Holography in artificial neural networks

Demetri Psaltis, David Brady, Xiang-Guang Gu & Steven Lin

The dense interconnections that characterize neural networks are most readily implemented using optical signal processing. Optoelectronic 'neurons' fabricated from semiconducting materials can be connected by holographic images recorded in photorefractive crystals. Processes such as learning can be demonstrated using holographic optical neural networks.

IN recent years there has been a marked resurgence of interest in artificial neural networks[1]. The structure and the principles of operation of such 'neural' computers are to a large extent biologically motivated. The reason for attempting to derive clues from neurobiology to build computers is the sharp distinction that exists between the types of problems for which modern digital computers are useful and the tasks at which humans and other animals excel. The classic example is the distinction between pattern recognition and arithmetic: humans do arithmetic poorly, but easily outperform computers in the simplest recognition tasks. Research in neural computing is based on the premise that this difference in capabilities is due to basic differences in the hardware and the ways in which the hardware is programmed. Although in some instances, particularly in the case of sensory systems that are relatively well understood, it may be possible to construct circuits that are reasonably accurate replicas of biological systems, more often one extracts only the basic properties evident in the nervous system and uses this information to guide the design of the computer. The hope is that if one is successful in identifying truly relevant properties, then the neural analogy should provide a valuable contribution to the design process even though the circuitry of the nervous system is not understood in detail.

The rapid progress in modern computer design is owed in part to the separation of algorithmic and architectural issues from the physics of the devices used to implement the algorithm; in this way, a programmer may develop algorithms without detailed knowledge of the physics of the semiconductor hardware. Unfortunately, this separation is difficult to maintain in highly parallel architectures, such as neural networks. One of the main issues in parallel computing is the relationship between the properties of a parallel architecture and the efficiency with which it solves problems. The physics of the hardware is a central consideration in this debate. The efficiency of communication between nodes of the network is a major limiting factor in hardware design. In parallel electronic circuitry most of the system area must be devoted to interconnecting wires[2]. This imposes limitations on the hardware architecture and therefore solutions have been sought that go beyond the traditional silicon technology. We will discuss one such solution, holographic optical interconnections[3], in the context of the most highly interconnected parallel architecture, artificial neural networks[4–11].

Basic principles

The principles that are most commonly used in the design of neural computers are as follows.

Massive parallelism. If we think of an individual neuron as a computational element, then the human brain consists of 10^{11}-10^{12} such elements. For the execution of a task that involves the visual pathway, a large portion of these units are active. This is in sharp contrast to a conventional digital computer, where perhaps a few thousand out of more than 10^8 transistors are active at any one time during a typical computation.

Dense interconnections. A neuron in the brain typically receives inputs from several hundred to several thousand other units to produce its own output, which is broadcast to roughly the same number of units. It is widely believed that the synaptic strengths serve as part of the memory of the system. This connectionist view has been adopted almost universally in the design of artificial neural networks. Connectionist networks achieve high computation speed through specialization by encoding general knowledge about the computational problem in the structure of the network.

Learning. The ability of the nervous system to perform desirable computations accurately is acquired through the genetic code and adaptation. The ability of a conventional digital computer to perform useful tasks derives from the design of its circuit architecture (the hardware) and its programming (the software). We can think of the hardware design as being analogous to the role of genetics and the software to adaptation. There is, however, a sharp distinction between programming and adaptation. A programmer explicitly designs every step of the computation as a sequence of precise mathematical statements. The neural approach to computation develops the required software for tackling a particular problem through a sequence of trials during which the parameters of the system are adapted to achieve a desired goal. The success of such learning procedures is critically dependent on the learning algorithm and the hardware architecture. The system designer must make both of these choices *a priori*. Even though these choices are typically very difficult to make, they are relatively few. In practical terms, for certain problems the learning approach can provide a method for transferring some of the burden of programming the computer from the user to the computer.

The architecture of optical neural networks

Figure 1*a* is a schematic diagram showing the various components of a pair of neurons and the way in which they are

