Multilayer optical learning networks

Kelvin Wagner and Demetri Psaltis

A new approach to learning in a multilayer optical neural network based on holographically interconnected nonlinear devices is presented. The proposed network can learn the interconnections that form a distributed representation of a desired pattern transformation operation. The interconnections are formed in an adaptive and self-aligning fashion as volume holographic gratings in photorefractive crystals. Parallel arrays of globally space-integrated inner products diffracted by the interconnecting hologram illuminate arrays of nonlinear Fabry-Perot etalons for fast thresholding of the transformed patterns. A phase conjugated reference wave interferes with a backward propagating error signal to form holographic interference patterns which are time integrated in the volume of a photorefractive crystal to modify slowly and learn the appropriate self-aligning interconnections. This multilayer system performs an approximate implementation of the backpropagation learning procedure in a massively parallel high-speed nonlinear optical network.

I. Introduction

There has been considerable interest in the optics community in recent years in the optical implementation of neural network models,[1-5] and these have been considered principally for associative memory applications.[6-15] Incoherent optoelectronic implementations of matrix vector multipliers with nonlinear electrical feedback were used to demonstrate that imperfect analog hardware worked surprisingly well in the robust environment of a neural network.[7] Holographic association with coherent light can be combined with optical nonlinearities within a strongly pumped phase conjugate mirror,[8-10] or with the nonlinear thresholding capabilities of an optical spatial light modulator,[6] to implement image association. Volume holograms can be repetitively exposed to a number of Bragg angle multiplexed connectivity patterns to produce a holographic interconnection matrix.[11] These systems are programmed to perform a fixed operation by precalculating the interconnections with an easy learning procedure, so that fixed points of the idealized neural dynamics are the desired associative recall. One of the most intriguing properties of a neural network is the ability to learn dynamically the interconnections that correspond to a desired behavior through an iterative adaptation of the weight matrix through outer product perturbations.[1,4,5] Optical implementations of adaptive associative memories using optoelectronic components and spatial light modulator technology have

been suggested.[12] A fascinating all-optical nonlinear dynamical system for adaptive association based on a saturating cubic nonlinearity in a phase conjugating dynamic volume holographic resonator has been proposed.[13] An even more powerful learning paradigm, sometimes called hard learning, involves either error driven learning, reinforcement learning, or self-organizing principles.[4,5] A hybrid electrooptical approach to Boltzmann learning has been proposed that is based on an incoherent optoelectronic matrix–vector multiplier interfaced with a microcomputer.[14] Error driven behavioral modification has the ability to sense system performance and adapt the synaptic weights in a manner which will compensate for some of the device imperfections and interconnection misprogrammings that caused the unwanted behavior. This paper explores the match between the backpropagation error driven multilayer learning procedure[1,2] and optical networks,[15,16] while ignoring the biological implausibility of bidirectional synapses, because of the intrinsic bidirectionality of optical interconnections. This system is a feed forward multilayer perceptron which has the potential of more general computationally universal behavior than single-layer associative networks. However, it differs from the recurrent networks because all the feedback dynamics are involved in training the modifiable interconnections and not in processing the input. We propose a new optical implementation of this multilayer learning system which uses self-aligning volume holograms to bidirectionally interconnect nonlinear etalons which act as the bidirectional optical neurons. This architecture combines the robustness of the distributed neural computation and the backpropagation learning procedure with the high speed processing of nonlinear etalons, the self-aligning ability of phase conjugate mirrors, and the massive storage capacity of volume holograms to pro-
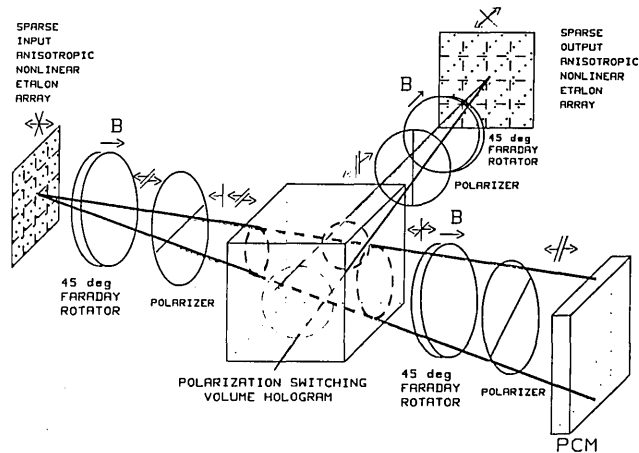
Fig. 1. Optical backward error propagation architecture with polarization multiplexed forward and backward waves, nonreciprocal polarization filtering, and self-aligning polarization switching volume hologram.

duce a powerful and flexible parallel optical processor.

One version of a single layer of this optical backpropagation architecture is shown in Fig. 1, and the operation is briefly described before discussing the idealized backpropagation algorithm and the details of this optical implementation. The learning algorithm in this single-layer optical perceptron begins with the repetitive presentation to the network input of the set of training patterns in a uniformly random sequence. Initially, the system gives rise to a sequence of output patterns through the holographic interconnection and output nonlinearity, which is different from the desired target response sequence. An error pattern is formed, either electronically or optically, by taking the difference between the actual output pattern and the targeted response. The difference pattern is sent backward through the output neurons and into the network using the same etalons and holographic interconnections, but encoding the error with an orthogonal polarization, or a slightly different frequency, or pulsed at a jittered time than the forward-propagating signal. This multiplexing of the forward and backward waves in orthogonal eigenmodes avoids direct interference between these waves. Meanwhile, the undiffracted portion of the input pattern is phase conjugated by an auxiliary phase conjugate mirror, which retroreflects each component of the input wavefront back toward the position at the input from which it originated. The phase conjugate beam has the polarization rotated or the wavelength shifted to match the error encoding to act as a self-aligning reference beam for the backward-propagating error wavefront. A volume hologram is recorded within the photorefractive crystal as the interference pattern between the phase conjugated input pattern and the backward propagating error signal. This is mathematically equivalent to changing the holographic connectivity matrix by the outer product of signal and error pattern vectors. The next time that this particular input pattern is present-

ed to the network, it produces a diffraction pattern that more closely resembles the desired output pattern. Eventually, the hologram will learn the correspondence between a set of input patterns and the associated responses as long as the set of input patterns is linearly separable, which implies that a holographic interconnection exists that produces the desired pattern transformation. Since the holographic reference wave is generated by a phase conjugate mirror, as the network learns it will also self-align as well as correct for some of the optical imperfections present in the system components.

When the desired pattern transformation is not linearly separable, as in most difficult problems of interest, it is necessary to adaptively implement more complex nonlinear decision surfaces.[17] One way that this can be accomplished is by stacking these single-layer networks up to form a multilayer network of holographically interconnected nonlinear devices that is trainable by backpropagating the error signal through the layers. When the error pattern strikes the hologram, part of it is diffracted toward the previous layer of nonlinear devices, known as hidden units, by the transpose of the interconnection matrix seen by the forward-propagating patterns, which is the necessary connectivity for backpropagating the error. The backpropagation algorithm also requires that the transmission function of the hidden units to backward-propagating signals be the derivative of the forward mode sigmoid transfer function evaluated at the current operating level of each device. The derivative is peaked where the nonlinear sigmoid transfer characteristic has a large differential gain, so that if the hidden unit is operating in this region the connections leading to it will be strongly modified by the efficiently transmitted error signal, thereby helping that neuron to decide that it should be either high or low on subsequent presentations and not between. The multiple layers of interconnections will be continuously modified until all the patterns within the training set pro-
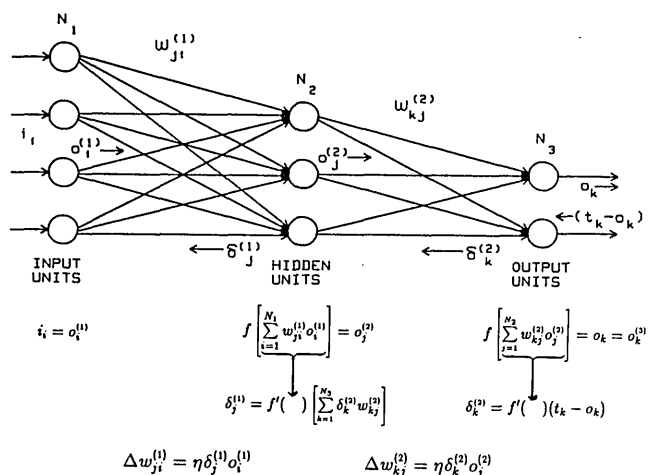

Fig. 2. Two-layer network for backpropagation learning, feed forward equations, backward-error-propagation equations, and learning rule.

duce outputs very near the flat upper or lower levels of the nonlinear device sigmoid response, so that the error signals are not allowed to backpropagate through the network. When convergence is reached, the error signals that are generated at the final layer become very small for all members of the training set.

