Optical Neural Computers

Can computers be built to solve problems, such as recognizing patterns, that entail memorizing all possible solutions? The key may be to arrange optical elements in the same way as neurons are arranged in the brain

by Yaser S. Abu-Mostafa and Demetri Psaltis

Computer scientists find it increasingly frustrating to see how casually a three-year-old picks out a tree in a picture. Sophisticated programs running on the most powerful supercomputers are capable of only a mediocre performance in doing what essentially amounts to the same task: pattern recognition. What makes this state of affairs so paradoxical is the fact that solutions to many problems that overtax the human brain can be arrived at quickly by computers. Indeed, a simple pocket calculator can easily outperform the human brain in such tasks as finding the product of two 10-digit numbers. What is the difference between the multiplication of numbers and the recognition of objects that makes the latter so much tougher to achieve in computers? In other words, why is it so difficult to make a computer recognize a tree?

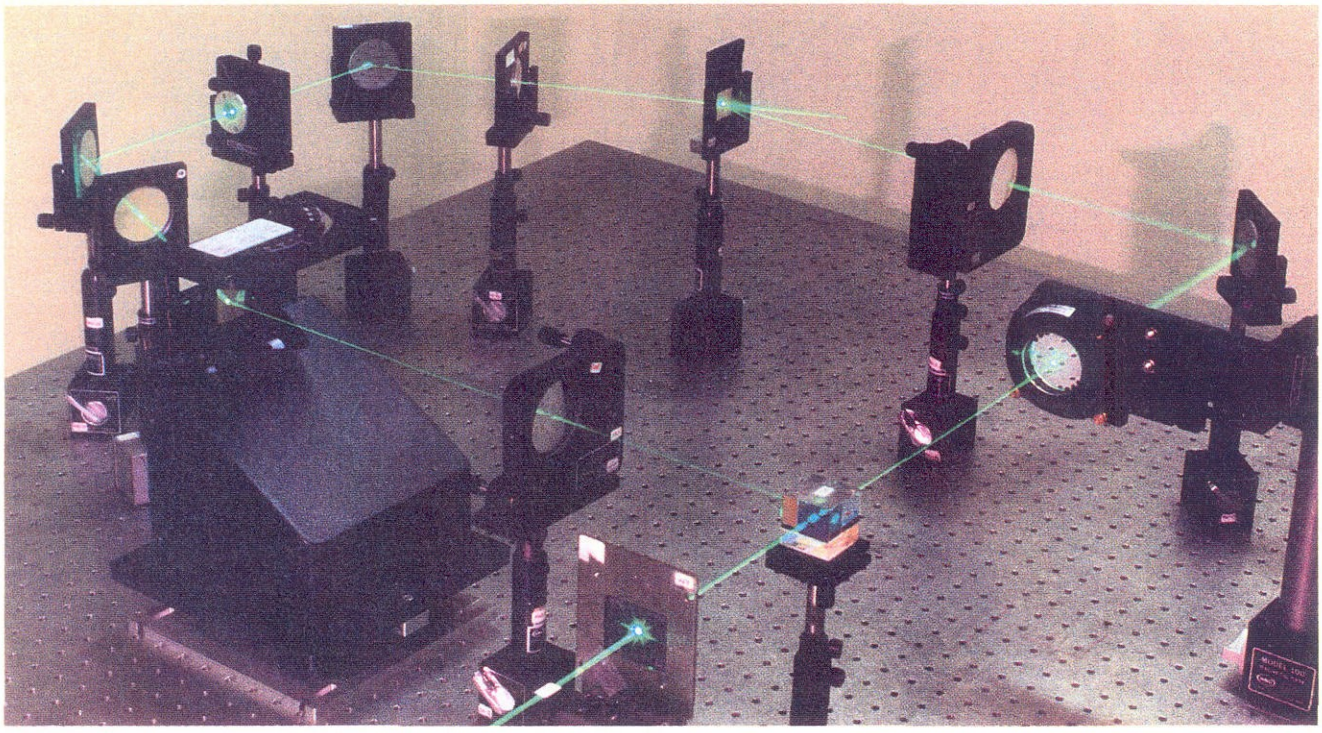
The answers to these questions ultimately hinge on the fact that pattern-recognition problems cannot be compactly defined. In order to recognize trees a comprehensive definition of a tree is required, and such a definition would be tantamount to a description of every conceivable variant. Problems such as those posed by pattern-recognition tasks constitute a subset of what we call random problems: problems whose solution requires knowledge of essentially every possible state of a system. Solving a random problem therefore entails memorizing the set of all possible solutions and quickly selecting the best solution from the set, given the input data. In contrast, the solution to such a classical computation problem as multiplication can typically be expressed succinctly in terms of an algorithm: a sequence of precise instructions specifying how the input data are to be manipulated to arrive at the solution.