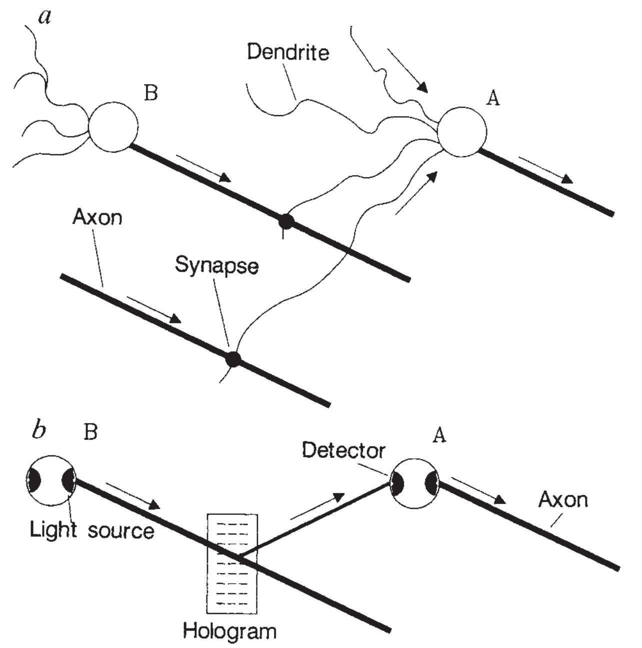

FIG. 1 Conceptual structure of *a*, a neural system and *b*, its optical analogue. See text for details.

connected. Each neuron receives inputs through the synapses on its dendrites, it processes these inputs in some fashion, and then it broadcasts the result on the axon where it is picked up by the dendrites of other neurons. This description is somewhat simplified (for example, dendrites can be two-way channels[12]), but it is a view that has been adopted almost universally in neural-network models. The diagram in Fig. 1*b* shows the holographic analogue of the two neurons in Fig. 1*a*. The output of each neuron is a light beam, and the activity of the neuron is coded in the amplitude or intensity of the optical signal. The input of each neuron is a light detector which senses the amount of light that is directed towards it. A holographic grating is placed in the path of the output beam of neuron B, which diffracts the incident light. The direction of the diffracted beam is determined by the period and the orientation of the grating. With an appropriate holographic grating, light from neuron B illuminates the detector of neuron A. In this way a signal is generated in A as a result of the activity in B, and we say that the hologram connects the two neurons. The strength of the connection can be modified by adjusting the modulation strength of the holographic grating. There are some direct analogies between the component structure of a neuron and its optical simulation. The output light beam has the role of the axon, broadcasting the signal from each neuron. The holographic grating plays the part of the synapse, directing the signal from one neuron to the next, and the optical pathways along which light is transferred from the hologram to the detector area of the neuron are analogous to the dendrites. The device consisting of the optical beam generator, the detector and the circuits that process the detected signal is reminiscent of the soma of the neuron. In real neurons some processing tasks can take place on the branches of the dendritic tree, but in the optical simulation all integration and computation tasks are concentrated on this integrated unit, which we refer to from now on as the 'neuron'.

Having established the basic structure of the optical analogue of a single neuron, the next issue is the way in which we may connect many such units together to form optical neural networks. Suppose that the third neuron is to be connected to neuron A as well. This can be optically simulated by superimposing a second holographic grating in the same crystal that was used to store the interconnection pattern between the first two neurons. This second grating has a distinct spatial orientation and period and it is tuned to redirect light from the third neuron onto the same input (the detector) of neuron A. This is a second difference between the real neuron and its holographic realization: the synapses are not localized at the intersection between axon and dendrite, but are implemented instead in a distributed manner, each sharing the entire volume of the holographic medium.

Distributed connections have advantages and disadvantages—the disadvantage compared to the localized implementation is the reduced control of individual synapses. The adjustment of the strength of one synapse may inadvertently affect other synapses as well. Accommodating this limitation is perhaps the most crucial research issue in this field and we discuss this further below. The advantages of the distributed holographic synapses are high storage density and ease of fabrication. The number of distinct synapses that can be packed in a hologram of volume V is V/λ^3, where λ is the wavelength of the light[13]. This corresponds to 10^{12} synapses per cm^3 for the typical operating wavelength $\lambda \approx 1\ \mu m$. There are other factors, primarily relating to the physical properties of the material used, that can prevent us from realizing this upper limit. In practice it is possible to realize 10^9–10^{10} interconnections per cm^3. The distributed holographic storage of the weights is also the ultimate in simplicity in terms of device fabrication: it involves only crystal growth.

