signals are returned to the headset for display on two forward-facing CRTs. The CRTs are mounted in the rear half of the headset's arms, which run along the wearer's temple. Aspheric optical projection systems image the CRT screens on the wearer's eyes at infinity, subtending 50 degrees horizontally by 38 degrees vertically. Given the CRT's resolution of 300 vertical line pairs, the image resolution is 5 arc-minutes—possibly a limiting factor for wearers with visual acuity of more than 20/100.

To customize the system, the 8-mm exit pupils of the projection systems are centered on the person's pupils, and his (or her) spherical and cylindrical eyeglass prescription is built into the system. A system of four straps supports the 1.0-kg weight of the overall system and keeps the projection system aligned with the eyes.

Under a grant from the Department of Veterans Affairs' Division of Rehabilitation R&D, evaluation research is now under way at Wilmer Eye Institute and at four Veterans Administration medical centers. The system is also being prescribed by over 30 centers in the United States and at centers in Waterloo, Ont., Canada; Heidelberg, Germany; and Oxford, England. —*G.D. and R.W.M.*

the tools of vision rehabilitation and to provide a "front end" for the electroneural type of visual prosthesis. Other functions, especially image remapping, interpretation, recoding, and translation into other sensory modalities, may become superfluous once a fully functional visual prosthesis is available.

• Track 2: **Retinal preprocessing.**

Of the three tracks under consideration, we estimate the one toward modeling of retinal preprocessing to be nearest our goal. The neuromorphic chips described in this report by Frank Werblin and his co-workers [pp. 30–37] and by Christof Koch and Bimal Mathur [pp. 38–46] can already perform sophisticated operations in the areas of dynamic range, temporal processing, and lateral interaction. Further advances in retinal research will be required before the many functions performed by inner retinal cells can be adequately modeled (though Werblin's reconfigurable models provide a good platform for experimentation); thus they are not addressed in detail in the pioneering implantation work of John Wyatt and Joseph Rizzo [pp. 47–53] and Richard A. Normann and his co-workers [pp. 54–59].

What particularly remains to be provided is the gradient in the human retina from the central, high-resolution fovea to the low-resolution periphery in cell density, cell population, and cell connectivity. Coupled with an imaging camera, such chips could be housed in a belt pack and provide input to an electrode array. Particularly useful would be versions of a single chip containing the gradients in density and connectivity from center to periphery.

• Track 3: **The biological interface.**

The above-mentioned articles by Wyatt and Rizzo and by Normann's group on retinal and cortical implants, respectively, demonstrate the potential of their techniques for building a machine-person interface. However, it is also clear that major research and engineering efforts will be required to achieve viable long-term signal transduction.

Topics to be addressed range from biocompatibility of implant and human host to the effects of long-term electrical stimulation on retinal and cortical tissue, from effective stimulus waveforms and thresholds to the spatial and temporal resolution of the electrical signal, and from implantation and stabilization techniques to signal transmission and power supply. Progress is being made in all these areas, and the short-term feasibility of both methods has been demonstrated in human subjects as well as in other primates.

Yet, of the three tracks, we estimate this track towards completion of a durable and high-resolution prosthetic device to have the longest distance to cover. Normann and his co-workers estimate the arrival of a working device within 10 years. We, however, submit that any such device appearing within that time would be a simple "proof-of-concept" model.

Yet, even if a conservative view of the available elements is adopted, a prototype visual prosthesis can be put together. The device might have no other purpose than to provide totally blind persons with enough light perception and rudimentary resolution to enable them to navigate and in general to perform enough daily activities to live independently. To be sure, the prosthesis would only complement a person's current skills: as an example, a cane would still be better than this device at detecting a low curb.

The retinal version of the device might consist of a small CCD camera mounted inside a pair of sunglasses. From the end of the visor's temple arm, a cable would carry the camera signal down to a belt pack containing the battery plus switches with which the user would manipulate zoom and contrast. The return signal would go to an infrared laser or light-emitting diode for signal and power transmission into the eye.

A chip over the retina would optimally perform not only the power management and control functions contained in Wyatt and Rizzo's current chip, but would also boast a neuromorphic processor and a microelectrode matrix. The matrix should have the highest resolution available at the time; for the tasks sketched above, an 8-by-8 matrix might already be a big help.

