1.0

4.5
5.0

2.8

2.5

3.2

2.2

3.6

1.1

4.0

2.0

1.8

1.25

1.4

1.6

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963

②

AD-A190 030

	Form Approved OMB No. 0704-0188

1b. RESTRICTIVE MARKINGS

UNCLASSIFIED

DTIC
SELECTE
FEB 2 9 1988
H
D

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	Approved for public release, distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
	AFOSR-TR- 88 - 0148

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
VERAC INC		AFOSR/NE

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
P O Box 26669 San Diego, CA 92121-1771	Bldg 410 Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR	NE	F49620-86-C-0070

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO
Bldg 410 Bolling AFB, DC 20332-6448	61102F	2305	B1	

11. TITLE (Include Security Classification)
OPTICAL CONCEPTUAL COMPUTING AND ASSOCIATIVE MEMOARY (OCCAM)

12. PERSONAL AUTHOR(S)
Duent

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final Report	FROM 01/06/86 TO 01/06/87		

16. SUPPLEMENTARY NOTATION

17.	COSATI CODES		18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
The Optical Conceptual Computing and Associative Memory (OCCAM) Program applied the techniques of neural-network dynamical systems analysis and fussy theory to problems in conceptual computing. Research results included both pure theory and optical implementations. OCCAM research was jointly perform by research teams at VERAC Incorporated and at UC San Diego's Electrical Engineering Department.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	UNCLASSIFIED

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL
GILES	(202) 767-4933	NE

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

⬡ **VERAC**
incorporated

OPTICAL CONCEPTUAL
COMPUTING AND
ASSOCIATIVE
MEMORY (OCCAM)

FINAL REPORT

R-088-87

SEPTEMBER 1987

88 2 25 068

TABLE OF CONTENTS

🔺 **VERAC**
Incorporated

TABLE OF CONTENTS (Cont'd)

VERAC
Incorporated

EXECUTIVE SUMMARY

The Optical Conceptual Computing and Associative Memory (OCCAM) Program applied the techniques of neural-network dynamical systems analysis and fuzzy theory to problems in conceptual computing. Research results included both pure theory and optical implementations. OCCAM research was jointly performed by research teams at VERAC Incorporated and at UC San Diego's Electrical Engineering Department.

VERAC's key contributions are detailed in the ten technical papers included in the Appendix. Three fundamental results were: (1) the development of bidirectional associative memories (BAMs), (2) a fuzzy knowledge combination scheme that allows arbitrarily many neural (causal) network "expert systems" from experts with arbitrary credibility to be naturally synthesized into a single, representative associative knowledge network, and (3) the development of pure fuzzy associative memories (FAMs).

UCSD focused on the theory and application of optical neurocomputing. Four fundamental research thrusts were: (1) the design and implementation of optical neural networks, especially optical BAMs, (2) the investigation of physical properties important to neural network implementations, (3) the application of optics to fuzzy knowledge processing and to fuzzy computing (approximate reasoning) in general, and (4) the application of neural network principles to optical pattern recognition. A precursor to an optical continuous BAM, the all-optical Cohen-Grossberg (Hopfield) symmetric unidirectional neural circuit, was implemented. Several holographic and nonholographic BAM systems were devised, with emphasis on volume holography. The fundamental fuzzy operations of minimum and maximum were optically implemented. A backpropagation (hierarchical supervised learning) network was taught to perform rotation-invariant pattern recognition.

⊕VERAC
Incorporated

1

CHAPTER 1: OCCAM OVERVIEW

1.0   The First-Year of OCCAM

The first-year Optical Conceptual Computing and Associative Memory
(OCCAM) Program investigated theoretical and optical properties of
neurocomputing and associative-memory processing. VERAC and UC San Diego
jointly performed the research. A wide variety of research results were
achieved, many of which have since been published.

1.1 OCCAM at VERAC

At VERAC, the primary achievement was the development of a family of
bidirectional associative memory (BAM) architectures. The BAM is the
minimal two-layer nonlinear feedback neural network. The discrete BAM
extends the symmetric unidirectional asynchronous model of Hopfield. A
popular exposition of the discrete BAM appears in the September issue of
Byte. The continuous BAM extends the symmetric unidirectional
differentiable model of Cohen and Grossberg. The adaptive BAM further
extends these results to realtime unsupervised learning. The adaptive BAM
is the first extension of global Lyapunov or "energy function" stability to
learning networks. Unfortunately, it is the only extension compatible with
the standard neuron state-transition model, where each neuron processes a
sum of path-weighted output signals.

At VERAC other results were achieved in associative conceptual computing.
These were in the areas of differential Hebbian learning and fuzzy set
theory. Analytic and simulation properties of the differential Hebbian
learning law, which correlates time derivatives of neuronal outputs, were
explored. Difference learning was applied primarily to causal networks.
The result is a type of inductive inference. The causal analogue of
learning is called adaptive inference. The causal network is a fuzzy signed
digraph with feedback, a fuzzy cognitive map (FCM). FCM connection
strengths are adaptively inferred from data with the differential Hebbian
law, which learns a weighted time average of lagged change. Data generated

VERAC
Incorporated

2

from a FCM was used to "back infer" the original FCM with the differential Hebbian learning law.

More practically, a general scheme was developed for combining arbitrary weighted FCMs from arbitrarily many experts with arbitrary credibility weights. This knowledge combination scheme is a powerful alternative to expert systems technology, which is based on search tree construction. Search trees cannot naturally be combined to yield another tree. To produce a tree, knowledge representation accuracy must be compromised. In contrast, with the FCM combination technique, the Kolmogorov's Strong Law of Large Numbers insures that as expert sample size increases, knowledge representation accuracy increases as well.