Although holograms can simulate the function of the synapses, the axons and the structure of the dendritic tree, optoelectronics are used to perform the nonlinear computations of the neuron and to provide the optical energy needed to incorporate such devices into large networks. Typically, the neuron has a 'soft threshold', which is to say that the relationship between the summed post-synaptic signals into a particular unit and the output of this unit is an S-shaped curve. To implement this nonlinearity we must provide a mechanism by which the incoming signals can interact. Optical signals do not interact in free space, so a material is introduced at the location of the neuron that senses the presence of the light and, in response, changes its electronic state. This modifies the optical properties of the material and an optical beam travelling through this medium can therefore be affected by the presence of another beam. Various potentially appropriate devices based on optical nonlinearities are under development[14-16], but at present the most effective method for achieving such nonlinear optical effects involves detection of the optical signals, followed by nonlinear processing of the electronic signals and subsequent regeneration of optical signals through a light modulator or source. Semiconductor materials are particularly well suited for this purpose because it is possible to fabricate monolithic devices that incorporate all three functions. The unit shown in Fig. 2 is an example of a 10×10 array of optoelectronic neurons fabricated in gallium arsenide[17]. Each unit consists of a pair of transistors, a photodetector and a light-emitting diode (LED). An incident signal above threshold is sensed by the photodetector amplified by the transistors and rebroadcast by the LED. It is possible to construct devices with $\sim 10^4$ neurons per cm^2 that can be switched in parallel at rates of over 10,000 Hz. Such a capability, although not readily available at present, can be obtained by rather straightforward advances.

The common feature of the various types of optical and optoelectronic neuron arrays is planar layout. This is a consequence of the techniques used in fabrication, namely epitaxial growth or deposition of a sequence of thin layers on a two-dimensional substrate. As a result, we obtain two-dimensional arrays of neurons or neural planes. This technological constraint leads to an optical neural network that consists of planes of neurons separated by a space in which reflection holograms specify the connection between units in a single layer and transmission holograms interconnect neurons in two different layers. Although this structure of layers of neurons separated by interconnections emerges from technological constraints of the optical implementation, it is nevertheless similar to the structures used in almost all neural-network models.

The three-dimensional structure of the optical architecture has a potential advantage over electronic implementations. In electronics, the neurons (simulated by transistors) and the interconnections (implemented with wires, resistors and transistors) are both fabricated on the same planar surface[18]. Typically, the area devoted to the interconnections is the largest fraction of the available area and this severely limits the size of the network that can fit on a single chip. In the optical implementation, the active planes can be populated with neurons alone because the interconnections are realized holographically using the third dimension. As a result, the number of neurons that can fit on a plane of a given area is at least C times larger for the optical implementation, where C is the average number of connections per neuron. To make the comparison another way, the electronic implementation requires at least C times more chips to form a network with a given number of neurons. C can be well over 100, so optical connections provide a large savings in the number of semiconductor devices needed. The relative disadvantage of optics is that, with only a few exceptions, the technology is at an early stage. Two basic components are needed: the neural planes and the holographic interconnections. We have discussed briefly how optoelectronics can provide the answer for fabricating the neurons and hereafter we focus on the holographic interconnections.

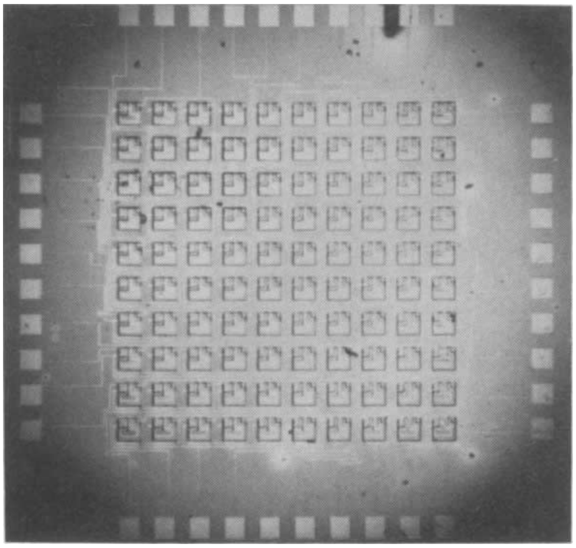

FIG. 2 Photograph of an optoelectronic implementation of a 10×10 array of neurons in gallium arsenide. The area of the array is 5×5 mm.