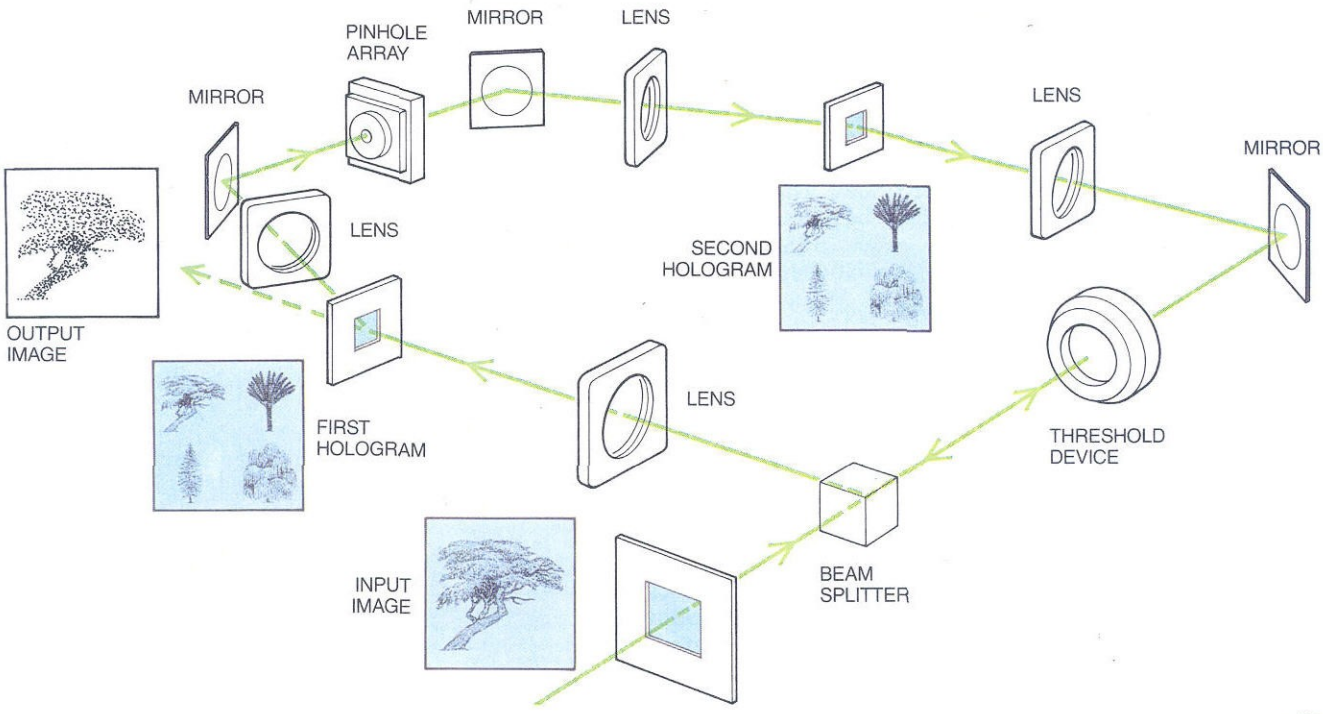
A conventional computer is adept at mechanically executing the instructions in an algorithm, but it cannot match the memorization and recollection capability of the human brain, which regularly and effortlessly conquers pattern-recognition problems. Because the brain is unique in its capability to solve random problems, many computer scientists and mathematicians have taken a closer look at how the brain works in the hope that the principles of its operation can be fruitfully applied in designing machines capable of solving random problems. Devices designed to model the workings of the human brain by emulating its anatomic structure are called neural computers; like the brain, they would consist of a large number of simple processors that are extensively interconnected. In this respect one technology stands out as being particularly promising for constructing neural computers: optics.

Optical technology dovetails nicely with the notion of a neural computer because the technology's strengths lie in exactly those areas that distinguish a neural computer, such as the interconnection of a large number of processing elements; its weaknesses lie in areas that are less critical for the functioning of a neural computer, such as the ability to perform intricate logic operations at the processor level. Whereas semiconductor technology in conventional computers has proved to be capable of tackling classical computation problems by means of algorithms, optical technology in the neural computers we envision may one day make it possible to solve random problems efficiently. Indeed, in our laboratory at the California Institute of Technology we and our colleagues have already built experimental pattern-recognition systems that represent a first step toward an optical neural computer.

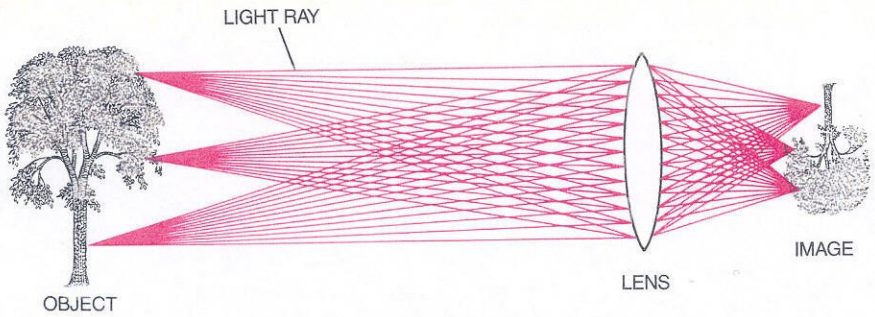
Regardless of the technology a computer incorporates (be it optical or electronic) or the functions it executes (be they multiplication or pattern recognition), two principal activities take place in it as it solves a problem: logic operations and data transmission. Viewing computation in such fundamental terms helps to get to the source of a particular computer technology's strengths and weaknesses. Semiconductor technology can be applied to build sophisticated logic circuits from electronic switches of very small size that have very reliable characteristics, yet such integrated circuits are rather limited in the amount of data that can be transmitted among the circuit elements. The reason is that on a silicon chip communication links consist of wires that must be kept separated by at least a minimum critical distance; otherwise the electrical signals they carry interfere with one another. This practical restriction places an effective limit on the number of wires that can be placed on a chip and hence on the amount of data communication that can take place on the chip.