Fuzzy theory was investigated in two other ways. First, we proved that most continuous neural networks--BAMs, Boltzmann machines, competitive learning systems, and the brain-state-in-a-box model--proceed in their distributive computation so as to minimize the fuzzy entropy of the input pattern. Such networks continuously and nonalgorithmically disambiguate inputs in the unit hypercubes $[0, 1]^n$ or $[0, 1]^n \times [0, 1]^p$. Second, fuzzy associative memories (FAMs) were constructed. FAMs map fuzzy sets to subsets of stored fuzzy subsets. The storage mechanism is a type of fuzzy Hebb law. VERAC produced ten (10) technical journal articles on the OCCAM Program. These articles are repeated in full below in the Appendix. Details of optical software simulations are provided below.

1.2 OCCAM at UC San Diego

At UCSD, OCCAM research fell into four categories: (1) design and implementation of optical neural networks, (2) investigation of physical properties important to network implementations, (3) application of optics to implementation of fuzzy cognitive maps and (4) application of neural network principles to optical pattern recognition problems. A brief summary of these accomplishments is given below. The following sections contain full technical details.

VERAC
Incorporated

Working in close coordination with VERAC, the bidirectional associative memory the bidirectional associative memory (BAM) was identified as a powerful neural network architecture well-suited for optical implementation. In preparation for implementation of an optical BAM, an all-optical Cohen-Grossberg symmetric unidirectional autoassociator was implemented and evaluated. Also, a variation of the BAM, the Bidirectional Optical Optimal Memory (BOOM), which uses encoding based on Kohonen's optimal linear (pseudo-inverse) associative memory, was investigated and found memory storage efficient. Several designs for optical and opto-electronic BAMs were developed. These include optically programmable electronic crossbars, matrix-vector multiplier BAMs, and volume holographic BAMs.

Physical processes instrumental for implementing holographic BAMs were investigated. Correlation of images using counter-propagating two-wave mixing was demonstrated. Characteristics of volume holographic association of pseudo-random phase-encoded images was studied.

Fuzzy cognitive maps (FCMs) are fuzzy signed directed graphics with feedback, used as an associative nonlinear dynamical system. FCMs are the neural-network paradigm for conceptual computing. They represent complex problem domains where causal connections of varying degree are heavily interlocked. The operations of minimum and maximum play fundamental roles in fuzzy theory. They are the basic operations of intersection/conjunction (AND) and union/disjunction (OR). Optical systems to implement these operations for fuzzy cognitive processing, as in FAMs, were designed and analyzed. For example, the maximum operation

$$\max(x, y) = 1/2[x + y + |x - y|] \; ,$$

is precisely the operation used in competitive learning winner-take-all networks, such as the adaptive resonance theory, Kohonen's self-organizing feature maps, and counter-propagation.

Finally, aspects of the operating system that must be developed to support optical neural networks for real applications was investigated. Preprocessing of raw image data by both conventional and neural network

VERAC
Incorporated

4

techniques was simulated and compared. A backpropagation network was trained for rotation-invariant pattern recognition. This study indicated important results regarding the proper size of network parameters.

⊕VERAC
Incorporated

2.1 Holographic Associative Memory System

Parallels between optical holography and associative memory have been recognized since either field has existed. To implement completely the associative memory system with the optical hologrphy, the following devices are necessary: a storage medium for the holograms and discriminating devices for correct recall from a partial input. Photorefractive crystals are the devices that satisfy the above requirements. In this section, experimental results using photorefractive crystals for an optical associative memory system are described.

1. Counter - Propagating Two Wave Mixing

When two optical waves are incident into the photorefractive crystal, their interference pattern causes a photo-induced space charge pattern. This space charge pattern modulates a refractive index of the crystal through an electro-optic effect, as follows ;

$$n = n_0 + n_I \exp(j\phi) A_1 A_2^* \exp[j(\underline{k_1} - \underline{k_2})\underline{r}] / 2I_0 + C.C$$

where A_i, $k_i (i = 1,2)$; amplitudes & wave vectors of incident

optical waves to the crystal

$$I_0 = |A_1|^2 + |A_2|^2 \qquad n_I = \gamma_{eff} n_0^3 E_{sc} / 2$$

$$\gamma_{eff} = \hat{e}_{1i} \epsilon_i^w (r_{ijk} \ell_{gk}) \epsilon_{jj}^w \hat{e}_{2j} / n_0^3 n$$

$$\underline{k_q} = \underline{k_1} - \underline{k_2}$$

$$\tan\phi = E_d(E_d + E_q) + E_0^2 / E_0 E_q$$

$E_{d,q,0}$; electric field by diffuson & maximum space

charge, and applied field

γ_{eff} ; effective electro-optic coefficient determined

by a crystal orientation, polarizations of incident

two beams and index grating vector

This spatial refractive index grating causes beam coupling in

the crystal. The following coupled wave equations are obtained:

$$2c/\omega \cos\theta \; dA_1/dz = -jn_1 \exp(j\phi)A_1 A_2^* A_2 / I_0$$

$$2c/\omega \cos\theta \; dA_2/dz = -jn_1 \exp(-j\phi)A_1^* A_2 A_1 / I_0$$

Defining a coupling coefficient γ,

$$\gamma = j\omega n_1 \exp(-j\phi) / 2c\cos\theta$$

the above equations become

$$dA_1/dz = \gamma^* A_1 A_2^* A_2 / I_0$$

$$dA_2/dz = -\gamma A_1 A_2 A_1^* / I_0 .$$

Let $A_{12}(z) = A_1(z) / A_2^*(z)$. Then

$$dA_{12}(z)/dz = 1/A_2^* \; dA_1/dz - A_1(z)/[A_2(z)] \; dA_2^*/dz$$

$$= \gamma^*(I_1(z) + I_2(z))A_{12}(z) / I_0$$

$$= \gamma A_{12}(z)$$

where $I_0 = I_{10} + I_{20} = I_1(z) + I_2(z)$

if negligible absorption in the crystal is assumed. The solution
of $A_{12}(z)$ is

$$A_{12}(z) = A_1(z) / A_2^*(z) = A_{12}(0) \exp(\gamma^* z)$$

Thus ,

$$I_1(z) / I_2(z) = I_{12}(z) = A_{12}(z)A_{12}^*(z)$$

$$= I_{10} \exp(2\mathrm{Re}(\gamma)z) / I_{20}$$

$$= I_{10}\exp(\Gamma L) / I_{20}$$

where $\Gamma = 2\mathrm{Re}(\gamma)$, called exponential gain coefficeint

and L is a thickness of the crystal.

This means that the energy of one beam is transferred to the other,

ie, one beam is amplified. This is called Two - Wave Mixing (TWM).

The maximum energy transfer is achieved at $\phi = \pi / 2$. The gain is defined as

$$G_i = I_i \text{ with } I_j / I_i \text{ without } I_j \quad (i = 1,2)$$

which leads to

$$I_1(L) = I_2(L) \, I_{10} \exp(\Gamma L) / I_{20}$$

$$= (I_0 - I_1(L)) I_{10} \exp(\Gamma L) / I_{20}$$

Solving $I_1(z)$, the gain is

$$G_1 = (1 + r) \exp(\Gamma L) / [1 + r \exp(\Gamma L)]$$

where r is the initial intensity ratio of the two beams.

In TWM, the direction of energy transfer is along the c-axis of the crystal. When two optical beams are incident onto the same surface of the crystal, the resultant refractive index grating is a transmissive type. On the other hand, when two beams propagate in opposite directions and the c-axis is oriented as in Fig. 1, the energy transfer still occurs along the c-axis; this is called a Counter - Propagating Two Wave Mixing (CPTWM) . So, the refractive index grating of CPTWM is a reflective type as in the Fig. 1.

Experiments on CPTWM with regular-cut $BaTiO_3$ were performed; its physical dimensions were 5.5mm x 4.35mm x 3.75mm. Fig.2 shows the experimental set-up to characterize CPTWM. The intensity ratio of the pumping beam to the counter-propagating signal beam was 10^4 to 1. The gain was measured with varying θ, the the angle between c-axis and the refractive index grating vector inside the crystal; and with varying β, the angle between two beams inside the crystal.

The maximum gain with varying θ is achieved at $\theta = 88^O$. From the paper of Y. Fainman, et.al., the maximum gain was obtained $\theta = 2^O$, the same angle after the two beams cross each other, ie , $\theta = 2^O = 90^O - 88^O$. Also, the other characteristics of the gain of CPTWM are the same as those for TWM. But in this experiment, the maximum gain with varying β was not obtained

because of the geometical limits of the regular – cut $BaTiO_3$ as in TWM.

In TWM, or CPTWM, optical gain can be achieved only when two beams interact and form a refractive index grating in the crystal whose phase is shifted to that of the interference pattern of the two beams. The gain will depend on the modulation function when the two beams are modulated spatially. For example, if two beams are modulated with orthogonal functions, and therefore they don't interact in the crystal, no gain is achieved. In the next experiment, both the signal beam and the pumping beam in CPTWM were modulated with two identical patterns but one was rotated with respect to the other, as in Fig. 3. A measurement was taken of the gain for the same crystal . The gain was measured to be proportional to the degree of similarity between the two patterns. This means that the nearest one of a set of stored patterns to a given pattern can be maximally amplified through this preferential gain property, and it can be selected through a threshold device.

2. Thick Holograms in $LiNbO_3$

The optical hologram is classified into two kinds (a thin hologram and a thick hologram) by the thickness of the recording medium and by the wavelength of the recording optical beams. When a large amount of data must be recorded, the thick hologram is considered to be better than the thin hologram because of its high storage capacity and low crosstalk noise.

Conventional photographic plates for the holograms need a wet process for development, so that real-time media have been investigated and the photorefractive crystal has been recognized as one of good real-time thick holographic mediums. In the photorefractive crystal, the refractive index is modulated by the inteference pattern formed by two incident beams . Thus, the hologram formed in the crystal is a thick phase hologram.

The thick holograms were recored in $LiNbO_3$ through angular multiplexing. ; the crystal is doped with 0.005% Fe , its size is 10mm x 2mm x 10mm (X x Y x Z) , and the y-faces are polished. Its

VERAC
Incorporated

9

c-axis is along the x-axis. During recording and recalling , it was electrically open-circuited.

To get a high diffraction efficiency, the beam polariztion from the argon laser (its wavelength is 514.5nm) was rotated horizontally and the c - axis of the crytal is also aligned horizontally. The angle between the two incident beams was $18^O - 22^O$. Masks for the signal beam are shown in Fig. 4. Normal patterns are recorded in the crystal; for high modulation index to achieve high diffraction efficiency, intensities of two beam were controlled to be equal to $10mw/cm^2$ and $40mw/cm^2$. The exposure time was 200sec ~ 400sec. The maximum diffraction efficiencies for different exposure energies were approximately constant; at 10^{-3}%. Even though they were very low, the reconstructed patterns could be seen clearly.