II. Backpropagation Learning Procedure

In this section we briefly review the derivation of the backward error propagation learning procedure[1,2] to establish the notation and encapsulate the system characteristics that the optical architecture must incorporate. A schematic representation of a two-layer network is shown in Fig. 2, which consists of an input layer globally interconnected to a hidden layer, which is interconnected through a second weighted communication network to an output layer. The interconnection strengths are modifiable, so that the system can be trained to perform a desired pattern transformation from the input space to the output space. The binary signals applied to the input layer of N_1 neurons are reproduced at the output of these neurons as binary outputs, which are the inputs to the first layer, so that $i_i = o_i^{(1)}$. The outputs of the first layer are interconnected through an $N_2 \times N_1$ weight matrix $w_{ji}^{(1)}$ to a hidden layer consisting of N_2 neurons, forming presynaptic input strengths which are linear combinations of the outputs from the previous layer:

$$s_j^{(1)} = \sum_{i=1}^{N_1} w_{ji}^{(1)} o_i^{(1)}. \tag{1}$$

The hidden layer of neurons performs a soft thresholding operation on these presynaptic inputs, with a nonlinear sigmoid response $f(s)$, forming the outputs of the hidden layer which become the inputs to the second layer:

$$o_j^{(2)} = f[s_j^{(1)}] = f\left[\sum_{i=1}^{N_1} w_{ji}^{(1)} o_i^{(1)}\right]. \tag{2}$$

The outputs of the hidden layer are interconnected through the $N_3 \times N_2$ weight matrix $w_{kj}^{(2)}$, which gives the N_3 presynaptic network input to the final output layer, as a linear combination of the hidden layer outputs:

$$s_k^{(2)} = \sum_{j=1}^{N_2} w_{kj}^{(2)} o_j^{(2)}. \tag{3}$$

The final layer performs the same nonlinear soft thresholding operation as the hidden layer giving the N_3 network outputs:

$$o_k = o_k^{(3)} = f[s_k^{(2)}] = f\left[\sum_{j=1}^{N_2} w_{kj}^{(2)} o_j^{(2)}\right]. \tag{4}$$

These outputs represent the response of the network for a given set of inputs i_i, and it is the job of the training procedure to modify the interconnection weight matrices so that the actual response closely approximates the desired system response. Not all input–output mappings are possible in a network of a specific size, but complex problems of a cognitive nature with fuzzy decision boundaries have been efficiently performed in a multilayer network of this type.[18]

The desired response for the input $i_i(n)$, presented at the input of the network on the nth machine cycle, is given by a target vector $t_k(n)$, which differs from the network output $o_k(n)$, so the network error vector is given by $\delta_k(n) = [t_k(n) - o_k(n)]$. A positive definite mean-squares error (MSE) energy functional can be formed to characterize the systems behavior, and minimizing this function for all n will improve the quality of the behavior of the multilayer network:

$$E(n) = \frac{1}{2} \sum_{k=1}^{N_3} [t_k(n) - o_k(n)]^2. \tag{5}$$

A gradient descent procedure can be employed to modify the elements of the weight matrices and push them in the direction that improves the network performance, as measured by the MSE energy function, on subsequent presentations of a given pattern:

$$\Delta w_{kj}^{(m)} \propto -\frac{\partial E}{\partial w_{kj}^{(m)}}. \tag{6}$$

This weight update rule is designed to move the weights in a direction that rolls down the gradient of the energy surface in an amount which is proportional to the local slope. Ideally, the energy function should be averaged over the entire set of training patterns, so that the modification of the weight matrices is in the appropriate direction to improve the system response for the entire training set. However, a temporally localized learning can be performed by using a small acceleration coefficient η and modifying the weights after individual pattern presentations. The modification of the weights that results after cyclically presenting the training set in arbitrary order many times can approximate the desired change. The gradient descent is calculated by using the chain rule and representing the derivative of the energy function with respect to the weight matrix elements as a product of two parts, the backpropagating error and the forward-propagating signal:

$$\frac{\partial E}{\partial w_{kj}^{(m)}} = \frac{\partial E}{\partial s_k^{(m)}} \frac{\partial s_k^{(m)}}{\partial \omega_{kj}^{(m)}} = -\delta_k^{(m)} o_j^{(m)}. \tag{7}$$

The derivative of the energy with respect to the presynaptic input to the mth layer is defined to be $-\delta_k^{(m)}$, which is the backpropagating error signal in that layer. In the final layer this term is similar to the standard form of a least-mean-squares (LMS) error signal, as originally derived for the single-layer Adaline:[19]

$$\delta_k^{(2)} = -\frac{\partial E}{\partial s_k^{(2)}} = -\frac{\partial E}{\partial o_k^{(3)}} \frac{\partial o_k^{(3)}}{\partial s_k^{(2)}} = (t_k - o_k) f'[s_k^{(2)}]. \tag{8}$$

The first term is found by directly differentiating the energy function, which yields the standard error signal used in adaptive filters, and the second term is found by differentiating the nonlinear response of the
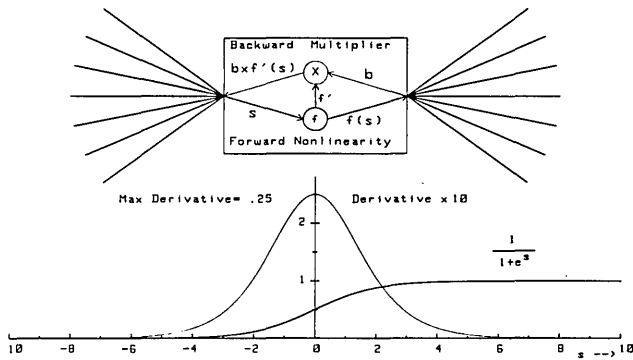
Fig. 3. Bidirectional neuron for backpropagation, its forward-mode saturating nonlinearity, and magnified derivative.

neurons. The significance of Eq. (9) is that to translate the kth component of the output error vector back into the final layer of the network it must simply be multiplied by a value which is locally computable within the kth output neuron. Thus the network output error function $\delta_k = (t_k - o_k)$ can be sent back through the corresponding output layer neurons, which multiply the error component by the derivative of the nonlinear sigmoid response at the current operating level of that output neuron. The error signal which is used to program the weights of the final layer is propagated back through those weights by multiplying by $w_{kj}^{(2)}$, and all the appropriately weighted error signals converging on the jth hidden neuron are summed to form a backpropagating presynaptic network input. The weighted sum of the error functions transmitted in the backward direction by the final layer is computed using the same interconnection matrix seen in the forward processing mode, but summing over the N_3 output neurons using the transpose of the matrix which is used for the forward-propagating interconnection:

$$\delta_j^{(1)} = -\frac{\partial E}{\partial s_j^{(1)}} = -\frac{\partial E}{\partial o_j^{(2)}}\frac{\partial o_j^{(2)}}{\partial s_j^{(1)}}$$

$$= -\left[\sum_{k=1}^{N_3}\frac{\partial E}{\partial s_k^{(2)}}\frac{\partial s_k^{(2)}}{\partial o_j^{(2)}}\right]\frac{\partial o_j^{(2)}}{\partial s_j^{(1)}} = \left[\sum_{k=1}^{N_3}\delta_k^{(2)}w_{kj}^{(2)}\right]f'[s_j^{(1)}]. \quad (9)$$

This represents an iterative algorithm for successively computing the error function at deeper layers back toward the beginning of the network in terms of the error function injected back into the final layer. Alternatively, this algorithm can be considered to represent a wavefront that backpropagates through the network, multiplying by the weights, accumulating at the neurons, and multiplying by the neurons backward transmittance to compute the appropriate error to program the previous layer. The network is highly nonlinear in the forward-propagating direction, but the backpropagating wavefront is computed using only linear operations.

The neurons must, therefore, have two signal pathways as shown in Fig. 3. The two pathways share the same weights on the connected layers, but the neuron response is a nonlinear soft thresholding for forward-propagating signals and a multiplier that only allows a backward propagating error through the neuron when the slope of the forward mode operation is large. The transmitted components of the backpropagating error vector are only large when the corresponding output neurons are operating in the steep thresholding regime where the derivative is large, and that component was significantly in error at the network output. Any neuron that had decided that it is a one or zero by being well above or below the threshold knee inefficiently transmits the error back into the previous layer of the network. From the definition of the change in the weight matrix given in Eq. (6) and the chain rule expansion of Eq. (7) we can write the form of the weight update rule for the mth layer according to this first-order gradient descent procedure:

$$w_{kj}^{(m)}(n+1) = w_{kj}^{(m)}(n) + \eta\delta_k^{(m)}o_j^{(m)}. \quad (10)$$

The error transmitted by the neurons back into the previous layer of interconnections is used to modify the weights of that layer through this outer product update rule. The weighted interconnection $w_{kj}^{(m)}$ is carrying the output from the jth neuron in the mth layer $o_j^{(m)}$ to the kth neuron in the $(m + 1)$st layer, which is simultaneously broadcasting the error function $\delta_k^{(m)}$ back into the mth layer of the network. The product of this forward-propagating signal and backward-propagating error takes place within each weighted synaptic connection as the desired weight update contribution, completely independently of what is taking place within all the other weighted connections, and this is the only information needed to update that weight, so this learning rule can be said to be a local update rule. The training procedure for the final layer of weights is given by an appropriate outer product learning rule, which is a local update rule that takes place within each weighted signal pathway, but the problem of credit assignment of the MSE energy to the earlier layer has been solved by nonlocally backpropagating the error vector. This is referred to as the backward error propagation algorithm for training multilayer networks, and it can be further generalized to N layer networks or networks with feed forward interconnections, e.g., when the first layer connects directly with the output layer as well as indirectly through the hidden layer to the output layer. For more details of the derivation, operation and utility of this multilayer network training algorithm the reader is referred to Refs. 1, 2, and 18.

III. Optical Implementation

The optical implementation of a backpropagation network requires two basic bidirectional components, the interconnection matrices, and the nonlinear units. Volume holograms appear as the most promising candidate for implementation of an interconnection matrix because of the large storage capacity possible within the volume of a crystal and the dynamic response possible with a photorefractive crystal. The readout

of a volume hologram can be accomplished with either a forward-propagating beam or a backward-propagation. Spatial light modulators (SLMs) could also be used as the interconnection element for small networks, but they would have to be both bidirectional and optically addressable to be used in a backpropagation network. In this paper a new self-aligning approach to adaptively forming optical interconnections based on phase conjugating one of the undiffracted beams is presented. This technique uses interferometric detection in the volume of a photorefractive crystal to accomplish all the outer product multiplications necessary for weight matrix perturbation. These $N_1 \times N_2$ weight updates are calculated in parallel by exposing the crystal with N_1 phase conjugated collapsing spherical waves and N_2 expanding spherical waves simultaneously.