Is there another technology from which computers could be built that does not suffer from this limitation in data communication? The operation of the eye's lens suggests one. The lens takes light from each of millions of points in the entrance pupil of the lens and redistributes it to millions of sensors in the retina. It is in this sense that the lens can be thought of as a highly capable interconnection device: light from every point at the pupil is "connected" to every point in the image focused on the retina. Moreover, multiple beams of light can pass through lenses or prisms and still remain separate. Indeed, two beams of light, unlike a pair of current-carrying wires, can cross without affecting each other. It is the ability to establish an extensive communication network among processing elements that primarily distinguishes optical technology from semiconductor technology in its application to computation.





PATTERN-RECOGNITION SYSTEM (*top*) developed by the authors and their colleagues at the California Institute of Technology can quickly find the best match between an input image and a set of holographic images that represents its "memory." The input image is projected into the system (*diagram*) through a beam splitter—a partially reflecting mirror—by illuminating a transparency (in this case one carrying an image of a cypress tree) with a laser beam (*bottom left*). The light that passes through the beam splitter hits the front of a threshold device, reflects off it and retraces its path back to the beam splitter, where it is reflected at an angle to initiate an optical "loop." A lens focuses the input image on a hologram, where the image interacts with each of four holographically stored images (here, of trees), creating patterns of light whose brightness varies according to how well the input and stored images match. A lens and a mirror direct the light issuing from the hologram to a pinhole array that spatially separates the four light patterns associated with each combination of input image and stored image. Another lens and mirror collimate the light and illuminate a second hologram with it. This hologram contains the same set of stored images as the first and is designed to produce a superposition of the four image combinations. The beam bearing the superposed images is focused by a third lens-and-mirror pair on the back of the threshold device. The pattern of light impinging on the back of the threshold device determines what light is reflected off its front. Since the brightest image reaching the back of the threshold device represents the best match to the input image from the set of stored images, it is essentially an image of the best match that is reflected from the front of the device into the optical loop for a second pass. Successive passes around the loop continue to enhance the best match from the set of stored images, which can ultimately be retrieved as the output image leaving the system through the first hologram.



OPTICAL LENS is an inherently powerful interconnection device: it connects every light ray that originates at a point on an object and passes through the lens with every point of the object's image. Unlike wires on an integrated-circuit chip, light rays can come close to one another and even cross without affecting one another. Hence millions of light rays could conceivably carry data simultaneously into a processing device, whereas electronic devices on a chip are limited to accepting input from a few wires at a time.

Because optical processing elements communicate through beams of light, they can be hooked up to one another without attaching a cumbersome wire between each pair of elements, and they need not be confined to the restrictive planar configurations of silicon chips. Indeed, optical connections are being considered as a means of relieving communication bottlenecks encountered in very-large-scale-integration chips. In such a hybrid optoelectronic system the processing units are electronic but the connections between them are optical, typically consisting of light sources and light detectors fabricated on the same chip as the processing units.

The most promising device for establishing arbitrary optical connections is not a lens but a hologram. Holograms are best known as a means of generating three-dimensional images, but more generally they represent an effective technique for recording and reconstructing the intensity of a light ray as well as the direction from which it came. Whereas a conventional lens maps each light ray entering the lens to a particular point on the image plane, holograms can readily be "programmed" to allow a variety of such mappings.

A planar hologram, produced on a relatively thin medium such as photographic film, can direct any light beam on one side of it to any point on the other side, provided the total number of points and light beams does not exceed the number of resolvable spots on the film. The number of resolvable spots in a one-inch-square hologram can be as high as 100 million. This would allow each of 10,000 light sources to be fully interconnected with each of 10,000 light sensors. A similar interconnection scheme by means of wires would be extremely difficult to accomplish on a silicon chip.

Even more prodigious in its capability to connect light emitters to light detectors is a volume hologram made from a photorefractive crystal. When such a crystal is exposed to light, electric charges are generated in it that redistribute themselves according to the pattern of the illumination's intensity. Because the local charge density in a photorefractive crystal determines the local refractive index (a measure of how fast light travels through the material), holographic images projected onto the crystal are recorded in terms of the spatially varying refractive index. The image information can then be extracted from the hologram simply by illuminating the crystal with a light beam.

Other hardware associated with traditional computation can also be realized optically, namely switching elements (from which processors are constructed) and memory elements (in which data are stored). Switching elements can be made from a nonlinear optical material. An optical material is nonlinear if its transmittance properties, such as its opaqueness or its refractive index, change as the brightness of the light incident on the medium changes. Gallium arsenide is an example of a nonlinear optical material from which two-dimensional arrays of optical switches have been fabricated. Nonlinear optical materials make possible the construction of an "optical transistor," in which the brightness of one light beam controls the transmission of another light beam.