Holographic interconnections

Figure 3 shows a particular realization of a holographic neural-network architecture[10]. Each resolution element (or pixel) at the first neural plane can be a location for a neuron. The light from a pixel is collimated and diffracted by a holographic grating. The diffracted light is focused by a lens onto a pixel at the output neural plane. We now discuss how, given the location of two neurons at the input and output planes, the appropriate grating can be formed. As shown in Fig. 3, the portion of the light from the input neuron that is not diffracted by the hologram is reflected by a phase-conjugate mirror (PCM). A PCM, unlike a conventional mirror, reflects a collimated beam back along its original path. A PCM is a nonlinear crystal that forms a hologram of the incident beam. The recorded hologram is then illuminated from the opposite direction, resulting in a reconstructed beam that precisely retraces the path of the beam incident on the PCM. Light from the output neuron propagates towards the left in Fig. 3 and is collimated by the lens. This beam and the beam that is reflected by the PCM form a sinusoidal interference pattern at the hologram. The interference pattern exposes the hologram and a sinusoidal grating is recorded. This arrangement ensures that the grating will interconnect the input and output pixels or neurons that were used to record it. The strength of the connection is controlled by varying the exposure during recording. The connection is stronger when the light is brighter or the exposure time longer. If x_i and x_j are the amplitudes of the light from the input and output neurons, respectively, then the change in the strength or the weight of the connection formed by the grating is $\Delta\omega_{ij} \propto x_i x_j$. x_i and x_j can be, in general, bipolar signals. When the sign of x_i matches the sign of x_j then ω_{ij} increases; otherwise it decreases. Alternatively, the system can be designed with x_i and x_j being positive only. In this case $\Delta\omega_{ij}$ can be only positive, and separate excitation and inhibition channels are constructed. Because the algorithm causes only positive changes in the weights, they will eventually exceed the operating range of the holographic medium and the medium will be saturated. This can be avoided by introducing a weight decay mechanism. Both modes of operation are possible, but here we consider only the case of bipolar neurons.

We now consider connections between multiple neurons at the input and output planes. If N is the number of pixels in one dimension for both input and output planes, there are N^2 pixels in each plane. If we use each of these pixels as a site for a neuron, and we want to interconnect independently each of the input neurons to all the output neurons, we need N^4 interconnections, which implies that N^4 weights or gratings must be stored in the hologram. This poses the question of the maximum number of gratings that can be distinguished in a hologram of volume V. Assuming that the resolution of the holographic medium in any direction is $\delta > \lambda$, the total number of resolution cells or distinct samples in the volume of the holographic crystal is V/δ^3. From the sampling theorem, the number of distinct sinusoids, or gratings, that can be recorded in this medium is equal to the number of samples, so the maximum number of independent interconnections that can be supported by a hologram of volume V is V/δ^3. This guides the design of specific architectures. As an example, we consider a network in which neurons in two adjacent layers are arbitrarily interconnected. The product of the number of units in the first plane, $N_1 \leq N^2$, and the number of units in the second plane, $N_2 \leq N^2$, cannot exceed the available holographic interconnections in the crystal:

$$N_1 N_2 \leq \frac{V}{\delta^3} \tag{1}$$

If this relationship is violated then an attempt to establish the interconnection between one pair of input and output neurons automatically specifies the interconnections for other pairs as well. To avoid this we can either increase the volume of the crystal or decrease the density with which the neural planes are

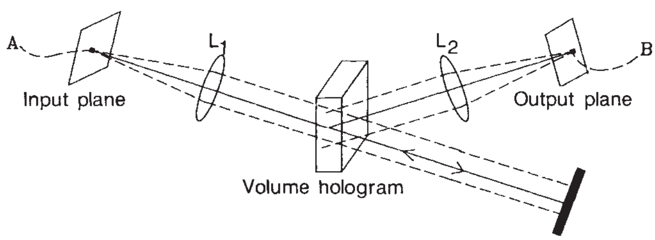

FIG. 3 Holographic neural-network architecture. The light amplitude at each pixel on the input plane corresponds to the output of a neuron. The signal detected at each pixel on the output plane corresponds to the summed inputs to a neuron. The phase-conjugating mirror reflects the input exactly back on itself. Light propagating back towards the input interferes with control signals generated at the output plane to adaptively update the interconnection pattern stored in the volume hologram.

populated (that is, decrease N_1 and N_2). Here we will use the second strategy. In particular, we assume that the hologram is a cube, with edges of length L, and $L/\delta = N$. In other words, the number of resolvable points in any one dimension is the same for both the neural planes and the hologram. This is a special case, but it is also the most sensible way to design such optical systems. In this canonical system, the total number of gratings available in the hologram is $V/\delta^3 = (L/\delta)^3 = N^3$. If the density of neurons remains constant from layer to layer ($N_1 = N_2$), then from equation (1) the maximum number of neurons per plane is $N^{3/2}$. The remaining task is the selection of the correct $N^{3/2}$ pixels from the N^2 available sites at each plane for the placement of neurons.