The nonlinear units or neurons need to threshold the forward-propagating beam while transmitting the backward beam only when the forward beam nonlinearity is in the high slope, or undecided, regime of operation. A special purpose, bidirectional, detector modulator pair array structure could be tailored to generate the desired backpropagation neuron responses by utilizing the appropriate integrated electronic circuitry, but the individual neurons could become quite complicated with this conventional optoelectronic integrated circuit approach. Appropriately modified transmissive spatial light modulators might be considered for backpropagation neurons, and one possible structure of this type is illustrated in Fig. 4. In this type of birefringent SLM, crossed polarizers are placed on either side of an electrooptic medium which is optically addressed by a photoconductor. A high voltage is applied across a transparent conductor in contact with the photoconductor on one side and a transparent conductor on the other. To use this type of electrooptic device as a backpropagation neuron the induced birefringence must be doubled for the back-

ward propagating wave to obtain a saturating forward nonlinearity while obtaining a derivative backward multiplication. This can be accomplished by a pair of photoconductors, both addressed by the same forward propagating beam, where one is used to modulate the forward-propagating device which is biased with a voltage V_π, while the other is used to modulate an EO device with a saturation voltage $2V_\pi$. The forward-propagating modulator is used to modulate a fixed intensity pump I_0 so that a single half-cycle of a saturating nonlinearity can be generated, $I_f^m = I_0 \sin^2(I_f^{m-1}/I_{sat})$ for $I_f^{m-1} < I_{sat}$, and $I_f^m = I_0$ otherwise. The backward-propagating modulator is used to multiply the backward-propagating error signal by a function $I_b^{m-1} = I_b^m \sin^2(2I_f^{m-1}/I_{sat})$ for $I_f^{m-1} < I_{sat}$, and $I_b^{m-1} = 0$ otherwise, and this is of the form of the desired derivative multiplication. Since the two functions required of a backpropagation neuron can also be accomplished with a simpler nonlinear resonator structure, and the response time of these nonlinear etalons can be extremely fast compared to SLM technology, they were chosen for study in the architecture presented in this paper.

A. Nonlinear Fabry-Perot Backpropagation Neurons

Nonlinear Fabry-Perot etalons[20] are a promising candidate for implementing the neurons in an optical learning network because they can perform nonlinear operations on arrays of coherent beams, which allows the outputs to be used to record and modify interconnection holograms. A soft thresholding operation can be performed on a forward-propagating beam by decreasing the cavity detuning below the critical detuning needed for bistability.[20] These optical neurons cannot easily implement the idealized derivative transmission required for backpropagation, but a similar peaked response can be obtained by operating a nonlinear etalon in the probe mode[21] for the backward-propagating error signal. In this mode, the Fabry-Perot resonance is scanned by the nonlinear dependence of the index on the intracavity intensity, which varies in response to the high power forward beam intensity. The weak backward-propagating probe beam does not scan the cavity, but it is modulated by the current state of the cavity transmission function, which is the appropriate multiplication type of response needed in the backward direction. The probe mode transmission is peaked at the resonance of the Fabry-Perot, which occurs when the sigmoid response to the forward beam reaches the upper level. The peak maximum is not exactly at the region of the highest slope of the forward beams nonlinear sigmoid response, but since the forward and backward beams have different polarizations, or different wavelengths, the resonance function can be offset to achieve a properly positioned probe beam resonance peak.

In the polarization multiplexed case this shift can be induced by including a thin birefringent sheet in the cavity,[22] or perhaps a tunable birefringence can be caused by applying a static external field to the cavity. This type of birefringent nonlinear Fabry-Perot etalon
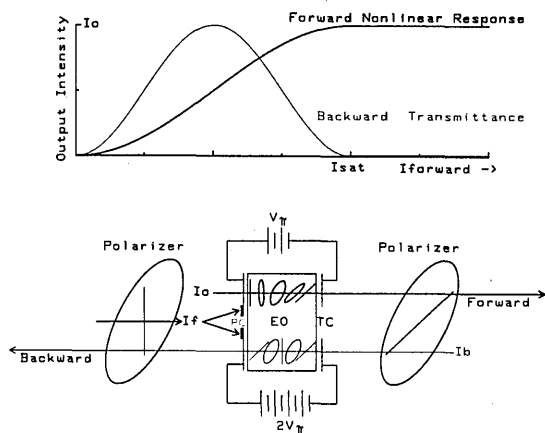
Fig. 4. Input output relations for a special purpose bidirectional optically addressed spatial light modulator backpropagation neuron: PC = photoconductor; EO = electrooptic; TC = transparent conductor.
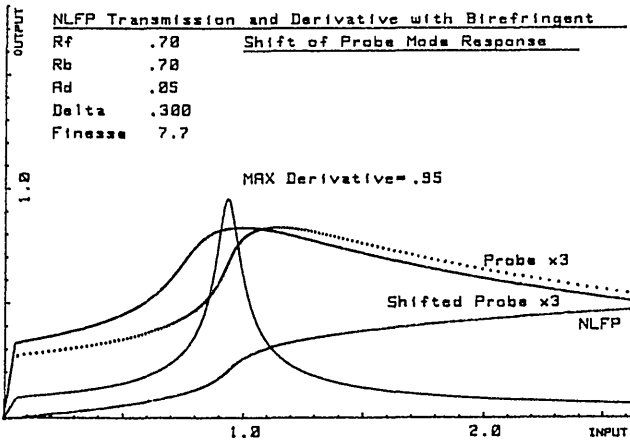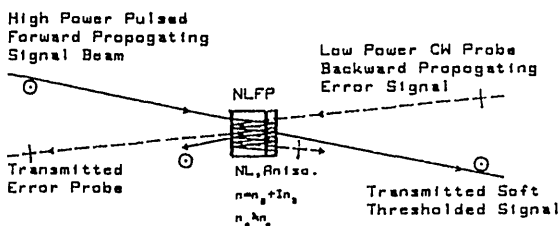
Fig. 5. Nonlinear Fabry-Perot etalon sigmoid response, its deriva-
tive, and the probe mode transmission for the two polarizations with
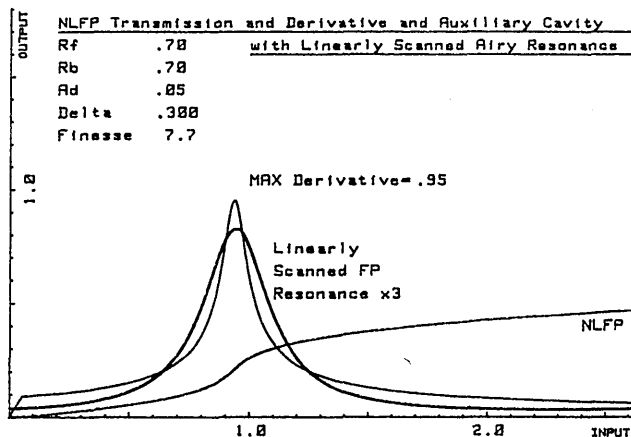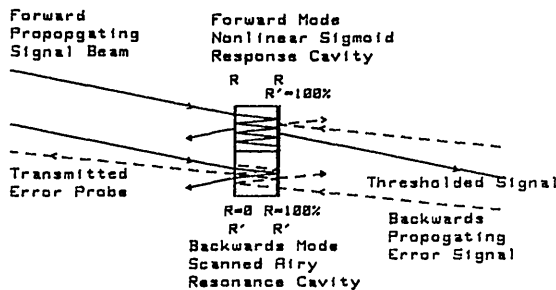an auxiliary intracavity birefringence.

Fig. 6. Dual-cavity nonlinear Fabry-Perot etalon with forward-
propagating nonlinear response and backward-propagating scanned
resonance probe mode transmission

and a simulation of the forward-mode sigmoid transfer
function is shown in Fig. 5 along with its derivative and
the shifted probe mode response approximation to this
derivative. This device can implement the desired
sigmoid nonlinearity of the high intensity forward-
propagating signals with a differential gain greater
than one, although the actual gain in transmission is
less than one. The probe mode response is not sym-
metric about the peak because the Airy function reso-
nance is scanned by the intracavity intensity which is
equal to the transmitted sigmoid response divided by
the backmirror transmittance. This asymmetry con-
tinues to allow signals that are above threshold to build
up interconnection gratings in the previous stage cor-
responding to correlated inputs, thereby partially
compensating for the slow forgetting of gratings by the
volume hologram. However, the high level of trans-
mission for the probe beam when the etalon pump is
below threshold is undesirable. By decreasing the
finesse of the cavity to the forward-propagating beam a
trade-off can be made between the peak width and off--
resonance transmission of the probe mode response,
with the switching energy for the forward-propagating
nonlinear device characteristic.

Another possibility would be to use two closely
spaced cavities, both addressed by the same forward-
and backward-propagating resolution spots, as illus-
trated in Fig. 6. In this case one cavity is optimized to
produce a sigmoid response of the forward-propagat-

ing beam while blocking the backward-propagating
error signal, while the other cavity is resonant to the
backward- propagating beam. The Fabry-Perot reso-
nance of the backpropagating cavity is linearly
scanned by the 100% reflected forward-propagating
incident intensity, thereby producing a good approxi-
mation to the desired symmetric derivative response.
We expect that learning and eventual convergence can
be achieved in a multilayer optical network with the
forward and backward response that can be obtained
from these scanned resonance devices, even though the
responses do not precisely match the nominal respons-
es of the backpropagation algorithm because of the
robustness of this gradient descent learning procedure.

B. Description of a Single Layer of the Optical
Architecture

A single layer of an architecture that can perform
this type of multilayer perceptron learning procedure,
using polarization multiplexing of the forward-propa-
gating processing beam and backward-propagating
teaching beam, is shown in Fig. 1. The illustrated
architecture is one implementation of this class of
backward-error-propagating holographic learning ma-
chines that serves to illustrate the principles involved.
Notice that no lenses are shown in this diagram be-
cause the volume hologram can perform the desired
weighted interconnection imaging by exposing it with

the proper expanding image and focusing reference beam to form a Fresnel volume hologram. If Fourier lenses are inserted between the etalon arrays and the volume holographic crystal, the exposed hologram will be a Fourier hologram with planar fringes, and the momentum space analysis will be simplified, but the processor learning and self-aligning operations will be similar.