The same experimental set-up was used to record Fourier - Transform(F.T.) patterns except lenses for Fourier Transform for record and Inverse Transform for reconstruction were introduced into the system. In the F.T. patterns, the dc - intensity is very much higher than the intensities of high frequencies, so care must be taken for intensity matching between the reference beam and the F.T. pattern. In this experiment, the dc intensity of the patterns was calculated roughly and the intensity of the reference beam was measured with a detector having a pin-hole whose diameter was 0.7mm, so that the intensity matching condition was satisfied by controlling intensities of two beams. The resultant intensities of the beams were $12mw/cm^2$ and $20mw/cm^2$. The noise fringe pattern(~ 6 lines / mm) was caused interference of the incident beam with its reflected beam from the crystal surface. It is inevitable and reduces the recording efficiency. The maximum diffraction efficiency of the F.T. patterns obtained was not significantly different from those of the normal patterns, around 10^{-3}%.

To erase the stored patterns, a uniform beam from the laser, or ultraviolet light from a Mercury lamp illuminated the crystal for two hours. The patterns were totally erased , but a random noise which would be caused by defects in the crystal remained.

3. Thick hologram recording using a pseudo-random phase

mask

When multiple patterns are recorded in a thick hologram, the recording medium should be rotated, or the recording beams should be deflected for angular multiplexing which eliminates the crosstalk noise between the stored patterns. Such methods require additional elaborate optical/electronic systems. To alleviate this requirement, a reference beam is coded with an orthogonal phase funtion , or with a random phase. The thick hologram made with the random phase coded reference beam, in addition, have other benefits ; noise immunity, fault tolerance, and use of the recording medium with a limited dynamic range.

For recording thick holograms in $LiNbO_3$, pseudorandom phase masks (PRPM's) were generated by a laser scanner.

Two kinds of PRPM's were generated with the laser scanner. computer-generated hologram (CGH) type PRPM's and bleached film PRPM's.

In the CGH type PRPM's, an input pattern to the CGH program is a pattern from a random number generator, instead of a pattern from an image. So, the generated output pattern was a CGH of the random numbers. The diffraction efficiency , however, of the PRPM's was too low (< 1%). For the bleached PRPM's by the laser scanner, the exposure time of each scanned pixel on the film was taken from the random number generator, so the phase transmittance of each pixel in the film was random. After that, the film was developed and bleached. A problem of these PRPM's was that they had regular spots after Fourier transformation due to the regularly scanned pixels on the film. So the PRPM's by the laser scanner could not be used for the experiment.

The other way to produce PRPM's was to use a speckle pattern. As the opening aperture of the iris is decreased, there occurs a speckle pattern. A picture of the speckle pattern was taken. Its film was developed and bleached. The resultant film showed an adequate property of PRPM's ; pseudorandom phase coding of patterns with a finite random phase.

With these PRPM's produced by the speckle pattern, Fourier

VERAC
Incorporated

Transform patterns as in Fig. 4 were recorded in the crystal. The dc
intensity of the F.T. pattern with the PRPM was reduced by a factor
of 2 ,compared with the F.T. pattern without PRPM's , while the
intensities of high frequencies were increased. The maximum
diffraction efficiency achieved was around $10^{-3}\%$.

Counter propagating two-wave mixing

Strong Pump Beam

Strengthened Signal Beam

C-axis

Weak Signal Beam

Weakened Pump Beam

rotating polarizer

beam splitter

Ar laser

attenuator

BaTiO3

detector
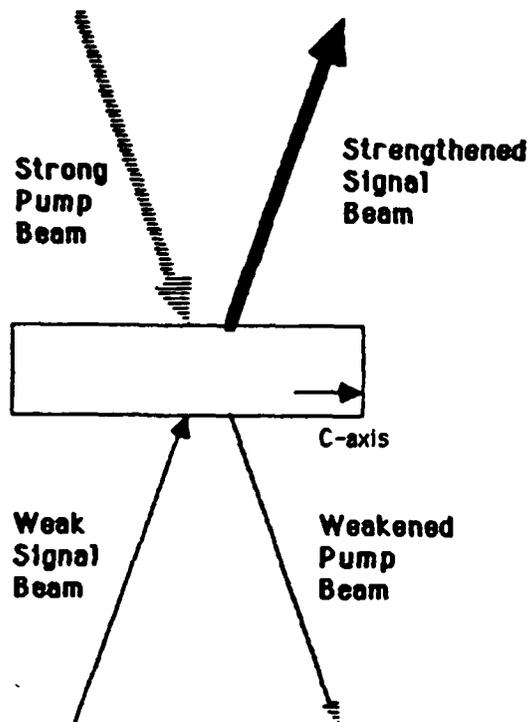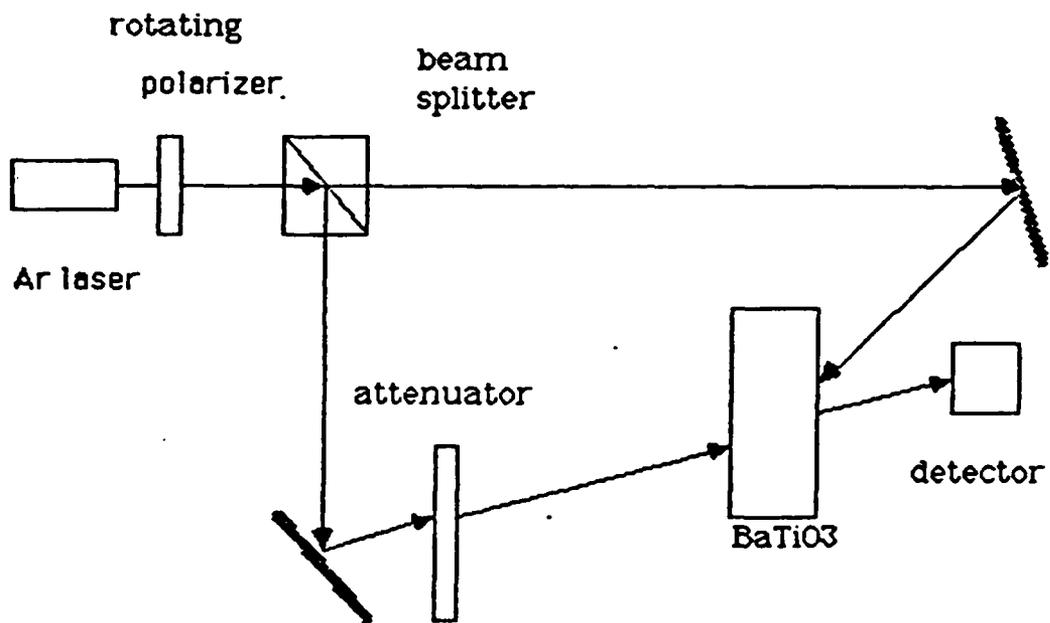
Experimental Set - Up   for CPTWM

