tant source of femtosecond pulses in the infrared.

Optical and optoelectronic information processing are other areas rich in potential where further advances in femtosecond technology can play an important role. However, just as the advances in femtosecond technology stemmed from a deeper understanding of the underlying scientific principles, so to will significant advances in optical information processing depend on similar strides in our understanding. One area of particular interest is the formation of images of femtosecond pulses. Such matrix arrays of femtosecond optical pulses should facilitate strategies for better using the potential for parallel computing inherent in optical phenomena.

Whichever strategy succeeds, it will be important for those involved in working with the technology of ultrashort pulses to respect the lessons history teaches. We should work hard to translate the advances of the 80's into useful tools and, at the same time, take responsibility for guiding the social impact those tools will inevitably produce along positive directions.

• • • • • •

R.L. Fork is professor of physics at the Rensselaer Polytechnic Institute, Troy, N.Y.

References

1. A useful general reference for material discussed in this review is R.L. Fork, *et al.*, "Ultrashort Light Pulses," American Scientist **78**, 1990, 216-223.
2. J. Shah and T.C. Damen, "Femtosecond luminescence spectroscopy with 60 fs compressed pulses," Appl. Phys. Lett. **52**, 1988, 1291-1293.
3. R.W. Schoenlein, *et al.*, "Femtosecond studies of image-potential dynamics in metals," Phys. Rev. Lett. **61**, 1988, 2596-2599.
4. C. Fattinger and D. Grischkowsky, "A Cherenkov source for freely-propagating terahertz beams," IEEE J. Quantum Elec. **25**, 1989, 2608-2610.
5. N.H. Abramson and K.G. Spears, "Single pulse light in flight recording by holography," Appl. Opt. **28**, 1989, 1834-1841.
6. K.M. Yoo, *et al.*, "Biological materials probed by the temporal and angular profiles of the backscattered ultrafast laser pulses," JOSA B**7**, 1990, 1685-1693.
7. J.D. Kafka, private communication.
8. C.C. Cutler, private communication.
9. R.L. Mollenauer, "Solitons in optical fibers and the soliton laser," Phil. Tran. Roy. Soc. London A, Vol. **315**, 1985, 437-450.

Optical Neural Networks

By Demitri Psaltis and Yong Quio

The optical implementation of artificial neural networks is a subject that combines optics and neural networks. The notion that links the two fields is connectionism. In optical computers, photons are used instead of electrons as the carriers of information. The advantage of doing this derives from the fact that photons do not directly interact with one another. This makes it easier to establish a communication network connecting a large number of processing elements. Therefore, the design of optical computers is naturally guided toward architectures that require many connections.

Neural computers, on the other hand, are computing structures whose design is motivated, at least partially, by the nervous system. One of the most striking features of the brain is its dense connectivity. Each neuron typically receives input from several thousand other units. Accordingly, artificial neural network architectures require very dense connectivity and optics is a device technology that is well suited for providing it.

A similar history of development is another common feature of optical and neural computers. Both topics became popular in the late 1950s and early 1960s and both fell out of favor soon thereafter, until recently. The difficulty that both approaches ran into was that they could not provide competitive, practical solutions to interesting problems, despite their promise for superior capabilities. This remains true today.

Nevertheless, there is renewed optimism and the level of interest in these two fields is more intense now than ever before. Whatever the reasons for this resurgence in interest may be, the remarkable fact is that the interest has been increasing for the last five years or so. The underlying reason is the transition, in both fields, from the single-layer machines of the early 1960s to the multilayer machines of today. Most of the work in neural networks in the 1960s focused on a single neuron (Figure 1), modelled as a simple threshold gate. Each of the inputs to the neuron is multiplied by a weight and the weighted signals are summed. If this sum exceeds a threshold, then the neuron turns on, otherwise it stays off.

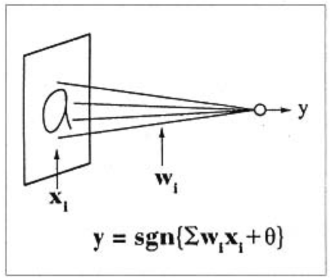

Figure 1. A single neuron simulated as a threshold gate.

What created excitement about this sort of thing was the discovery of algorithms for training such a system. If we are given a set of inputs along with the desired output for each of the inputs, then procedures such as the perceptron[1] and the adaline[2] can iteratively develop a set of neuron weights to map each of the given inputs correctly. Of course, these procedures can produce a set of weights to solve the problem at hand only if such a solution exists. The realization that there are not problems of practical interest that can be solved efficiently by a single neuron essentially brought to an end the early efforts in neural networks.

