

Optical neural networks: an introduction by the feature editors

Kelvin Wagner and Demetri Psaltis

This feature of *Applied Optics* is devoted to papers on the optical implementation of neural-network models of computation. Papers are included on optoelectronic neuron array devices, optical interconnection techniques using holograms and spatial light modulators, optical associative memories, demonstrations of optoelectronic systems for learning, classification, and target recognition, and on the demonstration, analysis, and simulation of adaptive interconnections for optical neural networks using photorefractive volume holograms.

The resurgence of interest in neural networks in the early 1980's was followed by numerous proposals and experimental demonstrations of optoelectronic implementations. Optics and neural networks appear to be an ideal match of requirements and capabilities. The primitives used in neural computation are analog inner products and simple point nonlinearities, precisely the computations that are implemented optically most naturally. Massive connectivity and parallelism, the two basic attributes of optics, are required in neural networks more than in any other computing paradigm. Finally, the large storage capacity and the parallel access of optical memories can be used to store and retrieve rapidly the potentially huge number of weights that are required for the implementation of large networks. This perfect match, it seemed, would quickly translate to practical optical systems that outperformed electronic implementations. The first *Applied Optics* feature on neural networks was edited by Gail Carpenter and Stephen Grossberg and included numerous papers on neural algorithms and architectures, electronic implementations, and optical architectures for neural networks.¹ In contrast, the papers in this feature are all devoted to optical and optoelectronic implementations of neural networks, and they give a snapshot of the progress made in the intervening five years. It is encouraging to see

the contrast between the first feature, which contained many pioneering optical architectural studies, and the current feature, in which the majority of papers consists of experimental demonstrations of systems and devices. However, research in optical neural networks is still concentrating on technology development and laboratory demonstrations rather than practical applications. The research of the type described in this feature has put us at the threshold of being able to address real-world applications that require huge training sets of measured data, and it will enable us to accomplish this goal for problems that are clearly beyond the capabilities of electronic approaches. To be sure, numerous research problems remain to be addressed to accomplish this goal, including a well-formulated prediction of the scalability of the optical approaches, the refinement of neural learning algorithms to match the optical technology, the development of techniques for large-scale manufacture of practical device technologies for the neurons and interconnections, and the use of effective methods to package commercial systems. It is our hope that the next feature will contain some examples of the successful application of optical neural networks to practical problems.

In this feature, 18 papers on current research in the field of the optical implementation of neural-network models of computation are presented. The papers span the range of devices for optical neurons, analysis and simulation of optical interconnections, experimental demonstrations of various optical neural networks and learning machines, and applications of optical neural networks for classification and target recognition.

The first group of papers emphasizes optoelectronic

K. Wagner is with the Optoelectronic Computer Systems Center, Department of Electrical and Computer Engineering, University of Colorado, Boulder, Campus Box 425, Boulder, Colorado 80309-0001. D. Psaltis is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, California 91125.

Received 21 December 1992.

0003-6935/93/081261-03\$05.00/0.

© 1993 Optical Society of America.

neuron arrays that are designed and fabricated for incorporation in specific optical neural-network architectures. Yoshikazu Nitta, Jun Ohta, Syuichi Tai, and Kazuo Kyuma present results of optical learning in a hybrid electro-optic neural network that uses the compact matrix-vector multiplier topology. Adaptive weights are implemented by using externally programmable variable-sensitivity photodiodes; the authors present experimental results for backpropagation learning of 14 patterns with 8 input units, 8 hidden units, and 3 outputs. Steven Lin, Annette Grot, Jiafu Luo, and Demetri Psaltis present results of the fabrication of GaAs devices that incorporate photodetectors, transistor amplifiers, and light-emitting diodes to implement thresholding and saturating optoelectronic neuron arrays. System design constraints motivate the technology selection for achieving large gain, high sensitivity, and low power dissipation, so that dense integration of neurons can be accomplished without thermal limitations. Chongchang Mao and Kristina Johnson present a 64-pixel neuron array that calculates the binarized error signals needed to implement delta-rule training of an optoelectronic matrix-vector-multiplier neural network. The device subtracts the excitatory and inhibitory inputs, compares this difference with the desired target, and uses the resulting error to turn on one of two reflective modulators that can be used to increase or decrease the corresponding weights. Stuart Collins, Jr., Stanley Ahalt, Ashok Krishnamurthy, and Daniel Stewart present an optical architecture for finding the maximum intensity of an array of inner products produced by an optical matrix-vector multiplier. They use vertically smearing optics, a wedge attenuator, and a thresholding two-dimensional (2-D) spatial light modulator (SLM) in combination with horizontal smearing optics and an optical logic-gate SLM to find the location of the maximum input.