VERAC
Incorporated

14
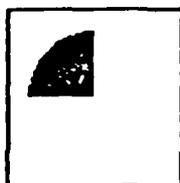
masks for modulation for two beams in CPTWM
(one of them was on the rotational stage)
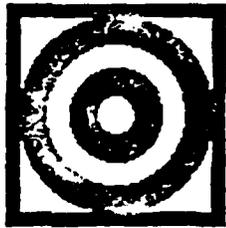
Masks for modulation of the signal beam

2.2 Optical Associative Memories

Storage Algorithms

Another research interest concerned an idea for an alternative storage prescription to the sum of outer products (SOP) method usually used with the HM. The problems with SOP are well known and include low storage capacity and recall of spurious states. The approach here was to use the Optimal Associative Mapping (OAM) described by Kohonen. Briefly, the OAM maps one set of vectors into another optimally in a least squares sense. Specifically, the OAM which optimally maps a set of column vectors $X = (\underline{x}_1,...,\underline{x}_n)$ to a set $Y = (\underline{y}_1,...,\underline{y}_n)$ is written as $M = YX^+$ where X^+ denotes the generalized inverse of X. The Hopfield Model is guaranteed convergence in the case of symmetric interconnections. For an autoassociative mapping, where X maps to X, the matrix M is always symmetric. In the heteroassociative case, M is not necessarily symmetric.

The HM can be used with outputs of ± 1 or 0,1. For the first case, the autoassociative OAM is appropriate. Each vector to be stored is associated with itself. For the second case, the heteroassociative OAM is more appropriate. The reason is that the threshold input value of the transfer function of the neurons is at zero, so for a zero output the input to the unit should be distinctly negative. Thus, a convenient prescription is to associate a binary vector with the bipolar transformed version of itself, i.e. zeros are associated with minus ones. The above two cases have some important differences in performance.

Both of the above systems have been simulated on a computer. Networks with up to 20 neurons were modelled, with 10 being the most common number. In the autoassociative case, only stable

VERAC
Incorporated

17

convergence was seen, as expected. The system did tend to recall intentionally stored states. In fact, for N processing units, all of the stored states appeared to be recallable where up to and beyond .5N states were stored. Unfortunately, the regions of attraction of some of these states became very small, and spurious states were often recalled, as well as the complements of states that were stored, a phenomenon seen in the HM. In other words, although the stored states all seemed to be stable, their recallability degraded sharply as their number increased. The spurious states seem to differ from those in the HM in that they are not stored states that have degraded. They are additional undesired stable states. These spurious states and the stable complements may stem from the property of the OAM to associate linear combinations of the key vectors with linear combinations of their associated vectors. Thus, the negative of a vector made of -1's and +1's, which is a linear combination of that vector and is also the compliment of that vector, might be expected to be recalled by a system that stored that vector. Similarly, the spurious states observed may be related to linear combinations of the stored states. Interestingly, the system with outputs of 0,1 are free of stable complements, perhaps since the complement of a vector made of 0's and 1's is not a linear combination of that vector, as was the case previously. Spurious states still plague this system. The storability and recallability seem to be at least as good as the previous system. An additional interesting feature of the 0,1 system is that despite the non-symmetric interconnection matrix, stable, non-oscillatory behaviour was always observed, leading to a suspicion that some particular properties of the OAM may imply stability.

VERAC
Incorporated

To summarize, the OAM prescription seems to be characterized by superior stability of stored states, but equal or even inferior recallability, depending on the number of states stored, and an absence of stable complements in the 0,1 system. Conceivably, forgetting algorithms such as those used with the HM might be developed to enhance its usefulness.

The Bidirectional Asociative Memory (BAM), developed by Kosko, is a heteroassociative memory. It has similar components and functionality to the HM and suffers from some of the same disadvantages. As above, replacement of a SOP storage scheme with an OAM scheme was examined. In this case, two OAM's were used, each mapping one of the two fields to the other. Thus, the requirement of a single valued set of interconnections between individual pairs of processing units which guarantees convergence is no longer met. There is a different matrix used depending on the direction of data flow.