The most intensively studied optical computing system that originated in that era is the optical correlator[3] (Figure 2). This system correlates an input image with a reference that is stored holographically. The value of the correlation function at the center of the output plane is the weighted sum of the input pixels, the weights being the corresponding pixel values of the reference image. At other positions, we obtain the weighted sum of a shifted version of the input image. Therefore, the optical correlator is functionally equivalent to an individual neuron with the additional feature of shift invariance. Optical correlators have not found significant practical applications for the same reason that single neuron systems have not.

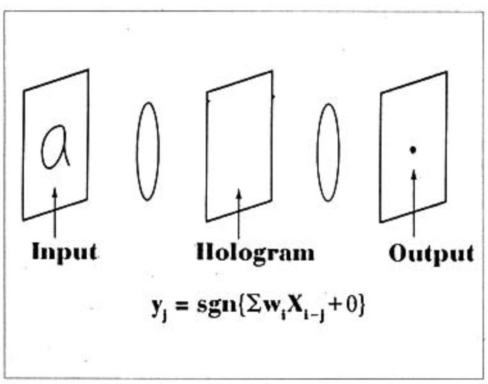

Figure 2. Optical correlator.

Recent advances

Research in neural networks in recent years[4] has focused on the study of large sets of neurons. The discovery of training algorithms for multi-layered feedforward networks[5,6] is the development that has had the strongest impact in the field. In a two-layer network (Figure 3), the pattern is first processed by a set of "hidden" neurons, and the final output is calculated by an additional neuron that accepts as its inputs the outputs from the hidden units. If we are given enough hidden units, any function can be approximated with arbitrarily good accuracy.[7] Therefore, the question for multi-layer networks is not whether there are any interesting problems they can solve, but whether they can solve some of them more efficiently than can other approaches.

While progress in neural networks occurred primarily through algorithmic developments, in optical computing the most significant development in the 1980s was the maturing of several key optoelectronics technologies. Spatial Light Modulators (SLMs) have always been considered the key component for optical computing systems. An optically addressed SLM is a 2-D array of devices, each device detecting light incident on it, electronically processing the detected signal (typically in a very simple way), and then regenerating an optically modulated signal through a light modulator or a source.

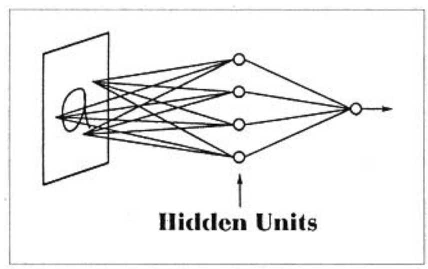

Figure 3. Multi-layer neural network.

The most significant advances in SLM technology in the past decade came from the development of ferroelectric liquid crystal devices[8] and semiconductor optoelectronic devices.[9,10] These advances have made it possible to construct multi-layer optical systems in which the output of one layer (or SLM) is optically interconnected to another. The linear portion of the system (the interconnections) is the strength of the optical approach, but the nonlinear, optoelectronic portion provides the system with computational power and diversity. Digital optical computing,[11] optical interconnections for electronic systems,[12] and optical neural networks[13] are all examples of this new breed of multi-layer optical machines.

The emergence of algorithms for training multi-layer neural networks and the advances in optoelectronics, both of which occurred during the 1980s, have made it possible to build multi-layer, optical neural networks. The activity in this area began with the optical implementation of a Hopfield network[14,15] in which every neuron is connected to all the rest. There are now many research efforts in optical neural networks, investigating optical system architectures[29-34] and methods for learning.[29-35] In what follows, we will describe briefly, as an example, a two-layer system we have built at Caltech.

An optical multi-layer network

The method used to set the values of the weights is a modification of a method described by Kanerva.[35] According to this method, the weights of the first layer are selected at random. The weights of the second layer are trained by presenting the training patterns at the input of the network that induce a response at the hidden layer through the random connections. If, for the current input, the desired response for the output neuron is high, then the response of the hidden layer is added to the weights of the second layer. This simple procedure is repeated for all the patterns in the training set.