Figure 4 shows three examples of patterns that appropriately sample $N^{3/2}$ points out of N^2 for $N = 16$ (refs 19, 20). In each case the larger circles represent neuron locations on the regular two-dimensional grid of dots. The number of neurons in each case is $16^{3/2} = 64$. We derive these patterns using a process of elimination. Each time we attempt to add a new neuron to the input (output) sampling grid, we check to see whether this new neuron is already connected to one of the neurons selected previously at the output (input) by an existing grating. If it is, we eliminate the position of the new neuron from the sampling grid; if it is not, we select this position for a new neuron, which implies that gratings are established to connect this new neuron to all the neurons now selected at the output (input). In addition, we eliminate all the positions at the output (input) to which this new neuron is connected by existing gratings. We arrive at the complete sampling grids by systematically iterating this procedure. The patterns that result are not unique. Different patterns emerge by adopting different methods for selecting the next available location. The most systematically derived and regular patterns are shown in Fig. 4*a*. We created the pattern in Fig. 4*c* by borrowing a concept from fractals. These sampling patterns can be thought of as fractals with dimension $\frac{3}{2}$. If we replace each neuron in any of the patterns in Fig. 4 with the entire pattern itself, we end up with a larger pattern of the same fractal dimension. We created the patterns in Fig. 4*c* by starting with the patterns of Fig. 4*b*, drawn for $N = 4$, and then following the above procedure. Interestingly, new patterns created in this manner have the property that each holographic grating interconnecting an input and output pair on these samplings grids is distinct if the seed pattern from which it was generated satisfies this same property.

a

b

c

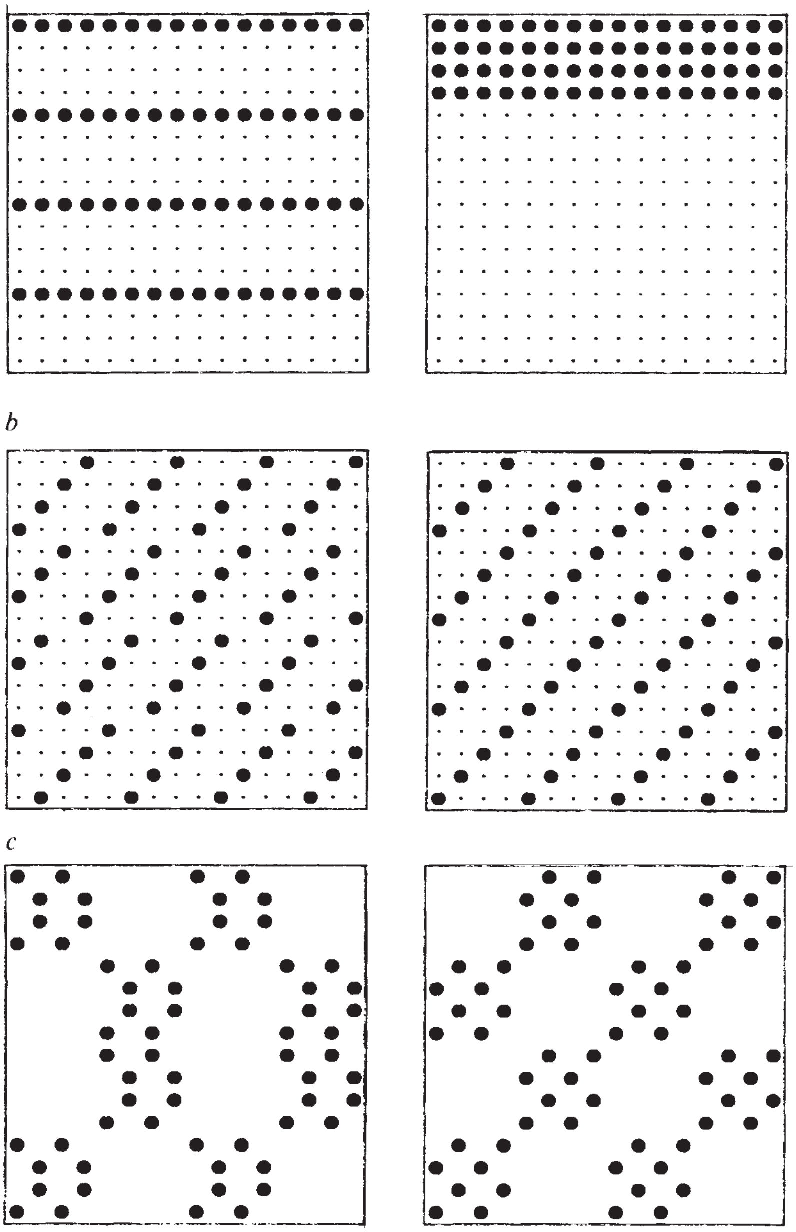

FIG. 4 Nondegenerate sampling grids. Each pair of sampling grids is such that the holographic grating that interconnects any point on the input grid with any point on the output grid is unique to that input–output pair.