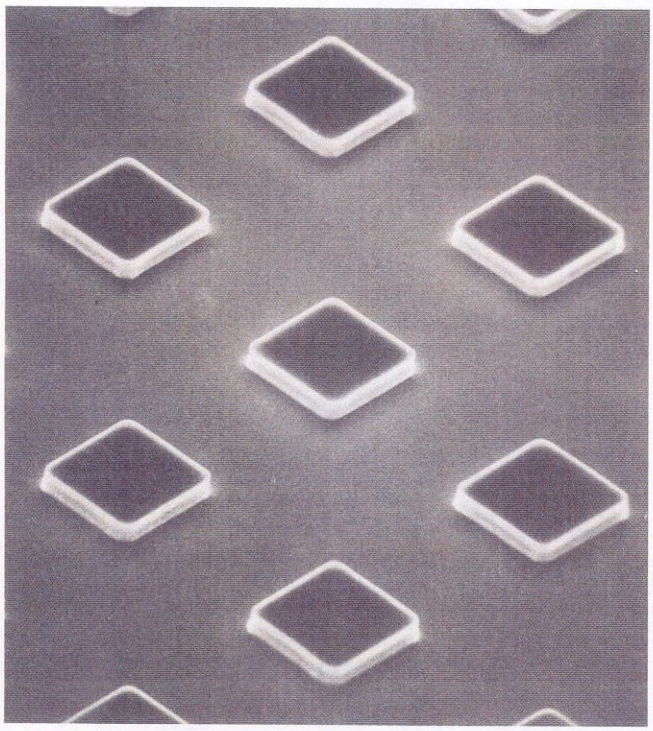
An optical memory element is essentially a device that can alter an input light beam into one of two possible states, each state corresponding to a bit of binary data (either a 1 or a 0). Optical memories have been developed for audio and video recording and more recently as digital mass memories for electronic computers. Yet in these devices the stored information is typically accessed serially by focusing a light beam on one stored bit of information at a time, much as information is read off a magnetic tape. These devices do not exploit the huge potential for increasing the speed with which data can be transferred from memory by allowing parallel access to stored data. Millions of bits of information could be read out and transferred at the same time merely by shining an unfocused light beam on a suitably designed optical memory device [*see illustration on page 92*].

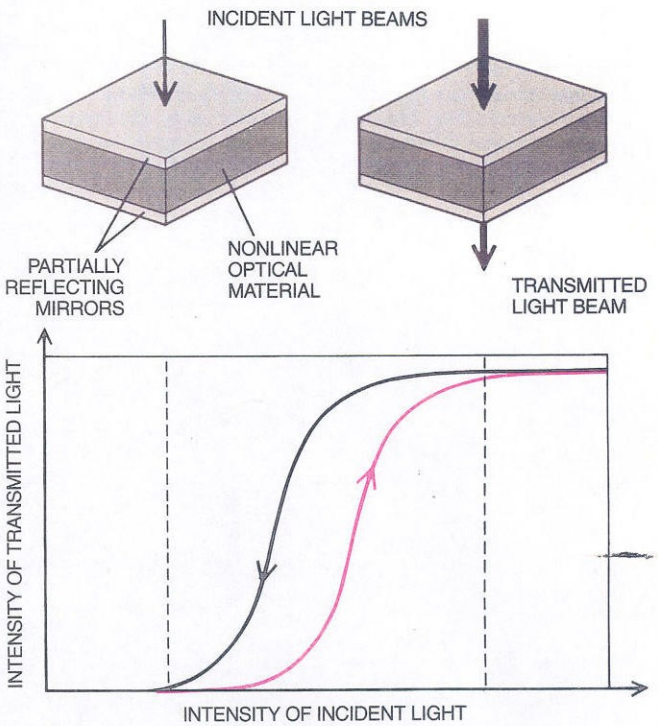
The fact that designers of optical memories have not exploited the potential for parallel access to data is an indication that most of the work in the development of optical switching and memory elements is done with the goal of implementing these devices in the execution of sequential, binary-logic functions. Hence these optical components would essentially duplicate (albeit perhaps more efficiently) the same operations that take place in conventional electronic computers. Although increased switching speed and a massive memory may ultimately result from such development efforts, they do not fundamentally change the mode of computation of conventional computers. Consequently devices in which electronic switches and memory elements have simply been replaced by optical analogues are just as likely as current computers to falter when they are confronted with pattern-recognition problems.

In order to understand why this is so it is necessary to consider how a conventional electronic computer solves a problem. As we have indicated above, the classical theory of computation, from which current computers were developed, is built around the notion of algorithms. The procedure for long division of two numbers is a good example of an algorithm. The procedure can be specified easily, and once it has been mastered—whether by a computer or by a sixth grader—it is universally applicable: it works as well for dividing a four-digit number by a three-digit one as it does for dividing a 1,000-digit number by a 900-digit one (although the algorithm may take longer to complete in the latter case, particularly for the grade schooler).

Computational problems that lend themselves to algorithmic solutions share a characteristic property: they are structured, meaning they can be stated clearly and concisely in mathematical terms. Most of the problems currently being solved with computers belong to this class of structured problems, and it is now a universal practice for computer programmers to look for

VOLUME HOLOGRAM (*on middle stand*), like the more familiar planar hologram recorded on photographic film, can distribute beams of laser light in "programmable" directions. Here, for example, two input laser beams coming from the bottom right are turned into four output beams heading toward the top left. A volume hologram is made from a photorefractive crystal. When such a crystal is exposed to light, electric charges are generated in it that redistribute themselves according to the pattern of the illumination's intensity. Because the local charge density in a photorefractive crystal determines the local refractive index (a measure of how fast light travels through the material), the crystal can record holographic images in terms of the spatially varying refractive index. Such a hologram can set up a pattern of optical connections between light sources and detectors for each image stored in it.