The forward-propagating pattern vector transmitted by the anisotropic nonlinear etalon array on the left-hand side of the figure is polarized at a $-45°$ angle and is rotated clockwise by $45°$ as it passes through the nonreciprocal Faraday rotator so that it becomes horizontally polarized. This aligns the forward-propagating beam with the polarizer allowing it to pass and illuminate the polarization switching volume hologram. The diffracted beam consists of a weighted interconnection of the forward-propagating pattern vector by the current state of the holographically represented interconnection matrix, stored as a superposition of curved and chirping space charge gratings within the photorefractive crystal. The diffracted beam is polarization switched by the birefringent diffraction mechanism to an orthogonal polarization to the input, and this vertical polarization state is rotated clockwise by $45°$ through the following Faraday rotator so that it falls on the next etalon array with the same $-45°$ polarization as the forward-propagating transmitted beam from the previous stage. The undiffracted beam passes straight through the volume hologram and has its vertical polarization rotated by $45°$ as it passes through the Faraday rotator, so that it falls on the phase conjugate mirror (PCM) with a polarization angle of $45°$, the same as the counterpropagating pump beams (which are not shown), producing an identically polarized phase conjugate beam, which is composed of an array of beams that retroreflect back toward the etalon sources that each originated from. This phase conjugate beam passes through the nonreciprocal Faraday rotator picking up another $45°$ rotation (instead of unwrapping the rotation as would occur with a reciprocal optical activity based rotator), emerging vertically polarized to act as the reference beam array for the self-aligning holographic outer product exposure with the backward-propagating error signal. The backward-propagating error signal emerges from the backside of the output etalon array with a $45°$ polarization that is orthogonal to the forward-propagating beam. This allows for the polarization filtering based separation of the reflected forward-propagating beam from the transmitted backward-propagating beam as well as the independent tuning of the relative Fabry-Perot resonance position of the forward- and backward-propagating beams. The backward propagating error signal is rotated to a vertical polarization by the Faraday rotator so that it interferes in the volume hologram with the vertically polarized phase conjugate reference beam and not with the horizontally polarized undiffracted forward-propagating signal. The interference of a backward-propagating error signal emerging from a particular etalon at the output with the phase conjugated forward-propagating beam emerging from a particular etalon from the input produces a self-aligning volume Fresnel holographic interference pattern that interconnects these two etalons for both forward- and backward-propagating beams with the exact same diffraction efficiency, or weight, due to the reciprocity of linear electromagnetic systems. The interference of the backward-propagating error beam with the phase conjugate of the forward-propagating beam records a Fresnel grating due to each pair of beams that is present, perturbing the weighted interconnection matrix represented by the hologram by the outer product of the signal and error vectors and thereby pushing the matrix toward the desired interconnection solution. The backward-propagating beam is polarization switched by the volume holographic diffraction mechanism, producing a horizontally polarized beam which is the appropriate weighted summation of the error signal by the transpose of the interconnection matrix seen by the forward-propagating beam. This passes through the polarizer and is Faraday rotated by $45°$ to be incident on the etalon array with a $45°$ polarization angle, orthogonal to the forward-propagating beam, and the same as the backward-propagating beam which emerged from the previous output layer. The undiffracted phase conjugate of the forward-propagating beam needs to be blocked so that it is not confused with the copropagating diffracted backward-error-propagating signal, and this is accomplished by the polarizer which blocks the vertical polarization of the undesired phase conjugate reference beam. The indicated nonreciprocal polarization filtering will also remove the unwanted reflections from the nonlinear etalons and unwanted diffraction terms produced by the hologram. The diffracted phase conjugate reference and undiffracted backward-error signal emerge at a different angle and will not focus on the etalon; thus they can be ignored, or they can be examined to determine intermediate states of the hidden neurons. Each layer is completely compatible with the previous and the following layers so this type of learning network can be stacked up to form a complex multilayer learning machine.

C. Requirements for the Holographic Interconnection

The dynamic holographic interconnection technique described in this paper is based on the photorefractive effect, which is a light-induced index of refraction modulation that occurs in photoconductive electrooptic crystals. A space charge grating image of an interference profile is created by carriers ionized from fixed traps into the conduction band, where the mobile carriers redistribute under the influence of drift, diffusion, and bulk photovoltaic effects, until they recombine with an unoccupied trap. The redistributed optically generated carriers produce a space charge grating with a fundamental Fourier component that may be phase shifted from the interference profile. The spatial variations of the resulting space charge pattern produce a corresponding electric field through Poisson's equation. This space charge field

induces an electrooptic modulation of the local impermeability tensor as long as the appropriate electrooptic tensor coefficient is nonzero. In turn, this couples the input field into a diffracted output field as long as the appropriate impermeability tensor coefficient is nonzero. The implementation presented in the previous section is based on a polarization switching diffraction mechanism for which it is required to have an off-diagonal impermeability tensor coefficient. This requires electrooptic tensor coefficients in the bottom half of the reduced subscript electroopotic matrix, which can take place in some electrooptic volume holographic materials, such as $Bi_{12}SiO_{20}$, $LiNbO_3$, $BaTiO_3$, and GaAs. Self-aligning recording combined with polarization switching diffraction between linear eigenmodes requires an optically isotropic medium (or one in which anisotropy can be eliminated through the application of a static field), with no optical activity, and these conditions imply that a material of symmetry group $\bar{4}3m$, such as some III–V semiconductors (e.g., GaAs or InP), should be used as the photorefractive holographic medium. The efficiency of the diffraction depends on the effective coupling strength which depends on the angle of the gratings, polarization of the input wave, and momentum matching (Bragg) condition in a rather complicated fashion. However, use of a Fresnel hologram can produce an averaging over all these effects for all the interconnection holograms, while in a Fourier hologram with planar fringes each grating has a different diffraction efficiency.

The polarization switching diffraction efficiency and the holographic storage capacity can be simultaneously maximized by having the input and output beams propagating at large angles, as indicated in the figure. The unwanted polarization switching grating exposures due to the simultaneous presence of multiple reference (or object) beams produce crosstalk of the undiffracted forward-propagating beam, which can be eliminated with the indicated polarization filtering. The storage capacity of the volume hologram will enforce limits on the number of nonlinear devices that can be interconnected and on their topology because of the cone of ambiguity associated with Bragg diffraction.[23] A sparse array of etalons will have to be utilized to implement a fully global interconnection without unwanted crosstalk, which will also facilitate the dissipation of heat generated in the nonlinear etalons and thereby allow a very high speed of operation. However, the learning operation must occur slowly for the backpropagation algorithm to converge properly, and this is well matched with most photorefractive crystal volume holograms, because the crystal response times are slow, and the perturbation of an existing space-charge grating by a single outer product exposure is very small.

It is necessary to be able to both selectively erase holographic gratings, thus decreasing the connection strength between particular etalons, as well as to strengthen individual gratings thereby increasing the corresponding elements of the interconnection matrix.

Selective erasure can be accomplished by using a phase encoded backward-propagating error signal, where a phase angle of 0 is used to represent all positive error signals, and a phase angle of π is used to represent all negative error signals. Fresnel gratings that are built up with a phase angle of 0 can have the corresponding interconnection decreased selectively by shifting the recording interference profile by π, as demonstrated in Sec. V, and by Huignard for Fourier holograms.[24] Alternatively, selective interconnection erasure might be accomplished by strengthening interconnection gratings when the applied bias field is in one direction, causing the resulting space charge grating to shift away from the optical intensity profile in the direction of the E field by approximately $\pi/2$, while decreasing interconnection gratings when the bias field is reversed, producing a canceling space charger grating with a phase shift of $-\pi/2$. Another approach to decreasing interconnection strength would be to rely on the simultaneous erasure of all the gratings by the optical readout and thermal effects, thereby inserting a forgetting term in the dynamical equation for the holographically represented interconnection matrix. This approach requires continuous reinforcement to avoid forgetting everything that has been learned. Once learning has been completed a mechanism of fixing the hologram could be used to make the interconnections permanent.[25]

A scheme must be devised to implement negative interconnection stengths, or else all the signals must be placed on an appropriate bias. An attractive possibility for the implementation of bipolar weights is to use the phase shift of each grating to represent its sign and count on destructive interference within each nonlinear etalon to subtract the positively and negatively weighted diffracted components. This approach is sensitive to the phase response of the etalons, so it is necessary to minimize (or to compensate for) nonlinear phase shifts produced by the etalons and to avoid phase sensitive switching behavior in the etalons.[26]

IV. Self-Aligning Bidirectional Volume Holographic Interconnections

The preliminary analysis of a bidirectional optical interconnection system begins with an explanation of the recording of a hologram by using a phase conjugated reference beam. The 1-D system used in this analysis is presented in Fig. 7 and consists of two lines (planes) of optical neurons which need to be interconnected by a volume hologram. The phase conjugate mirror is used to conjugate the expanding waves emitted from one plane of neurons and retroreflect them back toward the sources from which they emerged, and either direction can be the one chosen to be conjugated. The field emitted by a line of J neurons, separated by D, with an aperture profile $h(x_0,y_0)$, and propagating at an off-axis angle with spatial frequency α is a linear combination of off-axis spherically expanding waves propagating toward the right. The undiffracted field that passes straight through the volume hologram (in the undepleted pump approximation) and

strikes the phase conjugate mirror (PCM) is simple in the Fraunhofer regime of the individual apertures, which for an aperture profile width of d is valid for $z \gg \pi d^2/\lambda$, ($z \gg 1$ mm for a 10-μm aperture) and is always valid for Gaussian apertures:

$$A(x',y',z_s,t) = \exp(-i2\pi\nu t)\frac{\exp\left(i\frac{2\pi}{\lambda}z_s\right)}{i\lambda z_s}\int\int\sum_j \mathbf{a}_j h(x_0 - jD,y_0)\exp(i2\pi\alpha x_0)\exp\{i(\pi/\lambda z_s)[(x_0 - x')^2 + (y_0 - y')^2]\}dx_0 dy_0$$

$$\approx \exp(-i2\pi\nu t)\frac{\exp\left(i\frac{2\pi}{\lambda}z_s\right)}{i\lambda z_s}\sum_j \mathbf{a}_j\exp(i2\pi jD\alpha)\exp\{i(\pi/\lambda z_s)[(x' - jD)^2 + y'^2]\}H\left(\frac{x' - jD}{\lambda z_s} + \alpha, \frac{y'}{\lambda z_s}\right). \tag{11}$$

In this expression the neural activity pattern vector $\mathbf{a}_j$ is represented as a spatially multiplexed array of etalon output fields which will be used for learning, but the intensities $|\mathbf{a}_j|^2$ are a more likely representation of the neuron outputs that will be used for subsequent nonlinear processing.