The necessity of using the fractal sampling grids becomes apparent when we use the architecture of Fig. 3 as an associate memory. The goal is to record the appropriate holographic gratings in the crystal such that for a particular pair of images, the light at the output neural plane is image B if the illuminating pattern at the input is image A. We attempt to record the hologram by setting the state of the input neural plane to be the image A and the state of the output, image B (ref. 21). The input neural plane emits light towards the right in Fig. 3 and the output emits light towards the left. The light from the input is back-reflected by the PCM and interferes with the light from the output to form the hologram. The recorded hologram is reconstructed by placing pattern A at the input and the reconstruction forms at the output. The photographs in Fig. 5*a* show the result that was obtained when this procedure was performed in the laboratory. We recorded the holograms in a photorefractive lithium niobate crystal using an argon-ion laser. The two patterns used in the experiment were the capital letters A and B. The reconstruction obtained when the A was placed at the input is a smeared B. We can think of the holographic associative memory described above as an interconnection of each point on the letter A to all the points on the letter B. If this interconnection pattern is in fact recorded in the hologram, then when the A is present at the input, all the neurons that make up the letter A will produce the same reconstruction of B at the output. The result is a strong reconstruction of the letter B. In general, the output will not respond unless the activation at the input plane is sufficiently similar to the pattern used for recording the hologram. In the experiment the letters A and B were drawn on a regular rectangular grid. As a result, the establishment of interconnections of input and output points on the letters A and B inadvertently forms connections between points on the letter A at the input and points on the letter B at the output, thus giving the smeared reconstruction of B in Fig. 5*a*. When we repeated the same experiment but sampled the patterns on the fractal grids of Fig. 4*b*, we obtained the sampled patterns shown at the top of Fig. 5*b*. The reconstruction sampled by the fractal grid at the output, shown at the bottom of Fig. 5*b*, is now perfect.

This experiment vividly demonstrates how the sampling grids permit us to use each of the gratings recorded in a volume hologram to establish a distinct interconnection between two neurons. The high density with which gratings can be stored makes it possible to construct very large, densely connected networks. The capability to build large networks is useful, however, only if it is accompanied by the capability to fully load information in the connections of the network. In this experiment only a single association was stored in the hologram. In theory, this holographic associative memory has the storage capacity to support a number of associations between pairs of images equal to the number of neurons at each plane. We discuss the procedures for achieving this capacity next.

a

b

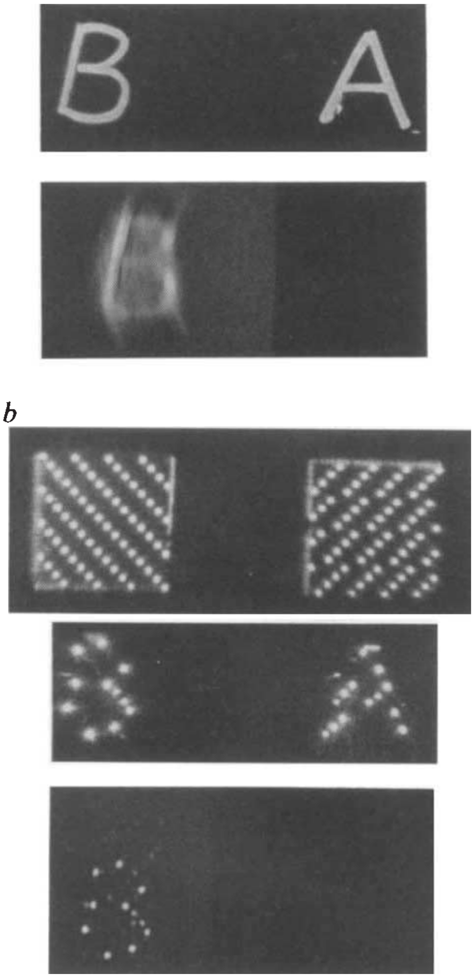

FIG. 5 *a*, Photographs of the letters B and A carried on an Ar^+ laser beam and the reconstruction of the hologram made by the interference of these two patterns. The hologram is reconstructed with A and yields a smeared version of B. The smearing is due to degeneracies in the hologram. *b*, Photographs of a pair of nondegenerate sampling grids, the letters B and A as sampled by these grids, and a reconstruction of a hologram made with the sampled patterns. The hologram is reconstructed with the letter A and yields a high-fidelity image of the sampled letter B.