OPTICAL SWITCHING ELEMENTS (*left*) are manufactured by sandwiching a nonlinear optical material (which alters its refractive index according to the intensity of the light to which it is exposed) between two partially reflecting mirrors. An element thus constructed (*top right*) can abruptly change its transmission properties depending on the intensity of the incident light beam. It also exhibits a so-called hysteresis cycle (*bottom right*) when it is switching. As the intensity of an incident light beam is gradually increased from zero (*color*), the element does not allow any light through until the incident beam reaches a certain threshold intensity; then transmission quickly rises to a maximum value. The intensity of the transmitted beam will not retrace the same path if the intensity of the incident beam is reduced back to zero. Instead the abrupt change in transmission (*black*) now occurs at a lower incident-beam intensity. The photomicrograph was provided by Thirumalai Venkatesan of Bell Communication Research, Inc.

an algorithm whenever they have to solve a problem.

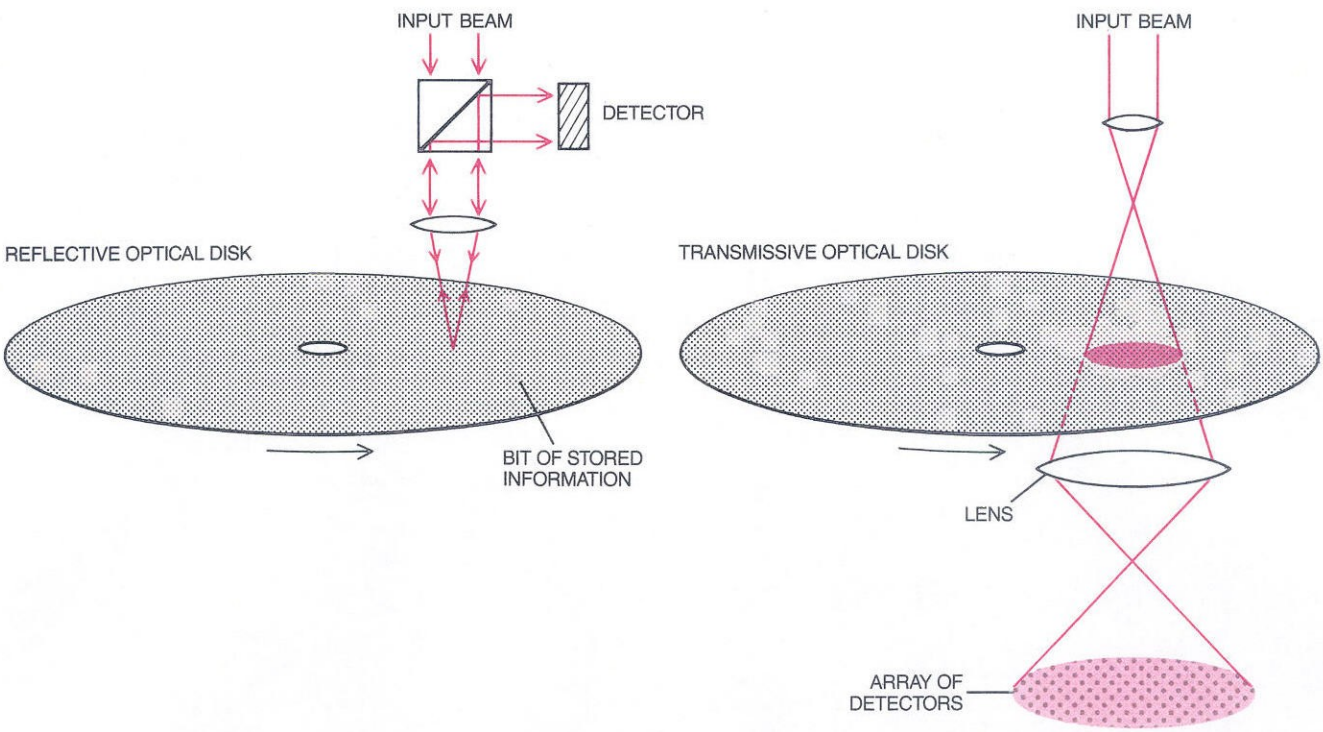
Problems such as pattern recognition in natural environments, however, lack the structure that would allow simple algorithmic solutions. It is this departure from the properties of structured problems and the methods for solving them that characterizes a random problem. The term "random," as we apply it here, is derived from the mathematical concept of randomness, namely the lack of a concise and complete definition. Randomness in this sense is linked to the mathematical notion of entropy, which can be thought of as the amount of disorder in a problem or, equivalently, the amount of information needed to define the problem. Because a formal description of a random problem would amount to a listing of essentially every possible solution to the problem, random problems have a much higher degree of entropy than structured problems.

To better understand what it means for a problem to be random, consider once again our tree-recognition example. Although it is clear to most people what a tree is, it would be very difficult to write down a concise definition for a visitor from another planet, who does not know what "branches" or "leaves" are, or for that matter what the color "green" is. Even if examples of each of these features could be presented to the alien visitor, there are innumerable types of branches, leaves and hues of green; a handful of examples is unlikely to be enough to cover all possible arboreal combinations.