The problem that we selected to test the operation of the system is hand-written character recognition. We created 104 training patterns by drawing characters on a 10″ × 10″ grid (four characters for each letter of the alphabet). The optical system architecture is shown in Figure 4. The network being simulated has 100 units at the input layer (matching the size of the 10″ × 10″ pixel grid), approximately 10^5 hidden units, and 26 output units—one for each letter of the alphabet.

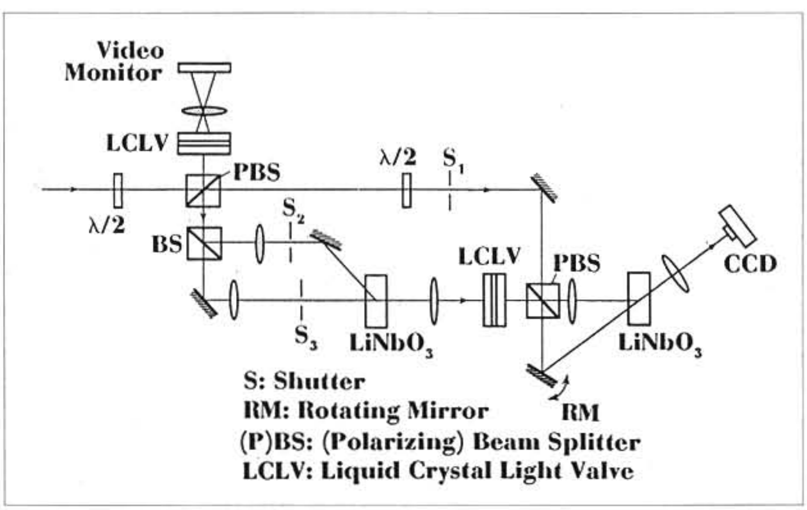

Figure 4. Optical two-layer network.

The network was trained so that the appropriate output unit was turned on when the corresponding character was presented at the input. The input and hidden layers were optically simulated by two liquid crystal, optically addressed SLMs, while the output layer consisted of a linear detector array with 26 elements. Only 100 pixels were used in the input SLM, but the hidden layer consisted of an array of approximately 300 × 300 pixels or neurons. The interconnections between the layers were implemented with holograms recorded in the two $LiNbO_3$ photorefractive crystals. Input patterns were presented to the system via a computer that displayed the pattern on a television monitor; it was then imaged onto the input SLM.

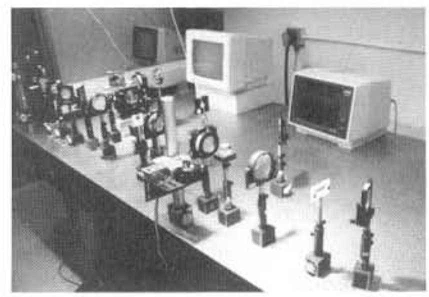

Figure 5. Experimental apparatus.

The random interconnections that the first hologram must simulate were realized by recording holograms of random dot patterns that were placed at the input and also allowing fanning to take place in the crystal, which further randomized the connections. The second hologram was trained by sequentially presenting each of the 104 training patterns at the input. Each input image induces the formation of a pattern at the SLM that simulates the hidden layer. A hologram of this hidden layer representation was recorded on the second crystal with a plane wave reference. There are 26 possible angles for the reference beam, one for each letter. The reference-beam angle can be selected through rotating a mirror controlled by the computer. The angle of the reference beam was set according to the identity of the input pattern. The photorefractive crystal was exposed 104 times to record the desired interconnection pattern.

Once the training is complete, the presentation of any one of the training patterns causes the second hologram to reproduce the reference beam with which it was recorded. This reconstructed beam codes, in the angle of propagation, the identity of the pattern. The final lens in the system focuses the reconstructed beam to a spot whose position on the output array is proportional to the angle of the reconstructed beam.

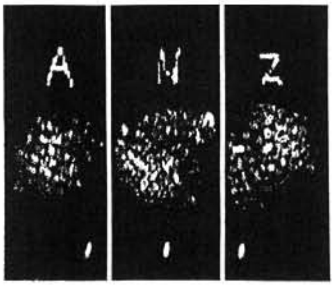

Figure 6. Examples of the signals at the input (top), hidden (middle), and output (bottom) layers in the experimental system.

A photograph of the optical system is shown in Figure 5; three examples of the signals produced by the system are shown in Figure 6. All 104 training patterns were recognized correctly by the system. This represents progress compared to the capabilities of single layer machines (correlators), which cannot recognize such a large number of small patterns. However, a lot more needs to be done before optical neural networks can have practical impact.