Simulations of this system were performed. Empirical evidence suggests that this system, as with the HM-OAM system, shows stable convergence. No oscillatory behavior has been seen. Only simulations of a system with outputs of ± 1 were attempted. The results indicate that complements of stored states were stable. Also, the stored states themselves were indeed stable, but with regions of attraction which varied greatly in size. Spurious states were present also. Qualitatively, the behavior was quite similar to that of the HM-OAM system. As an example, a BAM with fields of 8 and 10 units was modelled, using an OAM in which up to 5 pairs of vectors were stored. Initial states for the BAM were set at each of the 2^8 possible states of the 8 unit field, and the system was iterated until convergence. Roughly 1/2 of the time the system

converged to complements or other spurious states. Otherwise it converged to an intentionally stored state. These numbers appear to get worse with increasing numbers of stored vectors.

The results indicate that complements of stored states were stable. Also, the stored states themselves were indeed stable, but with regions of attraction which varied greatly in size. Spurious states were present also. Qualitatively, the behavior was quite similar to that of the HM-OAM system. As an example, a BAM with fields of 8 and 10 units was modelled, using an OAM in which up to 5 pairs of vectors were stored. Initial states for the BAM were set at each of the 2^8 possible states of the 8 unit field, and the system was iterated until convergence. Roughly 1/2 of the time the system converged to complements or other spurious states. Otherwise it converged to an intentionally stored state. These numbers appear to get worse with increasing numbers of stored vectors.

Optical Cohen-Grossberg System

Optical systems were studied for the implementation of the Cohen-Grossberg (CG) system. This a continuous version of the HM that obviates the need for asynchronous updating and threshold operations, and is more conducive to an optical implementation. Previous optical implementations have suffered from the need for electronic datapaths in the feedback loop in order to perform thresholding, and accomodate inhibitory (i.e. negative) signals. It is desirable to take full advantage of the benefits of optics. An architecture based on an electrooptic modulator that enables all-optical feedback has been designed and simulated, with good results.

The initial design, shown in fig. 1, is based around a Hughes Liquid Crystal Light Valve (LCLV). The processes to be carried out are multiplication of a matrix and a vector, performing a nonlinear sigmoid type operation on the result, and feeding the result back to the input. In this system, an optically addressable electrooptic spatial light modulator such as the LCLV is used to perform part of the matrix vector multiplication as well as provide the nonlinear transfer function. Electrooptic modulators work by changing the indices of refraction along the different principal axes of propagation of the material in response to an input signal, typically optical or electrical. The effect is to change the polarization state of light reflecting from or passing through the modulator. When the light then passes through a linear polarizer, the polarization change is transferred to an intensity variation. Typically, the

intensity as a function of input signal, for an output polarizer that is perpendicular to the input polarization, is roughly a $\sin^2$ dependence. Some modulators, such as the LCLV, tend to saturate at the upper end of the response curve, yielding a sigmoid function. In the perpendicular mode LCLV (in this case, perpendicular mode refers to the liquid crystal orientation), this function is monotonically increasing. It can easily be shown that if the input light is polarized parallel to the output polarizer, the response is precisely complementary to the case of perpendicular polarization, that is, the transfer function is a constant minus the perpendicular case value at that input level. Two orthogonally polarized beams with equal amplitudes, passing through an output polarizer that is parallel to one of them, will yield a constant intensity at the output as a function of the input signal to the modulator. This fact is the basis of an encoding scheme that will enable all-optical feedback.

If light whose intensity represents a vector component is presented to the input side of the LCLV in the form of a vertical strip of uniform intensity, then a matrix that is projected on the front of the LCLV will have a column multiplied by some function of the input light. As discussed earlier, this function will be an increasing or decreasing sigmoid, depending on the relative polarization of the read or output light and the output polarizer. Thus, it is possible to let the input light represent the signal entering a processing element and the LCLV act as the nonlinear element. The LCLV also then performs the requisite matrix vector multiply (with the exception of the final summation) between the interconnect matrix, which is encoded in the read beam, and the outputs of the processing units. The final summation of the

VERAC
Incorporated

22

products to form the new vector can be carried out using a cylindrical lens and a diffuser. The diffuser captures the intensity at the focus of the cylindrical lens, which corresponds to the sum, enabling it to be spread out and imaged onto the back of the LCLV. If an electrically addressable modulator were used, a set of photodetectors could be placed at the focus.

There remains the question of negative matrix element encoding. As alluded to earlier, the solution to this lies in the complementary response of the modulator system to orthogonally polarized beams. For an encoding scheme to work, we require that the matrix elements corresponding to zero will yield a product with an output vector component that is independent of the component value i.e. anything multiplying zero gives zero. This condition is fulfilled by 2 beams of equal amplitude but orthogonal polarizations, as described earlier. This means a zero matrix element will always contribute a fixed amount of light to the feedback signal. This is fortuitous, in fact, since the transfer function of modulators such as the LCLV has an inflection point at a greater than zero input signal. Since the inflection point in the continuous HM lies at zero input signal, one need only adjust the amount of light in the matrix beam until a matrix consisting of zeros yields enough light in the feedback to operate the modulator at the inflection point. The requirements of the network and the characteristics of the modulator are thus mutually satisfactory. A negative matrix element would then correspond to having more light of polarization parallel to the output polarizer than perpendicular, so that an increase in the output vector which multiplies that element will actually yield a decrease in the total amount of light being fed back, as desired of a negative contribution. It turns out that the

VERAC
Incorporated

best scheme is to have the same amount of one polarization compone: for all elements, while letting the other component vary in amplitude. Note that the system provides outputs between 0 and 1.