Among the inhabitants of the earth a universal understanding of what is meant by the term "tree" arises from a vast accumulation of common experience. A computer, like an extraterrestrial visitor, cannot draw on this reservoir of common experience; everything must be spelled out for it precisely and unambiguously. Although many properties of trees and other visual scenes have a fair amount of regularity, there is a major component of irregularity that does not fit any simple mathematical or algorithmic model. Any generalized definition that is grounded on an underlying regularity among trees runs the risk of subsuming objects that are not trees. Indeed, the only definition that would not assume any prior knowledge of trees and that would include all trees and exclude all other objects would amount to a description of all types of trees. It is important to recognize that this difficulty is inherent in random problems; it is not just a symptom of fuzzy thinking by human programmers or a poor choice of descriptions.

A simple algorithm will therefore never serve to solve a random problem, because an algorithm for the solution of a random problem would be tantamount to a definition of the problem, and hence it would have to contain all the problem's many possible solutions. For example, an algorithm to identify fingerprints would have to amount to a list of all possible fingerprints, but there is no way to pack such a list in a few lines of computer code. Fingerprints must ultimately be classified into a large number of basically unrelated types—each of which must be considered in order to identify a given print. The solution to random problems therefore lies essentially in memorizing all possible solutions.

Optical technology offers a potentially massive memory, but this alone does not suffice for a practical system to solve random problems. It would be pointless to store the vast data base of a random problem optically, only to search through it sequentially whenever a solution that fits the input data is needed; it would take a prohibitive amount of time. Moreover, the input data as well as the stored information are likely to be incomplete or inaccurate, precluding an exact match between them. The key additional ingredient for a practical system that solves random problems is

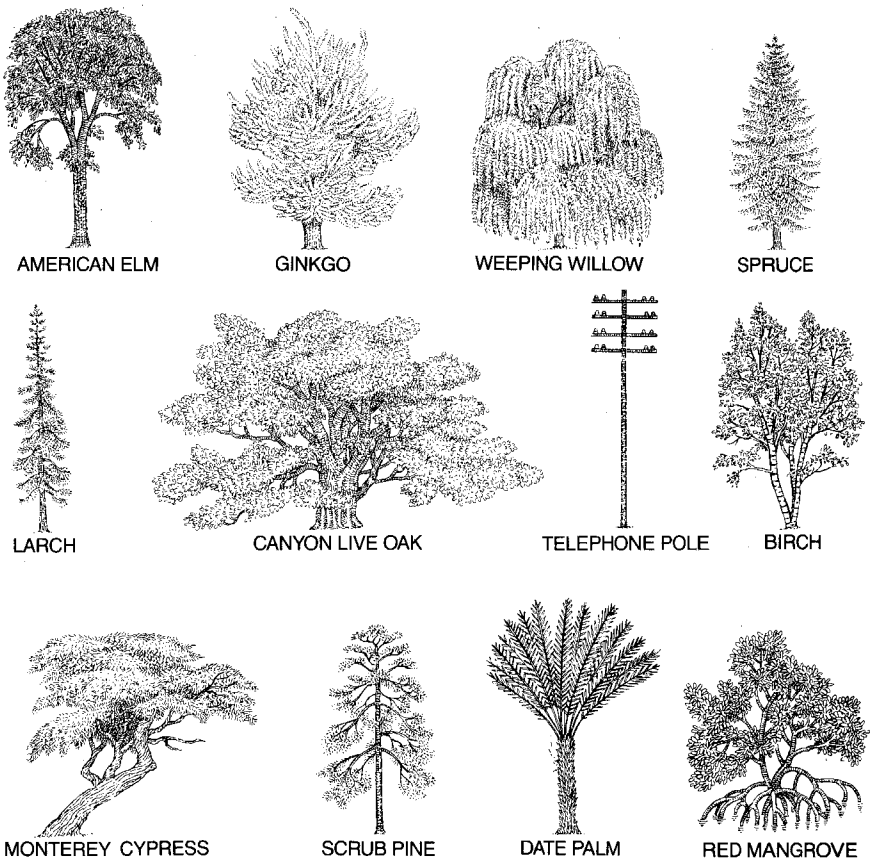


OPTICAL MEMORY DEVICES can be made from disks containing embedded spots that modulate light into two possible states. The states correspond to the value of a stored bit of binary data (either a 1 or a 0). In most current designs (*left*) the stored information is accessed serially by focusing a light beam on each data-storing spot and detecting the reflected signal. A similar device (*right*) with an unfocused beam and a transmissive disk could greatly increase the speed with which stored data is accessed: it would scan millions of spots at a time, simultaneously reading out and transferring the data to an array of detector elements.

a way to associate input data directly with the stored information without requiring an exact match.

Such a process of association is a major feature of biological memory, where partial features of an object trigger the retrieval of complete information about the object. Consider the train of associated reminiscences that courses through one's mind when one sees a familiar face: the person's name, one's general disposition toward him or her and perhaps the smell of his cologne or her perfume—to name a few. Similarly, human beings do not consciously follow an explicit step-by-step algorithm to recognize visual scenes; rather, they follow an unconscious process of association. Even in the case of highly structured problems, such as chess playing, experts develop skills that are associative in nature. (In fact, it is the inability of expert chess players to record explicitly the "algorithm" by which they made a brilliant move that so far has prevented the writing of a chess-playing program capable of beating world-class players.)

Can the anatomical structure of the brain provide an organizational principle by which associations can be readily established between what is stored in memory and the input data? Moreover, can such a model be implemented by taking advantage of the intrinsic strengths of optical technology?