Each source produces an off-axis expanding spherical wave with a linear phase factor given by the source position. The Fourier transform of a source aperture is given by $H(u,v)$ and its size and position shift with the propagation distance z_s. The distance between the input plane of neurons and the front size of the volume hologram is z_0, the thickness of the hologram is L, the index of refraction of the photorefractive medium is n_0, and the distance from the hologram to the

This equation represents a left propagating quadratically curved superposition of waves that are focusing toward the J source neurons. The profiles of these focusing beams are given by the transform of the individual source apertures $H(u + \alpha',v)$ that are scaled and shifted with the z coordinate within the holographic medium.

Similarly, the backpropagating error field emitted by a line of K neurons separated by D' in the second layer can be described as a superimposition of spherically expanding waves. The separation between the output neurons and a plane z within the hologram is $\bar{z} = z_1 + n_0 z$, which is a reversed coordinate from that used for the forward-propagating wave:

$$B(x,y,z,t) = \exp(-i2\pi\nu t)\frac{\exp\left(i\frac{2\pi}{\lambda}\bar{z}\right)}{i\lambda\bar{z}}\int\int\sum_k \mathbf{b}_k h(x_1 - kD',y_1)\exp(i2\pi\alpha x_1)\exp\{i(\pi/\lambda\bar{z})[(x_1 - x)^2 + (y_1 - y)^2]\}dx_1 dy_1$$

$$\approx \exp(-i2\pi\nu t)\frac{\exp\left(i\frac{2\pi}{\lambda}\bar{z}\right)}{i\lambda\bar{z}}\sum_k \mathbf{b}_k\exp(i2\pi kD'\alpha)\exp\{i(\pi/\lambda\bar{z})[(x - kD')^2 + y^2]\}H\left(\frac{x - kD'}{\lambda\bar{z}} + \alpha, \frac{y}{\lambda\bar{z}}\right). \tag{13}$$

phase conjugate mirror is z_c; thus the total optical path length between the neurons and the PCM is $z_s = z_0 + n_0 L + z_c$. This wavefront is phase conjugated by the PCM, which retroreflects each expanding spherical wavefront back toward its point of origin. The resulting field within the holographic crystal is dependent on the z coordinate, and since both writing waves are incident on the hologram from the right, z is defined to be zero at the right edge of the crystal and increased to L at the left edge. The phase conjugated reference wave within the hologram is most easily expressed in terms of the optical path length between the input neurons and a given plane within the hologram, $z' = z_0 + (L - z)n_0$.

$$A^*(x,y,z,t) = \exp(-i2\pi\nu t)\frac{\exp\left(-i\frac{2\pi}{\lambda}z'\right)}{-i\lambda z'}\sum_j \mathbf{a}_j^*\exp(-i2\pi jD\alpha)$$

$$\times \exp\{-i(\pi/\lambda z')[(x - jD)^2 + y^2]\}$$

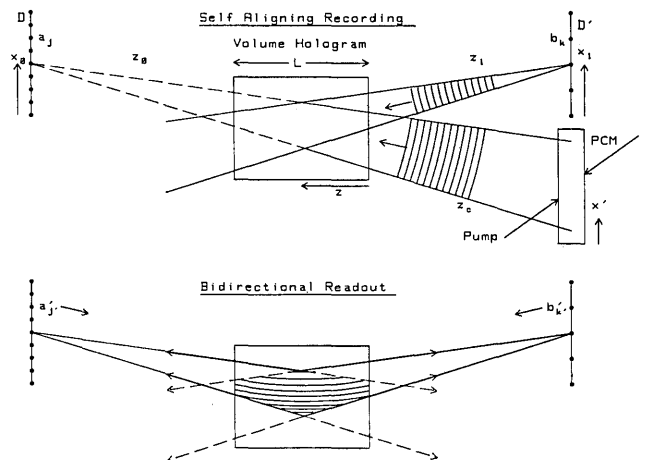$$\times H^*\left(\frac{x - jD}{\lambda z'} + \alpha, \frac{y}{\lambda z'}\right). \tag{12}$$

Fig. 7. Self-aligning bidirectional dynamic volume holographic interconnection using a phase conjugated reference

Fig. 8. Diffracted spot produced by a high diffraction efficiency lensless Fresnel volume hologram recorded in LiNbO₃ and an exposure which shows just the peak.

(a)

(b)

This backpropagating error vector $\mathbf{b}_k$ produces a wavefront that interferes with the retroreflected phase conjugate wavefronts due to the forward-propagating sources in the first layer within the volume hologram. This records a self-aligning interference pattern that modulates the index of refraction within the holographic medium. The modulation term can be expressed as a curved and chirping fringe pattern within the overlap region of the diffracting wavefronts in the crystal. The repetitive presentation of training patterns and bipolar error patterns to the front and back of the single layer being described will result in the time integration of successive outer product connectivity patterns:

$$T(x,y,z,t) \propto \int_0^t \mathcal{R}[A^*(x,y,z,t')B^*(x,y,z,t')]dt'$$

$$= \sum_j \sum_k \left[\int_0^t a_j^*(t')b_k^*(t')dt' \right] \frac{1}{\lambda^2 z'\bar{z}} \mathcal{R}\left[H^*\left(\frac{x-kD'}{\lambda\bar{z}} + \alpha, \frac{y}{\lambda\bar{z}} \right) H^*\left(\frac{x-jD}{\lambda z'} + \alpha, \frac{y}{\lambda z'} \right) \right]$$

$$\times 2\cos\left(\frac{2\pi}{\lambda} \left\{ z_t + \frac{[(x-jD)^2 + y^2]}{z'} + \frac{(x-kD')^2 + y^2}{\bar{z}} + \lambda(jD + kD') \right\} \right). \tag{14}$$

This expression represents a superposition of KJ families of elliptical fringes within the volume of the hologram with each pair of sources at the foci of a set of elliptical shells, and $z_t = z_0 + n_0L + z_1$ is the total distance between the etalon planes. This time-integrated interference pattern will be transformed into a proportional index modulation with a possible phase shift through the photorefractive effect.

For the chirped and curved volume Fresnel phase holograms being analyzed here a momentum space analysis is inappropriate, since spatial frequency and fringe tilt are spatially varying, resulting in a poorly defined perturbation momentum vector; thus an explicit integration of the diffracted field produced at each z-plane should be carried out instead. After the hologram is recorded, it is reilluminated by a weighted superposition of expanding spherical waves which are diffracted by all the index modulations that are present. This analysis can be carried out for either forward- or backward-propagating waves in an identical manner, but we only consider illumination with a forward-propagating wave here. When the volume hologram is illuminated by the diffracted wavefront from a new input neural activity pattern a'_j, the diffracted field at each plane z will contain a matched term, which will produce focusing wavefronts propagating toward the output neurons and a number of
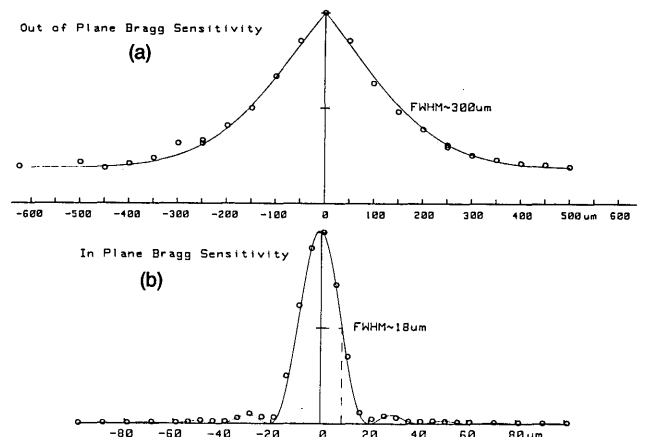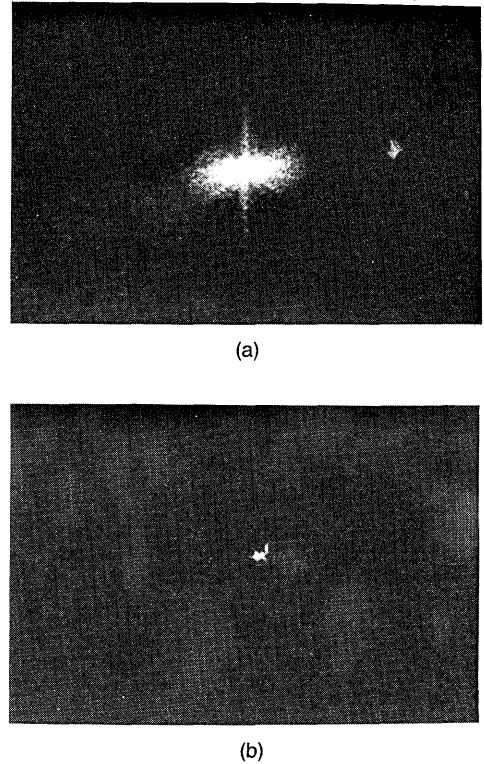

Fig. 9. Positional sensitivity of the Fresnel volume hologram: (a) out of the primary interaction plane; (b) in the interaction plane.