The system was also tested with 520 patterns that were not in the training set. The recognition rate was 60%, much better than random guessing (4%), but far below what is required for a useful character recognition system. The reason for the relatively poor performance on the test set is the choice of training algorithm used, specifically the random weights used in the first layer. This same system can be used to implement algorithms in which both layers are trained in response to the training patterns, which, in computer simulations, give much better performance. The implementation of such algorithms, however, requires the recording of thousands of holograms on each of the photorefractive crystals. Learning how to record such a large number of holograms in a single crystal is the major challenge in this field.

The power of the optical implementation is indicated by the following observation: The number of interconnections that the system of Figure 5 implements is 10^7 and the response time of the SLMs can be as low as 10 μsecs.[36] Therefore, the number of analog multiplications per second realized by this relatively simple optical system is 10^{12}.

Acknowledgement

This work at Caltech is funded by DARPA and AFOSR.

. . . . . .

Demetri Psaltis is a professor and Yong Qiao is a research assistant with the Department of Electrical Engineering, California Institute of Technology, Pasadena.

References

1. F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*, Spartan Books, Washington, D.C., 1962.
2. B. Widrow and M.E. Hoff, "Adaptive switching circuits," IRE Wescon, Conv. Rec. **4**, 1960, 96-104.
3. A.B. Vander Lugt, "Signal detection by complex spatial filtering," IEEE Trans. Inform. Theory **IT-10**, 1964, 139-145.
4. See, for example, J. Anderson and E. Rosenfeld, eds., *Neurocomputing*, MIT Press, Cambridge, Mass., 1988.
5. P. Werbos, Ph.D. Thesis, Harvard University, 1974.
6. D.E. Rumelhart, *et al.*, "Learning internal representations by error propagation," in *Parallel Distributed Processing, Vol. 1*, eds., D.E. Rumelhart and J.L. McClelland, MIT Press, Cambridge, Mass., 1986.
7. K. Hornik, *et al.*, "Multilayer feedforward networks are universal approximators," Neural Networks **2**, 1989, 359-366.
8. S. Fukushima, *et al.*, "Bistable spatial light modulator using a ferroelectric liquid crystal," Opt. Lett. **15**, 1990, 285-287.
9. A.L. Lentine, *et al.*, "Symmetric self-electro-optic effect device: optical set-reset latch," Appl. Phys. Lett. **52**, 1988, 1419-1421.
10. Y.H. Lee, *et al.*, "Room-temperature continuous wave vertical cavity single quantum well microlaser diodes," Electron. Lett. **25**, 1989, 1377-1378.
11. N. Streibl, *et al.*, "Digital optics," *Proc. IEEE* **77**, 1989, 1954-1969.
12. J.W. Goodman, *et al.*, "Optical interconnections for VLSI systems," *Proc. IEEE* **72**, 1984, 850-866.
13. Y.S. Abu-Mostafa and D. Psaltis, "Optical neural computers," Scientific American **256**, 1987, 88-95.
14. D. Psaltis and N.H. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," Opt. Lett. **10**, 1985, 98-100.
15. N.H. Farhat, *et al.*, "Optical implementation of the Hopfield model," Appl. Opt. **24**, 1985, 1469-1475.
16. B.H. Soffer, *et al.*, "Associative holographic memory with feedback using phase-conjugate mirrors," Opt. Lett. **11**, 1986, 118-120.
17. R.A. Athale, *et al.*, "Optical implementation of associative memory with controlled nonlinearity in the correlation domain," Opt. Lett. **11**, 1986, 482-484.
18. A. Yariv and S.K. Kwong, "Associative memories based on message-bearing optical modes in phase-conjugate resonators," Opt. Lett. **11**, 1986, 186-188.
19. D.Z. Anderson, "Coherent optical eigenstate memory," Opt. Lett. **11**, 1986, 56-58.
20. N.H. Farhat, "Optoelectronic analogs of self-programming neural nets: architecture and methodologies for implementing fast stochastic learning by simulated annealing," Appl. Opt. **26**, 1987, 5093-5103.
21. E.G. Paek and D. Psaltis, "Optical associative memory using Fourier transform holograms," Opt. Eng. **26**, 1987, 428-433.
22. D. Psaltis, *et al.*, "Higher order associative memories and their optical implementations," Neural Networks **1**, 1988, 149-163.
23. E.G. Paek and A. Von Lehmen, "Holographic associative memory for word-break recognition," Opt. Lett. **14**, 1989, 205-207.
24. C.C. Guest and R. TeKolste, "Designs and devices for optical bidirectional associative memories," Appl. Opt. **26**, 1987, 5055-5060.
25. J.M. Kinser, *et al.*, "Design for a massive all-optical bidirectional associative memory: The big BAM," Appl. Opt. **27**, 1988, 3442-3444.
26. M. Kranzdorf, *et al.*, "Optical connectionist machine with polarization-based bipolar weight values," Opt. Eng. **28**, 1989, 844-848.
27. L.S. Lee, *et al.*, "Continuous-time optical neural network associative memory," Opt. Lett. **14**, 1989, 162-164.
28. M. Oita, *et al.*, "Optical implementation of large-scale neural networks using a time-division-multiplexing technique," Opt. Lett. **15**, 1990, 227-229.
29. A.D. Fisher, *et al.*, "Optical implementations of associative net-