ARE ALL THESE OBJECTS TREES? Even a young child can answer correctly; a conventional computer, however, has enormous difficulty in doing so. Although there is a fair amount of regularity among the trees shown (each has a trunk and branches, for example), there is also a major component of arboreal irregularity among them. A generalized definition of a tree based on the underlying regularity could lead to erroneous identifications (such as mistaking a telephone pole, which has a "trunk" and "branches," for a tree). Hence any effective program designed to recognize trees would essentially have to be a list of all types of trees, which cannot be done in a few lines of computer code.

The brain consists of a very large number of neurons each of which is directly connected to a large number of other neurons. A neuron can be in one of two states (known as "firing" or "not firing") and is able to sense the states of its neighbors through its connections. During the course of cerebral "computation" each neuron independently examines the states of its neighbors and, based on the information, determines its own future state. Such a network of neurons is robust; if some neurons malfunction, the overall function of the network is not affected. (Indeed, neurons in the brain are continually dying off, and yet thought and memory are not appreciably hampered.) Computation in neural networks is done in a collective manner: the simple, simultaneous operation of individual neurons results in the sophisticated function of the neural network as a whole.

This form of organization enables thousands of neurons to collectively and simultaneously influence the state of an individual neuron according to the application of simple rules. More important, it also allows information to be encoded in the neural connections rather than in separate memory elements. Each distinct piece of stored information can be represented by a unique pattern of connections among neurons.

Computers whose processing elements are arranged in much the same way as neurons are arranged in the brain would exhibit several features making them remarkably suitable for the solution of random problems. For one thing, such neural computers would be versatile, since the connections between the elements (of which there are a huge number) serve as the programmable storage mechanisms that uniquely "tune" the computer's memory to a given problem. Essentially the connections in a neural computer could be reconfigured in a great many ways to make possible the storage of a random problem's many possible solutions.

Another major feature of the operation of neural computers is spontaneous learning. Imagine what it would be like if children had to be taught how to speak as they are taught how to carry out long division, that is, by teaching them a set of specific rules! Fortunately this is not necessary in most instances, since a child spontaneously associates spoken language with an experience. Learning to talk therefore begins as a process of mimicking the words heard in association with a particular experience. In this simple way the child starts to produce recognizable and sensible patterns of speech.

Similarly, the programmer of a neural computer does not have to understand in a formal, mathematical sense the problem for which he or she is programming. The programmer only has to provide enough "training" data (consisting of possible solutions) to the computer and allow it to set up a unique pattern of connections for each solution. In other words, it is possible for a neural computer to program itself. For example, if one wanted to program a neural computer to recognize different trees, one would provide images of trees as training, allowing a specific pattern of interconnections among the computer's processing el-

ements to be "imprinted" for each training image.

A neural computer built along these lines from optical elements consists of two main components. The first component is a two-dimensional array of optical switching elements to simulate neurons; the elements switch states depending on the states of the elements to which they are connected. Each element in the planar array can be interconnected to all other neurons by light beams. The second component is a hologram that specifies the interconnections among the elements. Since the connections constitute the memory, they must be modifiable—if different problems are to be tackled by a single optical neural computer.

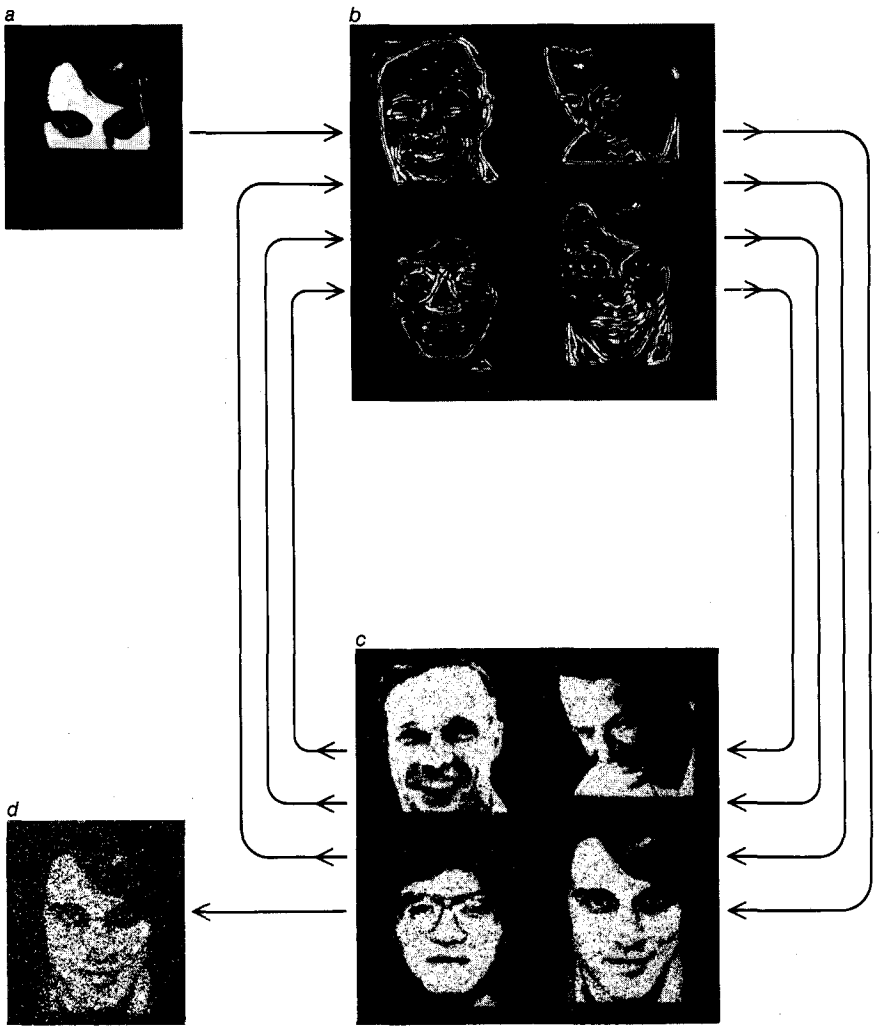
The array of switching elements can be made by well-established fabrication methods developed for semiconducting materials. Each element can be either a purely optical switch or an optoelectronic combination of light detector, electronic switch and light emitter. The total number of possible connections is the square of the number of elements. If a volume hologram is used to specify the interconnection scheme, the volume of the crystal must be proportional to the total number of connections.