unwanted crosstalk terms. Combining the results of Eqs. (12) and (13) we can derive an expression for the diffracted field produced at each plane of the volume hologram:

$$D(x,y,z,t) \propto A'(x,y,z,t) \int_0^t \mathcal{R}[A^*(x,y,z,t')B^*(x,y,z,t')]dt'$$

$$\propto T(x,y,z) \exp(-i2\pi\nu t) \frac{\exp\left(i\frac{2\pi}{\lambda}z'\right)}{i\lambda z'} \sum_{j'} \mathbf{a}'_{j'} \exp(i2\pi j'D\alpha) \exp\{i(\pi/\lambda z')[(x - j'D)^2 + y^2]\}H\left(\frac{x - j'D}{\lambda z'} + \alpha, \frac{y}{\lambda z'}\right)$$

$$= \exp(-i2\pi\nu t) \frac{1}{\lambda^2 z'^2} \sum_{j'} \alpha'_{j'} \sum_j \sum_k \int_0^t a_j^*(t')b_k^*(t')dt'$$

$$\times \frac{\exp\left(-i\frac{2\pi}{\lambda}\bar{z}\right)}{-i\lambda\bar{z}} \exp\{i2\pi\alpha[(j' - j)D + kD']\} \exp\{i(\pi/\lambda z')[(j'^2 - j^2)D^2 - 2Dx(j - j')]\} \exp\{-i(\pi/\lambda\bar{z})[(x - kD')^2 + y^2]\}$$

$$\times H^*\left(\frac{x - kD}{\lambda\bar{z}} + \alpha, \frac{y}{\lambda\bar{z}}\right)H\left(\frac{x - j'D}{\lambda z'} + \alpha, \frac{y}{\lambda z'}\right)H^*\left(\frac{x - jD}{\lambda z'} + \alpha, \frac{y}{\lambda z'}\right). \tag{15}$$

With the approximations that the last two terms approximately overlap so that the product is equal to a constant, and the paraxial approximation, the diffracted field can be propagated to the plane of output neurons from any diffracting plane within the volume hologram. This is assuming the undepleted pump approximation, which is a reasonable approximation for diffraction efficiency $\leq 10\%$. Each plane of infinitesimal thickness produces an appropriate focusing contribution, with the appropriate focal length, magnification, and phase to produce a focal spot with profile $h(x_1 - kD' - (j - j')D(\bar{z}/z'),y_1)$. We need to sum up all the contribution throughout the thickness L of the hologram to obtain a Bragg selection condition, which will require $j = j'$, so that focal spots are only produced at each of the K output neurons:

In this equation a number of simplifications have been made, but the neglected terms will lead to an increase in the Bragg selectivity. All the phase factors have been lumped into the term $\exp(i\varphi)$. The integration over z produces a sinc function of $(j - j')$, which is analogous to the thick hologram Bragg condition for these elliptical fringes. As long as the separations between the input neurons D, and output neurons D', are large enough for a given hologram thickness L and recording geometry, z_0,z_1,α, we can assume perfect Bragg selectivity, and, therefore, $j = j'$. This would be the normal Bragg condition if Fourier lenses were inserted in the processor, and its results in a positional selectivity in the plane of the lines of neurons which effectively eliminates all the unwanted shift-invariant crosstalk terms that are present with a thin holographic grating.

$$e(x_1,y_1,t) \propto \int_0^L \left[\frac{\exp\left(i\frac{2\pi}{\lambda}\bar{z}\right)}{i\lambda\bar{z}} \int\int D(x,y,z,t) \exp i(\pi/\lambda\bar{z})[(x_1 - x)^2 + (y_1 - y)^2]\}dxdy\right]dz$$

$$\propto \exp(-i2\pi\nu t) \exp(-i2\pi\alpha x_1) \exp(i\varphi_0) \sum_{j'} a'_{j'} \sum_j \sum_k \int_0^t a_j^*(t')b_k^*(t')dt'$$

$$\times \int_0^L h\left(x_1 - kD' - (j - j')D\frac{z}{z'}\right) \exp\left[-i\alpha D(j - j')\frac{z}{z'}\right]dz$$

$$\approx \exp(-i2\pi\nu t) \exp(-i2\pi\alpha x_1) \exp(i\varphi') \sum_{j'} a'_{j'} \sum_j \sum_k \int_0^t a_j^*(t')b_k^*(t')dt'$$

$$\times h\left(x_1 - kD' - (j - j')D\frac{z_1}{z_0 + nL}\right)L \operatorname{sinc}\left[\frac{nL\alpha z_t D(j - j')}{(z_0 + Ln)^2}\right]$$

$$\propto \exp(-i2\pi\nu t) \exp(-i2\pi\alpha x_1) \exp(i\varphi) \sum_j a'_j \sum_k \left[\int_0^t a_j(t')b_k(t')dt'\right]h(x_1 - kD',y). \tag{16}$$

For a hologram thickness $n_0L = 1$ cm, placed 1 cm off-axis and 10 cm from input and output neural planes we can separate the etalons by $D \approx 100$ μm, (at the fourth zero of the sinc), allowing 100 to be packed per centimeter. However, for a 90° diffraction angle, as illustrated in Fig. 6, we can bring the etalons to within $D \approx 10$ μm of each other allowing a linear packing density of 1000 etalons/cm. When 2-D arrays of neurons are to be interconnected additional constraints must be imposed on their topology to achieve the appropriate selectivity of the diffracted orders, and only a spare selection of a 2-D array of etalons may be utilized, containing between 10^4 and 10^5 etalons/in.[2,23] An appropriate topology for the utilized etalons can be derived by considering the interconnection to be space variant in the interaction plane and space invariant in the orthogonal dimension. In the case of a Fresnel volume hologram there is some space-variant widening of the impulse response in the direction orthogonal to the interaction plane, which is not present with Fourier holograms.

The resulting field incident on the spatially multiplexed output neurons is found to be proportional to the desired matrix vector product of the input activity pattern a_j' with the time-integrated outer product of the sequence of forward and backward waves. The intensity at the output neural plane is given by the modulus squared of the field, and this intensity will be detected by the neurons and used for subsequent nonlinear processing.

This is a fully self-aligning dynamic volume holographic global interconnection scheme which works reciprocally for forward and backward waves as required by the backpropagation algorithm. This interconnection technique requires no lenses, because the Fresnel hologram accomplishes the imaging operation.

V. Experimental Investigation of Fresnel Volume Hologram Interconnections

Volume Fresnel holograms were recorded in photorefractive crystals as the interference pattern between expanding and collapsing spherical waves to test their capabilities as lensless interconnection elements. When this volume hologram was reilluminated by one or more of an array of expanding reference beams, the collapsing spherical object waves were reproduced, which focused to an array of small spots at the output plane. First, a single input point was interconnected to a single output point using an expanding wave interfered with a collapsing wave in the volume of a LiNbO$_3$ crystal, and the interference pattern was time integrated for several minutes to build up a reasonably high diffraction efficiency grating. A magnified image of the diffracted focal spot which was produced at the output plane when the crystal was reilluminated by the expanding reference beam is shown in Fig. 8. A good focused spot is produced, but when the film is overexposed a large amount of sidelobe structure becomes apparent. A large amount of fanning of the diffracted light is produced in the plane of the crystal c axis due to the recording of additional gratings in the crystal be-

tween the reference beam and scattered object beam. A vertical line appears at the output plane which is due to the Bragg matched diffraction of the reference beam by the gratings formed between the scattered reference with the object beam, and this line is actually a small part of a large circle of confusion which passes through the reference source and the object virtual source. These fanning components built up over a longer time scale than the desired focusing diffracted light and were not visible with short holographic exposures. The weak additional spot is a multiple reflection artifact. A measurement of the Bragg positional sensitivity in and out of the principal interaction plane is shown in Fig. 9. In the plane a good approximation to a sinc function with 21-μm width is obtained, which is near the expected width for this experimental geometry.

The measurement was obtained by translating the Fresnel volume hologram and measuring the resulting diffraction efficiency. However, when the hologram was rotated in the plane, and any residual translation was compensated, a diffraction efficiency was measured that was essentially independent of angle, as expected for these angularly diverse volume holograms. Out of the interaction plane the Fresnel hologram diffraction efficiency was quite insensitive to the hologram position. However, the position of the diffracted focused spot translated across the detector array as the hologram was moved, indicating that the holographic interconnection was space invariant in this dimension. A Fresnel hologram that is thick in relation to the separation between planes produces a vertically widening impulse response as the out-of-plane offset is increased due to different offset magnifications at different hologram depths. This feature needs to be considered when selecting a 2-D neuron array topology for use with the Fresnel hologram interconnection scheme.

An optical neural network interconnection pattern requires many point sources to be imaged to many other virtual sources, and the Fresnel hologram was tested in this application by using lenslet arrays for the optical sources. A line array of real sources produced by a 1-D lenslet array was interconnected to a 2-D array of virtual sources that was produced by imaging the focal plane of a lenslet array through and beyond the volume hologram. In this $N \rightarrow N^2$ interconnection experiment approximately fifty sources were interconnected with a 50 × 50 array of output focal spots, thereby implementing more than 10^5 holographic interconnection lenses. A small portion of the diffracted output plane produced by this hologram when it was illuminated by the object wave is shown in Fig. 10. This looks almost identical to the image of the lenslet array produced by the object beam, and no fanning artifacts like those shown in Fig. 8 are visible. This is because the diffraction efficiency of each interconnection hologram is extremely small in this case, and the weak fanning artifacts produced by different sources do not add up constructively.