The cortical device promised by Normann and his team would send the return signal from the belt pack to a miniature connector behind the wearer's ear. From there it would run under the skin and through an entry port in the skull, to end at a demultiplexer, a neuromorphic chip, and an electrode array. A more advanced, and perhaps more adaptable, external neuromorphic chip would attempt to extend the retinal processing in view of research into signal processing in the lateral geniculate nucleus. This would be tentative at best, and this prosthesis might well be less "natural" than the retinal prosthesis.

To reiterate, the details of inner retinal processing (edge enhancement and temporal and spatial scaling) have not been completely unraveled. However, incorporating such processing into the retinomorphic chip may not be necessary and indeed may overshoot its objective: it is unclear which cell layers in the retina survive in various diseases, and which of the surviving cell types has the lowest threshold for electrical stimulation.

In fact, anatomical evidence, modeling, and preliminary experiments with blind patients performed at the Wilmer Eye Institute indicate that in patients with retinitis pigmentosa, a good percentage of not only ganglion cells survives, but probably also other inner retinal cells, so that the inner retinal processing may be done by the retina itself. This inference gives the retinal approach a leg up over the cortical approach—bearing in mind that the retinal approach can only work in the presence of a healthy optic nerve.

Toward the next generation

Once the first prototype has been implanted and the concept has been proven, development will become much more complex. Such diverse entities as government regulatory agencies, insurance companies, organizations for the blind, and venture capitalists will want to get in on the act. In addition, the number of avenues to be explored for improvement of implant survival, electrode density, threshold stability, resolution, and many other aspects will increase tremendously. A few reasonable expectations can be stated, however.

Vision enhancement hardware and algorithms will become more and more powerful, and by 2020 should be able to fulfill most of today's desires for visual image acquisition, preprocessing, and display. The limitation is and will be the level of vision left in any one patient: until experimentation and testing, it is unknown how much additional information bandwidth is achievable with an impaired human visual system through innovative presentation and training techniques.

The simulation of retinal preprocessing in hardware and software will probably become highly sophisticated. Moreover, it is possible that by 2020, working models of the lateral geniculate nucleus and primary visual cortex, including their interactions, will also be available on chip. If so, it would seem that the principal beneficiaries will be the designers of human-like machine vision systems, since it may be impossible to provide for elaborate feedback loops at the interface between such a chip and the cortical substrate.

The attainment of very high electrode densities, either in the retina or in the cortex, seems remote. A highly successful prosthesis for the peripheral retina and a low-density macular prosthesis may be possible, however, with obvious benefits for patients suffering from retinitis pigmentosa. For a normal-density foveal prosthesis, a cortical approach may be more successful at achieving high electrode densities: a much larger area for stimulation than in the retina is available.

A hybrid form of prosthesis is another possibility. Here a retinal prosthesis for the periphery of the visual field would be combined with a cortical one for the central field of vision mapped by the fovea. Other alternatives for the more distant future could be artificial neurons. These would interface with higher-order neurons, whether in the retina, the lateral geniculate nucleus, or the visual cortex, and could open new avenues to nerve-cell stimulation or even to the regrowth of retinal nerve fibers damaged by glaucoma or trauma, allowing true functional recovery. Initial experiments in retinal cell transplantation, currently with natural cells, suggest that at least partial recovery of visual function in this manner could become a reality over the next decade or two.

Throughout this development process there will be cross-pollination of ideas and techniques, not only among the three tracks indicated here, but also with other areas of artificial vision, imaging, and biomedicine. For too long, researchers of varied plumage have lacked a common language and a platform on which to exchange ideas. Our hope is that efforts such as the publication of this special report on bioelectronic vision will stimulate this exchange. ◆

About the authors

Gislin Dagnelie is assistant professor of ophthalmology and director of clinical low-vision services at the Johns Hopkins University School of Medicine, Baltimore, Md. He studies changes in temporal processing brought on by retinal degeneration, and is involved in developing the low vision enhancement system (LVES) and in retinal electrical stimulation research in blind volunteers. His e-mail address is gislin@lions.med.jhu.edu.

Robert W. Massof is professor of ophthalmology and neuroscience and director of the Lions Vision Research and Rehabilitation Center at the Wilmer Eye Institute of the Johns Hopkins University School of Medicine. He is lead inventor of the LVES effort. His e-mail address is bob@lions.med.jhu.edu.

Spectrum editor: Robert Braham