A hologram whose volume is one cubic centimeter can in principle specify more than a trillion connections, which means it can handle all possible interconnection schemes of more than a million optical elements. The ability to store the interconnection information in the three dimensions of a volume hologram creates a huge potential memory for optical neural computers. In the case of holograms used for pattern-recognition systems, for example, the interconnection scheme can easily be set up by making a hologram of all the images that are to be identified.

Several experiments are under way in our laboratory at Caltech to develop such optical neural computers. In one experimental setup [*see illustration on page 89*] the action of a two-dimensional array of more than 10,000 neurons is simulated by a threshold device consisting of 10,000 tiny elements that switch the reflectivity of their front surface whenever the intensity of a light beam hitting their back surface is greater than a certain threshold. In this sense the threshold elements act as neurons because they switch states depending on whether enough light reaches them from behind. A pair of planar holograms, a system of lenses and mirrors and an array of pinholes specify how much light each threshold element gets, in essence establishing the interconnections among the elements. Both holograms contain the same set of images, although the edges of the images stored in one hologram have been enhanced. The system is arranged in the form of an optical "loop" so that the system has continuous feedback.

An image to be recognized is projected into the system by reflecting it off the front of the threshold device. Lenses, mirrors and the pinhole array enable the reflected input image to interact with all the images stored in the two holograms in such a way that the best match between the input image and the stored holographic images is the brightest image issuing from the second hologram. The light from the second hologram is then directed to the back surface of the threshold device, causing individual threshold elements to change their reflectivity so that an image of the best match is primarily what is reflected off the front of the device for a second pass around the loop. Successive passes around the loop continue to reinforce the best match until the system "locks" onto the correct stored pattern, which can be retrieved as the output image. In this way the system is capable of recognizing any one of the stored images—even if only part of the image is projected into the system.

We believe the best way to design computers that solve random problems is through the implementa-



ASSOCIATIVE MEMORY is a feature of the authors' optical pattern-recognition system (*see illustration on page 89*): the system can "recognize" an image even if only part of it is projected into the system. If only half (*a*) of one of the four faces that are respectively stored as edge-enhanced images (*b*) and normal images (*c*) on a pair of holograms is projected into the system's self-reinforcing optical-feedback loop, the system nonetheless selects the correct whole image (*d*) as output. A similar process, in which one piece of information elicits the recollection of related stored information, is a feature of human memory and learning. Although only four images were recorded holographically on film in this example, a volume holograph, which has an enormous memory capacity, could conceivably make possible the memorization and swift recognition of millions of images.

tion of a neural architecture. Optical technology can be applied amazingly well in building such a computer. A neural computer demands a very large number of switches, each of which must perform only a simple operation in order to switch between two states. Similarly, it is possible to place a large number of simple optical switching elements on a plane. A neural computer requires extensive connectivity and data communication. Similarly, holograms can establish the necessary connections between numerous optical elements. Because light rays can cross one another without interfering and are not limited to traveling along the two dimensions of the surface of a silicon chip, simultaneous communication among numerous optical elements is readily achieved.

Although optical systems are vulnerable to various inaccuracies and local errors, neural computers are inherently fault-tolerant: a perfect match between input and output is not necessary. A neural computer would be programmed by establishing a unique interconnection pattern for each solution, and this could easily be done by recording various training images on a photorefractive crystal or photographic hologram. In contrast to conventional computers, the speed of the individual switching elements is not critical for the function of a neural computer, since a few iterations usually suffice to complete the association function. This is particularly fortunate for optical neural computers, since each firing of a "neuron" consumes a fixed amount of energy and speeding up translates directly into more power consumption and hence into excessive heat production.

Clearly many challenges must be met before optical hardware, arranged in a neural architecture, can produce practical computers that are capable of dealing with random problems. Advances in optical materials and manufacturing technologies and in the understanding of the organization of large-scale neural computers are needed. Equally important is better understanding of the operation of neurons in the brain and of how they collectively "learn" and "classify" patterns.

Engineers, computer scientists and mathematicians have reached significant turning points in three seemingly unrelated areas: optical components, neural computers and random problems. There is good reason to believe that, with progress in each of these areas, their interaction will ultimately yield systems capable of pattern recognition and other artificial-intelligence tasks that may never be duplicated by purely electronic means.