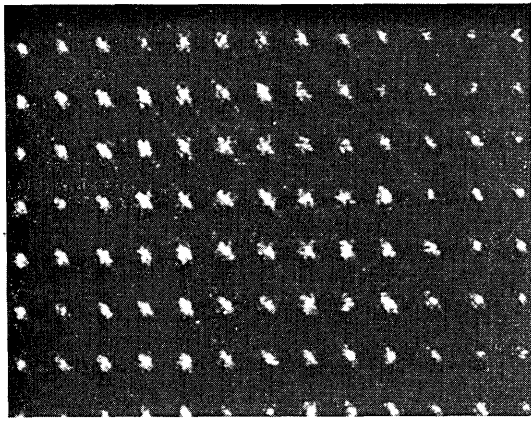Adaptive holographic interconnection networks

Fig. 10. Diffracted output produced by an $N \rightarrow N^2$ lensless holographic interconnecton.

must be able to represent bipolar weights and to decrease interconnection strengths, so selective erasure was examined as one possible technique that can be used for both these purposes. To show that Fresnel gratings could be selectively erased, $Bi_{12}SiO_{20}$ was used as the holographic recording medium so that photovoltaic effects could be eliminated, and faster response times could be obtained. A piezoelectric mirror was used to phase modulate an object beam with a π phase shift increment. Interconnection gratings were built up with one phase, then the object phase was shifted, and the diffraction efficiency of the reference beam into the object focal spot was measured as the hologram was erased and rewritten with a π phase shift. Because different wavelength probe beams cannot be used to measure the diffraction efficiency of Fresnel holograms, the object beam and diffracted beam were alternatively chopped in a nonoverlapping fashion to measure the diffraction efficiency seen by the reference beam as a function of time. An example of this type of selective erasure process is shown in Fig. 11(b), and it is to be compared with the incoherent erasure that was obtained by blocking the reference beam as shown in Fig. 11(a). The selective erasure was much faster than the incoherent erasure process, or the writing process after the previous grating was erased, because the incoherent erasure and phase shifted writing are cooperating processes during selective erasure, while they are competing processes when writing the hologram. The phase could be repetitively shifted by π as shown in Fig. 11(c), and a succession of out-of-phase gratings can be written and erased. Other gratings within the crystal were not erased any faster with this phase shifted reference approach than they were normally by incoherent erasure, which demonstrates that selective erasure of the Fresnel hologram is occurring throughout the volume of the crystal.
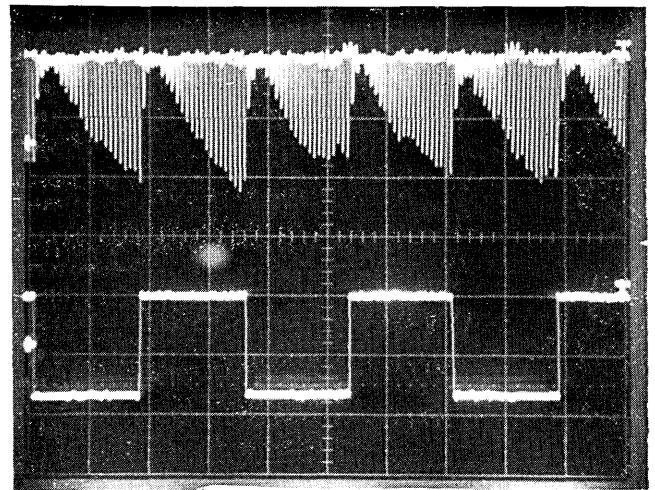
Polarization switching diffraction can be demonstrated in $Bi_{12}SiO_{20}$ by writing a grating in the 110 direction.[27] The propagating eigenmodes are circular without an applied field, because of optical activity, and the right mode can be coupled to the left mode
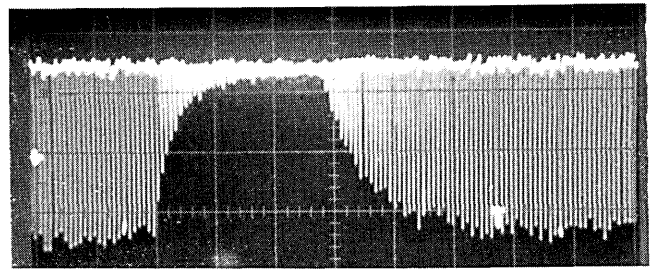
(a)
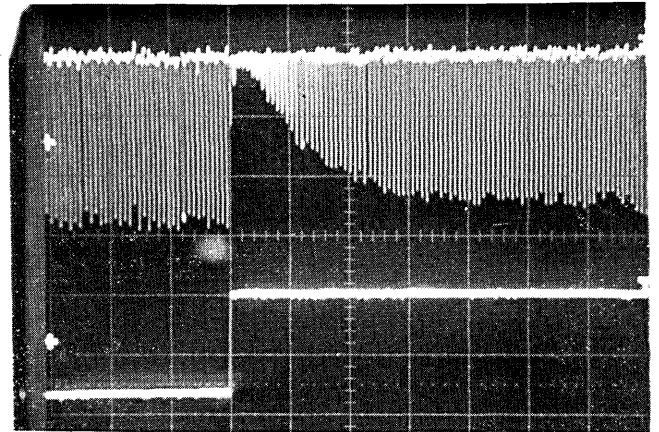
(b)

(c)

Fig. 11. Erasure processes in $Bi_{12}SiO_{20}$: (a) incoherent erasure process; (b) selective erasure process using a π phase shifted reference and the phase shift signal; (c) repetitive π phase shift writing and erasure (1 s/div).

through the off diagonal tensor components of a photorefractively induced pertubation grating. Right circularly polarized expanding spherical reference waves were interfered with right circularly polarized focusing object beams to record polarization switching Fresnel volume holograms in a properly rotated $Bi_{12}SiO_{20}$ crystal. The diffracted field focused to the object beam focal spot, and the polarization state was analyzed with a properly rotated quarterwave plate and polarizer, and it was found to be very nearly orthogonal to the

polarization of the input object beam. The orthogonality of these modes can be improved by careful alignment of the crystal axes with respect to the principal interaction plane, but perfect orthogonality is probably impossible with a Fresnel hologram because of the angular diversity of the birefringent gratings. When 99% of the object beam can be filtered with this polarization filtering scheme, a 5% diffraction efficiency hologram results in a 20% feedthrough of the undiffracted light measured with respect to the diffracted polarization switched light. This suppression ratio of the undiffracted beam must be improved if the polarization multiplexed architecture is to be used for a backpropagation network, so that the undiffracted phase conjugate reference will not corrupt the diffracted backpropagating error.

VI. System Requirements

A complete two-layer system, illustrated in Fig. 12, requires a high speed method of entering data for pattern transformation processing and another means of introducing the backward-propagating error signals for the learning phase. Probably the best approach to high speed data entry at the back end of the system is a sparse parallel laser diode array or a fiber-optic input array, demagnified onto the first layer nonlinear etalon array which is operated in the bistable regime. In this manner the subthreshold coherent bias beams transmitted by each addressed device can be modulated by the data signals, thereby using the input nonlinear Fabry-Perot etalon array as a high speed incoherent to coherent converter with memory. At the final layer of the system error signals need to be computed and injected back into the system with the appropriate polarization or wavelength and the phase shift or timing needed to represent the sign of the error. The system can be designed with either optical or electronic error detection and generation circuitry at the output to introduce the backpropagating error. Optical subtraction techniques can be considered for an optical approach to teaching the system. Image subtraction using a phase conjugated Michelson interferometer[28] appears to be a promising approach for this application since it produces subtracted fields with the appropriate phase shift to represent the sign of the error, without the accurate phase adjustments required by other interferometric approaches to image subtraction. Alternatively, since the computational load required at the output is relatively minor, optical detectors can be combined with electronic subtraction from the target vector to generate the bipolar error vector, which can be applied to a spatial light modulator at the output to introduce the backpropagating error. When the number of outputs of the pattern transformation procedure is <1000, they can be arrayed in a linear format which allows the utilization of high speed linear detector arrays for output, and the utilization of linear spatial light modulators, to introduce the backward-propagating error signals.

The fan-out capability of each layer is determined by the gain of the nonlinear devices, the holographic diffraction efficiency, and the polarization component throughput, and it will dictate an information collapsing network architecture. For example, if the product of optical efficiencies is only 3%, a network with 30,000-bit input pattern vectors might be processed by 1000 hidden units that communicate with thirty output devices, which simplifies the error generation process at the output. The ability of the system to process large amounts of data in parallel at a very high speed is limited by the electronic addressing of the input array, and the output photodetector array readout time, and not by the intervening optical system, because of the extremely fast response achievable with nonlinear etalons and the almost instantaneous optical interconnection delay. The optical power requirements of the system are primarily dictated by the first layer of nonlinear etalons, since there are many more in this layer than in the succeeding layers for a collapsing network. The first-layer etalons are not bidirectional and can be optimized to have a low switching energy. Bistable nonlinear etalons have been operated with a 3-pJ switching energy at a rate of ~100 MHz,[29] which leads to a power requirement of 0.4-mW/etalon or 12 W for 30,000 input etalons. Only a portion of this power is dissipated within the nonlinear etalons, and a heat dissipation requirement of only a few watts per cm^2 should be achievable with forced air or liquid cooling techniques. Most of this power is supplied by a high power coherent pump beam that is used to bias each bistable device just below the bistable loop, and ~10%, or 40 μW, is required per laser or fiber-optic input to
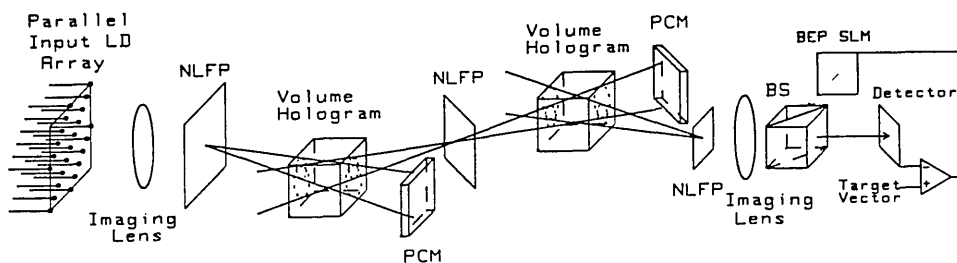

Fig. 12. Complete system for two-layer backpropagation optical learning including massively parallel input laser array and electronic error detection at the output. LD = laser diode (or fiber optics), NLFP = nonlinear Fabry-Perot, PCM = phase conjugate mirror, BEP SLM = spatial light modulator for backward-propagating error.

the network. The backpropagating neurons will require more optical power because the dual functions of the bidirectional cavities conflict with the requirements for a low power device. Since there are not as many hidden and output devices as input devices the systems power requirements are primarily dictated by the size of the input array.