works with versatile adaptive learning capabilities," Appl. Opt. **26**, 1987, 5039-5054.
30. K. Wagner and D. Psaltis, "Multilayer optical learning networks," Appl. Opt. **26**, 1987, 5061-5076.
31. D. Psaltis, *et al.*, "Adaptive optical networks using photorefractive crystals," Appl. Opt. **27**, 1988, 1752-1759.
32. E.G. Paek, *et al.*, "Holographic implementation of a learning machine based on a multicategory perceptron algorithm," Opt. Lett. **14**, 1989, 1303-1305.
33. J. H. Hong, *et al.*, "Optical pattern classifier with perceptron learning," Appl. Opt. **29**, 1990, 3019-3025.
34. M. Ishikawa, *et al.*, "Experimental studies on learning capabilities of optical associative memory," Appl. Opt. **29**, 1990, 289-295.
35. P. Kanerva, "Parallel structures in human and computer memory," in *Neural Networks for Computing*, ed., J.S. Denker, Am. Inst. Phys., New York, 1986, 247-258.
36. G. Moddel, *et al.*, "High-speed binary optically addressed spatial light modulator," Appl. Phys. Lett. **55**, 1989, 537-539.

Optical Interconnections in the '80s*

By Joseph W. Goodman

In the late 1970s and early 1980s, the spectacular contributions of fiber optics to long distance communication were motivating the examination of the use of the optics for solving interconnection problems at the highest levels of the interconnect hierarchy, namely at the processor-to-processor level. As early as 1979, Cathey and Smith[1] had proposed a parallel free-space optical data bus for use within a computer. In 1981, Tajima, Okada, and Tamura[2] proposed the development of free-space optical buses for a multiprocessor machine and, in 1983, these same authors reported development of an experimental free-space optical data bus running with a 100 MHz clock speed.[3] Further advances on this project were described in following years;[4] the computer in which this bus is used in often referred to as Dialog.H.

In parallel developments at the beginning of the '80s, it was becoming apparent, with projected progress in VLSI technology, that performance would be limited by interconnections, rather than by the raw speed of transistors and logic gates. While the speed of gates improved with reductions of VLSI feature size, the speed of interconnections was independent scaling, and indeed was exacerbated by the growth of chip size taking place simultaneously with scaling. Furthermore, speeds of communication between chips were already much slower than speeds within chips, and large pin-out requirements were becoming more and more difficult to satisfy.

One of the earliest proposals for optical chip-to-chip interconnects was published in 1983 by Anis Husain of Honeywell,[5] but unfortunately it did not appear in a widely accessible and archival journal. In the autumn of 1983, the U.S. Army Research Office sponsored a "Palentir Meeting on Optical Interconnects," with just a few U.S. participants, to focus on the possible use of optics for solving interconnect problems at all levels of the interconnect hierarchy. From this meeting came the most broadly based set of proposals, especially at the chip-to-chip and the intrachip levels. These proposals appeared in a widely accessible publication in 1984.[6]

Development of fundamental understanding

Thus the first half of the '80s witnessed the birth of the field of optical interconnects. The second half of the decade was occupied with exploration of the potentials and limits of these concepts, as well as the advancement of optical interconnect technology itself. In 1985, an examination of the powers required to drive 1

**The term "optical interconnection" as used here refers to the use of optics to solve interconnection problems within a single system. We exclude from consideration the important field of optical local area networks, which aims to provide interconnectivity between several or many separate machines.*