That the above system can perform the desired operations can be shown rigorously. Imagine a modulator with a transfer function $o(i)$ as shown in fig. 2. This is the transfer function seen for light perpendicularly polarized to the output polarizer. Light orthogonal to that would see a function $1 - o(i)$. Let I_0 be the intensity of light striking the optical interconnect matrix, $T = |M| + .5$, whose elements are the elements of the numerical interconnect matrix, M, linearly mapped to lie within the interval $[-.5,.5]$, and offset by .5, such that zero elements are half transmittance, negative values are less, and positive are more, with the maximum possible value of 1 and minimum of 0. (To maximize the dynamic range of the device displaying T, one would like the maximum element of the matrix as close to 1 and the minimum as close to zero, while keeping zero values at 1/2 transmittance, as such a linear representation allows.) Let A be the effective attenuation seen by a beam of unit intensity with area sufficient to cover the matrix, by the time it is presented to the writing side of the LCLV. Let .5 be the constant amplitude reference beam matrix, whose elements are equal to the transmittance of the zero elements in T, i.e. it is a uniform beam with intensity $I_0/2$. Then the vector on the input side of the LCLV is given by

$$\underline{i} = I_0 A T \underline{o}(i) + I_0 A(.5)(\underline{1} - \underline{o}(i))$$

$$= I_0 A[(|M| + .5)\underline{o}(i) + .5(\underline{1} - \underline{o}(i))]$$

$$= I_0 A[|M|\underline{o}(i) + .5(\underline{1})]$$

$$= I_0 A[\textbf{M}\underline{\varrho}(\underline{i}) + .5\underline{N}]$$

where $.5\underline{N}$ is a constant vector, each of whose components is $.5N$, where N is the number of components. Rewriting, we have

$$\textbf{M}\underline{\varrho}(\underline{i}) = \underline{i}/I_0 A - \underline{N}$$

Letting $\underline{i}/I_0 A - \underline{N} = \underline{i}'$ we have

$$\textbf{M}\underline{\varrho}((\underline{i}' + \underline{N})I_0 A) = \underline{i}' \qquad (3)$$

where $\underline{i}'$ is a unitless vector. Thus, the system executes the desired multiplication of a bipolar matrix $\textbf{M}$ and a shifted and scaled version of the transfer function $\underline{\varrho}(\underline{i})$.

Equation (3) can be used to numerically analyze the behavior of a given modulator in such a system using only its response curve. It might seem at first glance that the factor $I_0 A$ in eqn.(3) is system dependent. However, recall from earlier discussions that the input light must be adjusted until the intensity finally reaching the input side on the modulator will, for a matrix of zeros, bias the modulator at its inflection point. This point, denoted by h in Fig. 2, is obtained directly from the modulator response curve, and replaces the factor of $I_0 A$ in Eqn.3. Thus, only the response curve of the modulator is required to evaluate the performance of the system. Computer simulations of the above system support the validity of these arguments.

Whether or not a given system will execute a useful processing task is not guaranteed by the above arguments. What is guaranteed is that the system is equivalent to a Cohen-Grossberg system, and will thus correctly carry out the correct matrix-vector multiply, and, for a symmetric matrix, converge to a stable state. However, as discussed by Hopfield, when a sigmoid response curve is made less and less steep, the stable points move in from the corners of the unit hypercube and become closer and

VERAC
Incorporated

closer to the center. The point is that the response curve of a particular modulator may effectively be so flat that the 0's and 1's are so close to each other as to be indistinguishable. The only way to be certain is to simulate eqn. 3 on a computer.

Experimental Investigations

Experimental work centered on implementing the above system optically. Initial attempts involved a perpendicular mode LCLV provided by Hughes. This device has a monotonically increasing, roughly sigmoidal response curve. Using the set-up shown in Fig. 1, though, no distinguishable states were recalled for the matrix used. Subsequent computer simulations, based on the above arguments, showed that the response curve of the LCLV was not sufficient to yield distinguishable 1's and 0's.

Initial experiments have been performed using a special CRT driven LCLV. A small CRT provides the input to the writing side of the LCLV, and, instead of feeding the optical signal directly back, a CCD camera images the input vector and sends a video signal to the CRT (see Fig. 3). The system still demonstrates the coding scheme discussed above. The video conversion simply allows greater control over the response characteristics of the LCLV. Data on the response of this system indicate, via computer simulations, that the response of the combined CRT-LCLV system is sufficient to observe distinguishable 1's and 0's. The next step will be assembly of the complete system.

Initial work with the LCLV demonstrated some disadvantages of using a reflective mode device. If the read-matrix beam is incident normally on the surface of the LCLV, the requisite beamsplitters cause a significant light loss. The solution is to bring the beam in