The next group of papers is concerned with the optical implementation of associative memories. Paul Keller and Arthur Gmitro study the design of multifaceted computer-generated holograms for performing the interconnections between a 2-D array of input units to a 2-D array of outputs. They use simulations to compare the fidelity of the interconnections for off-axis and on-axis geometries, and they include fabrication error tolerancing and diffraction efficiency to optimize the hologram. Paul Horan, Andrew Jennings, Brian Kelly, and John Hegarty present the optical implementation of a second-order neural network for translation-invariant processing. Cascaded liquid-crystal SLM's are used for the outer-product input and for the weight matrix, and the system is shown to be able to adapt to nonuniformities during computer-controlled on-line training. Francis T. S. Yu, Chii-Maw Uang, and Shizhuo Yin present an approach to implementing associative memories for gray-level images by decomposing the image into binary slices. They demonstrate a sequential optical implementation of this system that use

cascaded liquid-crystal televisions to encode a binary-slice input and the weight matrix, and they use a lenslet array to produce the summations, which are accumulated in a computer to complete the association.

The next collection of papers addresses the issue of learning, so that useful classifications can be performed. S. Abramson, D. Saad, E. Marom, and N. Konforti implement a vector-matrix multiplier with bipolar weights and bipolar neuron activities by using cascaded liquid-crystal televisions, polarization logic, and spatial multiplexing. They use this computer-controlled system to implement a handwritten-digit recognition task that takes 10,000 cycles to converge, and they achieve 85% generalization on the test set. Stephen Brodsky, Gary Marsden, and Clark Guest implement a binary content-addressable network that uses polarization-coded logical XOR operations as its basic interconnection in order to measure distances, rather than using analog weights and dot products as is done typically in neural networks. A multilayer system with 12 units per layer is implemented with liquid-crystal light valves and is trained successfully to map 3×4 pixel images to other images. Alastair McAulay, Junqing Wang, and Xin Xu implement a modified positive-weight perceptron algorithm by using spatial light rebroadcasters for the adaptive interconnections. A single output neuron, implemented with a photodetector and a computer, is trained successfully to dichotomize four Chinese characters. Haruyoshi Toyoda, Naohisa Mukohzaka, Yoshiji Suzuki, and Masatoshi Ishikawa implement an adaptive optical processing system with microchannel SLM's that uses an optical Fourier-transform preprocessor to achieve shift-invariant classification of images. A small piece of the Fourier plane is used to train the second layer to autoassociate six images of handwritten letters to prototypes of these letters after 100 learning iterations. Tien-Hsin Chao and William Stoner demonstrate a multilayer image classifier that is implemented sequentially as cascaded optical correlators and that uses a wavelength-multiplexed hologram to represent the different layers. The correlation features that are needed to discriminate forward-looking infrared images of cones from spheres are trained off line; then they are recorded in the Fourier hologram used to test the recognition performance of the system. Mark Neifeld and Demetri Psaltis demonstrate two optical classifiers based on radial basis functions; optical disks are used to store a library of prototypes, which are then used to calculate the inner products with the input, and a computer is used to calculate the distance and to compute the second layer of processing. This system is used to store 650 training vectors that are rastered versions of handwritten images, and 300 test patterns are used as inputs to exhibit a generalization rate of 83% at a real-time recognition speed of 0.2 s/input.

The final series of papers examines the use of dynamic volume holographic media as the adaptive interconnections in optical learning systems. Yuri

Owechko uses cascaded gratings from a fanned reference beam to remove Bragg-degeneracy effects in a photorefractive adaptive neural network. This system is used to demonstrate single-layer perceptron learning of up to 96 patterns, and input SLM partitioning is used to implement multilayer backpropagation learning of two 570-element input patterns, using up to 1140 neurons and 244,000 weights. Donald Wunsch, Dave Morris, Rick McGann, and Thomas Caudell discuss an optical implementation of an adaptive resonance theory neural network that is based on a double phase-conjugate resonator containing an adaptive holographic interconnection and a reset SLM. The effects of two-wave mixing, self-pumped phase conjugation, four-wave mixing, fanning, crystal geometry, grating erasure, and multiple-exposure storage are considered and evaluated for implementing the adaptive-resonance theory architecture. Kelvin Wagner and Tim Slagle consider the optical implementation of competitive learning using VLSI/liquid-crystal winner-take-all arrays that are fabricated in the sparse-grid topology required for eliminating Bragg degeneracy in volume holograms. A beam-propagation simulation of the optical system that uses a simple model of a dynamic holographic me-

dium is presented and compared with the ideal algorithm. Claire Gu, Arthur Chiou, and John Hong present an analysis of cross talk resulting from nonuniform two-wave-mixing amplification in the Fourier domain of an optical interconnection system. The cross talk is observable in experimental results and is in qualitative agreement with the theoretical prediction. The final paper in the feature, by Praveen Asthana, Gregory Nordin, Armand Tanguay, and B. Keith Jenkins, is an analysis of the interconnection fidelity of the coherent-incoherent double angularly multiplexed photorefractive system, and it offers a comparison with coherent photorefractive interconnection. Beam-propagation simulations are used to evaluate interconnection systems of up to 10×10 inputs and outputs, and it is shown that, in the high-throughput regime corresponding to strong gratings, the mutually incoherent source approach can have an improved fidelity and diffraction efficiency.

Reference

1. G. A. Carpenter and S. Grossberg, "Neural networks: introduction to the 1 December 1987 issue of *Applied Optics*," *Appl. Opt.* **26**, 4909 (1987).