Learning in photorefractive holograms

The recording mechanism described in the previous section for establishing a connection between a single pair of neurons is very similar to hebbian learning, in which the strength of the connection between two neurons is reinforced if their activation patterns are correlated. The holographic associative memory is an extension of this basic learning mechanism, with the connection between sets of input and output neurons reinforced simultaneously and in parallel. Almost all of the learning algorithms that have been developed for neural-network models make use of this basic mechanism at the synapse level. The algorithms differ primarily in the procedures used to establish the activation of the neurons during training, but the modification of each weight during a training cycle is almost universally a simple product of activations of the two neurons that are connected by the weight. Therefore, at least in principle, holographic interconnections can be used for all such learning algorithms. Most learning algorithms require a very large number of training cycles. Here we describe how photorefractive crystals can be used to form holograms recorded by an arbitrarily long sequence of exposures, thereby extending the applicability of holographic interconnections to a broad class of learning algorithms.

Photorefractive crystals are a class of nonlinear optical materials that combine three properties: they are photoconductive (light causes current to flow), electro-optic (the presence of an electric field modifies the index of refraction), and they have defects (or 'traps') in their lattice that can be optically ionized[22]. Holograms are recorded in these crystals by exposing them to light with a photon energy that is matched to the energy required to ionize a trap. The recorded hologram is stored in the spatial distribution of the ionized traps in the crystal. When we make a sequence of exposures, the distribution is continuously rearranged but the material imposes no limitation on the number of exposures. This is in contrast to photographic film, for instance, where each exposure causes an irreversible chemical change and ultimately the material saturates. Although there is no limit in the number of exposures for a photorefractive crystal, we must still design a sequence of exposures that can load the appropriate weight values in the finite pool of trap sites that are available.

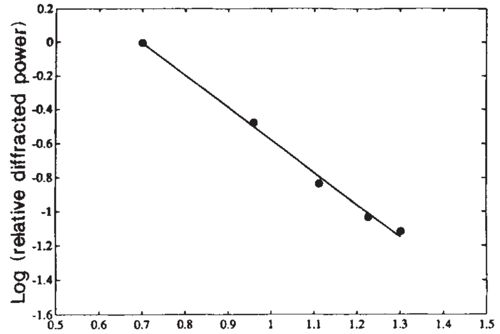

FIG. 6 Relative diffraction efficiency against the number of holograms stored for the fifth association stored between a name and a code. The slope of the least-squares fitted line is −1.95.

We will describe the extension of the holographic associative memory to multiple associations, as a simple example. We use again the basic arrangement of Fig. 3 with a photorefractive crystal used as the holographic recording medium. To establish the holographic connections that will associate M input and output image pairs, we superimpose a sequence of holograms. Each association is recorded in precisely the same way as before but the length of each exposure is different for each association. The first association is exposed for a long enough time to use all the available traps in the crystal for recording this one hologram. When the crystal is exposed to the second association, the resulting redistribution of the charge among the traps causes the first hologram to be partially erased as the second is recorded. We choose the length of the second exposure so that the two recorded holograms have equal strength. The subsequent exposure records the third association until all three have equal amplitude, and so on. At the end of M exposures, the hologram contains M associations, all with equal strength. This exposure schedule results, however, in an average diffraction efficiency for each association proportional to M^{-2} (ref. 11) rather than the ideal M^{-1} which would be obtained if the M holograms fully occupied the available dynamic range.

We can demonstrate the storage of multiple associations in a single hologram using a simple experiment. The patterns we associated were the first names of people affiliated with our research group and random dot patterns. Using the set-up in Fig. 3, the sampling grids in Fig. 4*a* and the exposure schedule above, 20 exposures were made in a strontium barium niobate photorefractive crystal. Once the holograms were recorded, each of the random dot patterns could be used to retrieve its associated name (or vice versa). We used a liquid-crystal spatial light modulator[23] to simulate the neural planes. The logarithm of the diffraction efficiency of each hologram is plotted as a function of the logarithm of the number of exposures in Fig. 6. The predicted M^{-2} relationship agrees well with the experimental results.

The reduction in efficiency that accompanies the increase in the number of exposures, ultimately limits the number of associations that can be superimposed on a single hologram. The limit is reached when the strength of the reconstruction of an individual association becomes comparable to the noise in the system. Mok *et al.* recently recorded 500 holograms of visual scenes on a single crystal using this technique[24]. This is probably close to the practical limit of the method. The number of training cycles required for the implementation of error-driven algorithms, such as error back-propagation, is not known *a priori* and it can easily be more than a thousand. As each training cycle corresponds to a holographic exposure, the present method must be extended so that a sequence of exposures of arbitrary length can be done. We are now experimenting with a method that involves two photorefractive crystals, serving as the short-term and long-term memory of the system[25]. Holographic exposures are accumulated in the short-term memory and its contents are periodically copied to the long-term memory. The long-term memory is also periodically rejuvenated by copying its contents into the short-term memory and back again. Our initial experiments confirm that this continuous exchange of information between the short-term-memory and long-term-memory crystals results in a non-decaying hologram for arbitrarily long training sequences.