VERAC
Incorporated

at a slight angle to the LCLV, as in Fig. 1, but this requires setting the optical components farther away from the modulator so as not to intercept the reflected beam. This necessitates imaging devices due to unavoidable diffraction. Also, operating the device with the incident light at an angle may yield suboptimal performance. Most of the problems can be avoided by the use of a transmission mode device. Such a system might take the form of Fig. 4. The modulators required seem quite feasible, given the work going on in this department and elsewhere. We hope to explore the possibility of fabricating such devices in the coming year.
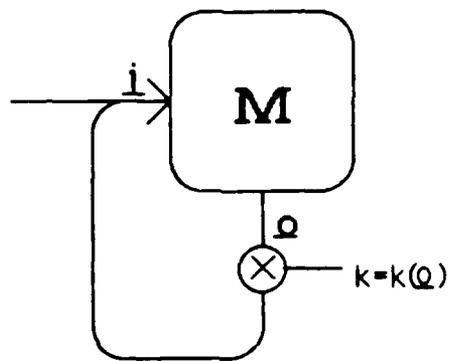
fig. 1
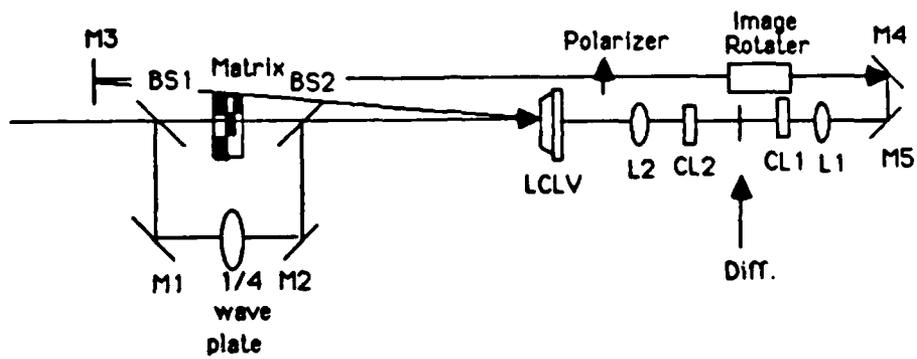
fig. 2

fig. 3a

fig. 3b

fig. 4

VERAC
Incorporated
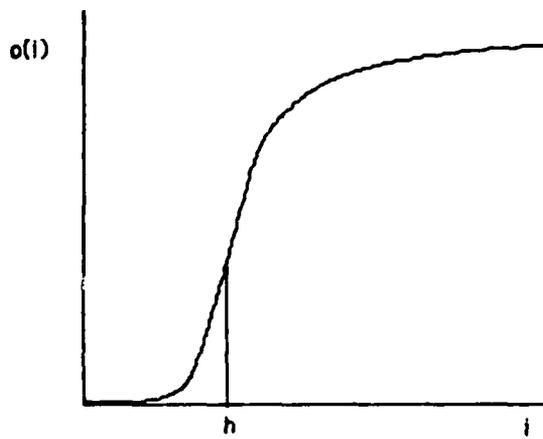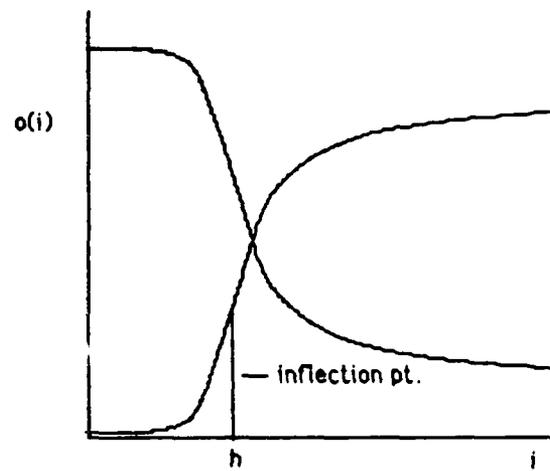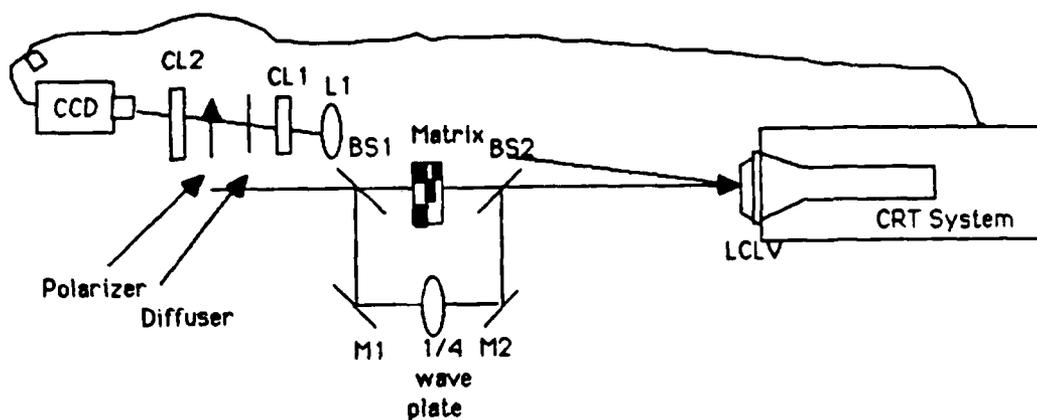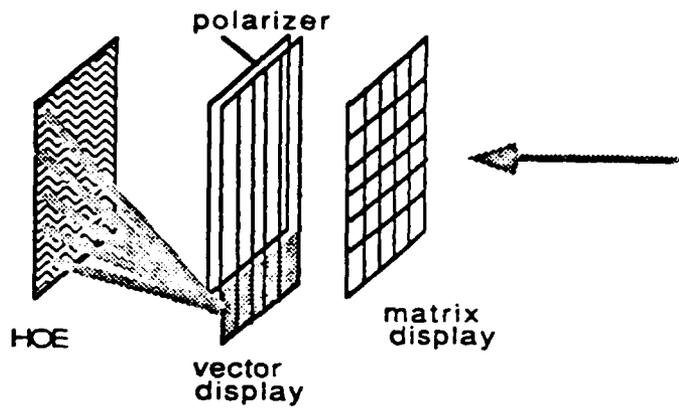
fig. 5

END

DATE

FILMED

APRIL

1988

DTIC