The forward-propagating signal can be a narrow pulse since the response of GaAs nonlinear Fabry-Perot etalon is determined by the peak power incident. In this case the backward-propagating error signal can be either pulsed or cw. In the pulsed mode the PCM would need to have practically instantaneous response, such as a nonlinear optical semiconductor might provide, and the forward- and backward-propagating pulses could be time jittered so they do not overlap in the volume hologram, but the phase conjugate reference and the backward-propagating error pulse would overlap within the crystal, thereby exposing a hologram. Alternatively, the backward-propagating error signal could be a low power cw beam that would not have a high enough peak power to nonlinearly modify the index within the Fabry-Perot etalons, and the forward-propagating pulse could be turned into a quasi cw phase conjugated reference by using a photorefractive crystal-based PCM which has a slow integrated response. The Fabry-Perot etalons would need to have a slow relaxation time of the nonlinearly shifted index (this requires long carrier life times and should lead to lower etalon switching energies); thus the probe beam would have the appropriate response for most of the interval between pulses of the forward beams. In this case the holographic exposure would be due to the time integral of the cw waves in the volume hologram, and the orthogonally polarized and pulsed forward-propagating beam would not contribute significantly to the hologram exposure.

VII. Conclusion

The 3-D storage capacity of volume holograms allows the construction of huge globally interconnected multilayer optical networks which are well beyond the projected capabilities of alternative technologies. The optical system seems well matched to the bidirectional requirements of a backpropagation learning system because of the intrinsically reciprocal nature of optical interconnections. Error driven learning operations, such as backpropagation, should be able to compensate for many of the technological flaws inherent to an optical implementation by adaptively sensing the misbehavior of the system and driving it in the appropriate direction necessary to overcome its imperfections. The nonideal optical implementation of a backpropagation network may actually have improved performance over that of an idealized digital simulation because noise will always be present in the system, helping it to avoid shallow local minima, and pushing the interconnection matrix away from solution boundaries. Imperfections of the holographic interconnection will help the system perform symmetry breaking, which the idealized model cannot perform spontane-

ously. The simultaneous self-aligning and learning of the optical system make this approach to multilayer optical neural processing experimentally feasible and allow the implementation of complicated systems that could not be completely specified *a priori* but can be learned and modified as the desired processing operation slowly changes. The slow learning of the holographic crystals combined with the extremely high speed processing of the nonlinear etalons gives this system an enormous throughput potential and the capability for solving complicated cognitive problems.

References

1. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Internal Representations by Error Propagation," in *Parallel Distributed Processing, Vol. 1,* D. E. Rumelhart and J. L. McClelland, Eds. (MIT Press, Cambridge, MA, 1986), Chap. 8.
2. D. B. Parker, "Learning Logic," Invention Report S81-64, File 1, Office of Technology Licensing, Stanford U. (Oct. 1982).
3. J. J. Hopfield, "Neurons with Graded Response have Collective Computational Properties like those of Two-State Neurons," Proc. Natl. Acad. Sci. USA 81, 3088 (1984).
4. S. Grossberg, *Studies of Mind and Brain* (Reidel, Boston, 1982).
5. T. Kohonen, *Self-Organization and Associative Memory* (Springer-Verlag, Berlin, 1984).
6. Y. S. Abu-Mostafa and D. Psaltis, "Optical Neural Computers," Sci. Am. 256, 88 (1987).
7. 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, 98 (1985).
8. D. Z. Anderson, "Coherent Optical Eigenstate Memory," Opt. Lett. 11, 56 (1986).
9. B. H. Soffer, G. J. Dunning, Y. Owechko, and E. Marom, "Associative Holographic Memory with Feedback Using Phase-Conjugate Mirrors," Opt. Lett. 11, 118 (1986).
10. A. Yariv and S. Kwong, "Associative Memories Based on Message-Bearing Optical Modes in Phase Conjugate Resonators," Opt. Lett. 11, 186 (1986).
11. T. Jannson *et al.,* "The Interconnectability of Neuro-Optic Processors," Proc. Soc. Photo-Opt. Instrum. Eng. 698, 157 (1986).
12. A. D. Fisher et al., "Implementation of Adaptive Associative Optical Computing Elements," Proc. Soc. Photo-Opt. Instrum. Eng. 525, 196 (1986).
13. M. Cohen, "Design of a New Medium for Volume Holographic Information Processing," Appl. Opt. 25, 2288 (1986).
14. N. Farhat, "Architectures for Opto-Electronic Analogs of Self-Organizing Neural Networks," in *Technical Digest of Topical Meeting on Optical Computing* (Optical Society of America, Washington, DC, 1987), p. 125.
15. K. Wagner and D. Psaltis, "Multilayer Optical Learning Networks," Proc. Soc. Photo-Opt. Instrum. Eng. 752, 16 (1987).
16. K. Wagner and D. Psaltis, "Multilayer Optical Learning Networks," in *Technical Digest, Topical Meeting on Optical Computing* (Optical Society of America, Washington, DC, 1987), p. 133.
17. D. Psaltis and C. Park, "Nonlinear Discriminant Functions and Associative Memories in *Proceedings, Conference on Neural*

Networks for Computing, Snowbird, UT, APS Conf. Proc. 151 (1986).

18. T. J. Sejnowski and C. R. Rosenberg, "NETtalk: a Parallel-Network that Learns to Read Aloud," John Hopkins U., JHU/EECS-86/01 (1986).

19. B. Widrow and M. E. Hoff, "Adaptive Switching Circuits," IRE Wescon Conv. Rec. 4, 96 (1960).

20. H. M. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic, New York, 1985).

21. H. M. Gibbs, *et al.*, "Optical Modulation by Optical Tuning of a Cavity," Appl. Phys. Lett. 34, 511 (1979).

22. A. W. Lohmann, "Polarization and Optical Logic," Appl. Opt. 25, 1594 (1986).

23. D. Psaltis *et al.*, "Optical Neural Nets Implemented with Volume Holograms," in *Technical Digest Topical Meeting on Optical Computing* (Optical Society of America, Washington, DC, 1987).

24. J. P. Huignard *et al.*, "Coherent Selective Erasure of Superimposed Volume Holograms in LiNbO$_3$," Appl. Phys. Lett. 26, 256 (1975).

25. D. L. Stabler *et al.*, "Multiplier Storage and Erasure of Fixed Holograms in Fe-Doped LiNbO$_3$," Appl. Phys. Lett. 26, 182 (1975).

26. J. D. Cresser and P. Meystre, "The Role of Phases in the Trasient Dynamics of Nonlinear Interferometers," in *Optical Bistability*, C. M. Bowden Ed. (Plenum, New York, 1980).

27. A. Merrakchi, R. V. Johnson, and A. Tanguay, Jr., "Polarization Properties of Photorefractive Diffraction in Electrooptic and Optically Active Sillenite Crystals (Bragg Regime)," J. Opt. Soc. Am. B. 3, 321 (1986).

28. A. E. Chiou and P. Yeh, "Parallel Image Subtraction Using a Phase Conjugate Michelson Interferometer," Opt. Lett. 11, 306 (1986).

29. J. L. Jewell *et al.*, "3pJ 82MHz Optical Logic Gates in a Room Temperature GaAs-AlGaAs Multiple Quantum Well Etalon," Appl. Phys. Lett. 46, 918 (1985).

Meetings Calendar

**1987
December**

1-3 Optical Information Systems Conf. & Exhibition, New York *J. Emard, 11 Ferry Lane West, Westport, CT 06880*

6-9 Electron Device Int. Mtg., Wash., DC *M. Widerkehr, Courtesy Assocs. Inc., 655 15th St., N.W., Wash., DC 20005*

7-11 Fundamentals & Applications of Lasers course, Albuquerque *Laser Inst. of Am., 5151 Monroe St., Toledo, OH 43623*

7-11 Lasers '87, Lake Tahoe *Lasers '87, P.O. Box 245, McLean, VA 22101*

7-11 Laser Beam Propagation & Interaction Effects course, Denver *Eng. Tech. Inst., P.O. Box 8859, Waco, TX 76714*

7-11 NAVSTAR Global Positioning System: Operation, Implementation & Applications course, Los Angeles *UCLA Ext., P.O. Box 24901, Los Angeles, CA 90024*

9-11 Fiber Optic Communications course, Tempe *Ctr. for Professional Development, Coll. of Eng. & Appl. Sci., Arizona State U., Tempe, AZ 85287*

9-15 Optics & Glass Expo/China '87, Beijing *China Promotion, Ltd., Rm. 1810 Shum Tak Centre, Office Tower, 200 Connaught Rd., Central, Hong Kong*

9-15 Fiber Optic Communications course, Tempe *Arizona State U., Elect. & Computer Eng. Dept., Ctr. for Professional Dev., Tempe, AZ 85287*

15-17 Solid State Lasers course, Williamsburg *C. Keen, Science & Tech. Inst., 101 Research Drive, Hampton, Va. 23666*

**1988
January**

4-5 Low Noise Amplifier Design course, Lake Buena Vista *V. Amico, U. Central Florida, Orlando, FL 32816*

4-5 Laser Principles Sem., Lake Buena Vista *V. Amico, U. Central Florida, Orlando, FL 32816*

5-16 Optical Science & Engineering course, Tucson *P. Slater, P.O. Box 18667, Tucson, AZ 85731*

6-8 Infrared Detectors & Systems course, Lake Buena Vista *V. Amico, U. Central Florida, Orlando, FL 32816*

6-8 Fiber Optics Workshop & Laboratory, Lake Buena Vista *V. Amico, U. Central Florida, Orlando, FL 32816*

6-8 Mechanical Design for Electro-Optical Systems course, Lake Buena Vista *V. Amico, U. Central Florida, Orlando, FL 32816*