Conclusion

There are remarkable analogies between the basic properties of neural-network models and simple holography. These similarities make it possible to construct, in the laboratory, very large densely connected multilayer networks with relative ease. At present, two obstacles prevent the practical use of such systems. The first obstacle is the lack of practical devices for the simulation of the neural planes; several candidate devices are, however, being investigated. The second obstacle may be more serious: we do not yet have sufficient knowledge, either from observation of the nervous system or from theoretical models, of the operation of large complex networks. Although recent progress has engendered a great deal of optimism for short-term practical benefits, we are, most probably, only at the beginning of a long journey. □

D. Psaltis and S. Lin are at the Department of Electrical Engineering, California Institute of Technology, Pasadena, California 91125, *USA. D. Brady is now at the Beckman Institute, University of Illinois, Urbane, Illinois* 61801, *USA. X.-G. Gu is now at Rockwell International,* 1045 *Camino dos Rios, Thousand Oaks, California,* 91360, *USA.*

1. Anderson, I. A. & Rosenfeld, E. (eds) *Neurocomputing* (MIT Press, 1988).
2. Sutherland, I. E. & Mead, C. A. *Scient. Am.* **237** (Sept.), 210–228 (1977).
3. Goodman, J. W., Leonberger, F. J., Kung, S. Y. & Athale, R. A. *Proc. IEEE* **72,** 850–866 (1984).
4. Psaltis, D. & Farhat, N. H. *Opt. Lett.* **10,** 98–100 (1985).
5. Abu-Mostafa, Y. S. & Psaltis, D. *Scient. Am.* **256** (March), 88–95 (1987).
6. Farhat, N. H. *Appl. Opt.* **26,** 5093–5103 (1987).
7. Owechko, Y., Dunning, G. J., Marom, E. & Soffer, B. H. *Appl. Opt.* **26,** 1900–1910 (1987).
8. Anderson, D. Z. *Opt. Lett.* **11,** 56–58 (1986).
9. Yariv, A. & Kwong, S. K. *Opt. Lett.* **11,** 186–188 (1986).
10. Wagner, K. & Psaltis, D. *Appl. Opt.* **26,** 5061–5076 (1987).
11. Psaltis, D., Brady, D. & Wagner, K. *Appl. Opt.* **27,** 1752–1759 (1988).
12. Kuffler, S. W., Nicholls, J. G. & Martin, A. R. *From Neuron to Brain* (Sinauer, Sunderland, 1984).
13. Van Heerden, P. J. *Appl. Opt.* **2,** 393–400 (1963).
14. Gibbs, H. M. *Optical Bistability: Controlling Light with Light* (Academic, New York, 1985).
15. Smith, S. D., Walker, A. C., Tooley, F. A. P. & Wherrett, B. S. *Nature* **325,** 27–31 (1987).
16. Miller, D. A. B. in *Electro-optic and Photorefractive materials* (ed. Günter, P.) 35–49 (Springer, Berlin, 1987).
17. Kim, J. H., Lin, S. H., Katz, J. & Psaltis, D. *Proc. Soc. Photoinstrum. Engnrs* **1043,** 44–52 (1989).
18. Mead, C. A. *Analog VLSI and Neural Systems* (Addison-Wesley, Reading, Mass., 1989).
19. Psaltis, D., Yu, J., Gu, X. G. & Lee, H. *2nd Topical Meet. Opt. Computing* (Opt. Soc. Am., 1987).
20. Lee, H., Gu, X. G. & Psaltis, D. *J. appl. Phys.* **65,** 2191–2194 (1989).
21. Gabor, D. *IBM J. Res. Develop.* **13,** 156–166 (1989).
22. Günter, P. & Huignard, J.-P. (eds) *Photorefractive Materials and Their Applications I and II* (Springer, Berlin, 1988).
23. Bleha, W. P. *et al. Opt. Engng* **17,** 371–384 (1978).
24. Mok, F. *Opt. Soc. Am. A. Mtg., 1989 Technical Digest Series* **12** (1989).
25. Brady, D., Hsu, K. & Psaltis, D. *Opt. Lett.* (submitted).

ACKNOWLEDGEMENTS. The authors thank H. Lee, K. Wagner, K. Hsu, J. Kim, S. Li, J. Yu and Y. A. Mostafa for their contributions and comments. S.L. is supported by an SDIO/ISTC Fellowship. D.B. acknowledges the support of the Aerojet Corporation. This research is supported by DARPA, the Air Force Office of Scientific Research, and the Army Research Office.