

Optical information processing based on an associative-memory model of neural nets with thresholding and feedback

Demetri Psaltis and Nabil Farhat*

Department of Electrical Engineering, California Institute of Technology, Pasadena, California 91125

Received July 9, 1984; accepted November 15, 1984

The remarkable collective computational properties of the Hopfield model for neural networks [Proc. Nat. Acad. Sci. USA 79, 2554 (1982)] are reviewed. These include recognition from partial input, robustness, and error-correction capability. Features of the model that make its optical implementation attractive are discussed, and specific optical implementation schemes are given.

Optical information-processing systems can have high processing power because of the large degree of parallelism as well as the interconnection capability that is achievable. Typically, more than 10^6 parallel processing channels are available in the optical system, and furthermore each of these channels can be optically interconnected (broadcasted) to 10^6 other channels. The majority of optical processors are analog systems, designed to perform linear operations. The accuracy of an analog processor is limited by the linear dynamic range of the devices used (detectors, light modulators). In principle, the accuracy and the repertoire of achievable operations can be improved with systems that perform nonlinear operations on binary encoded data using bistable optical devices. Optical bistability is a subject that has received considerable attention recently as a means of achieving efficient high-speed logic, and it has been demonstrated with several nonlinear optical materials and devices. If we are to use such bistable devices to realize powerful, nonlinear optical computers, it is important to find algorithms that are well matched to the characteristics of the optical processor and utilize effectively its parallelism and interconnection capability. In this Letter we examine a method for synthesizing optical processing systems, based on optical associative memory and threshold logic, that appears to meet these requirements well.

Associative (or content-addressable) memories are of interest in computer science, and it is theorized that information is stored in the human brain in this manner. Holographic associative memories have been described by Gabor,¹ who also commented on the similarity of the holographic memory to the way information may be stored in the human brain. More recently, Hopfield² introduced an associative-memory model to describe the collective behavior of neural networks. Hopfield's model consists basically of an associative memory similar to the holographic, with the addition of threshold and feedback. The incorporation of nonlinear feedback enhances dramatically the error-correcting capability of the holographic memory.

Let $v_i^{(m)}$ be a binary word that is N bits long. M such words are stored in a matrix T_{ij} according to

$$T_{ij} = \begin{cases} \sum_m^M [2v_i^{(m)} - 1][2v_j^{(m)} - 1] & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad (1)$$

If T_{ij} is multiplied by one of the stored binary vectors $v_i^{(m_0)}$, the product $\hat{v}_i^{(m_0)}$ is an estimate of the stored vector $[2v_i^{(m_0)} - 1]$:

$$\hat{v}_i^{(m_0)} = \sum_j^N T_{ij} v_j^{(m_0)} = N_0 [2v_i^{(m_0)} - 1] + \sum_{m \neq m_0}^M \left\{ \sum_j^N [2v_j^{(m)} - 1] v_j^{(m_0)} \right\} \times [2v_i^{(m)} - 1] - M v_i^{(m_0)}, \quad (2)$$

where the last term accounts for $T_{ij} = 0$ and N_0 is the number of 1's in $v_i^{(m_0)}$. We assume that for $m \neq m_0$ the binary words $v_i^{(m)}$ are statistically described in the following simple manner:

$$P[v_i^{(m)} = 1] = 1/2, \quad P[v_i^{(m)} = 0] = 1/2, \quad (3)$$

where $v_i^{(m)}$ are independent for all i and m . Then $E[\hat{v}_i^{(m_0)}] = (N/2)[2v_i^{(m_0)} - 1]$ and $\text{var}[\hat{v}_i^{(m_0)}] = N(M - 1)/2$. We define the signal-to-noise ratio (SNR) of the estimate $\hat{v}_i^{(m_0)}$ as the ratio of the magnitude of the expected value of $\hat{v}_i^{(m_0)}$ to the standard deviation of the estimate:

$$\text{SNR} = \frac{\|E[\hat{v}_i^{(m_0)}]\|}{\{\text{var}[\hat{v}_i^{(m_0)}]\}^{1/2}} = [N/2(M - 1)]^{1/2}. \quad (4)$$

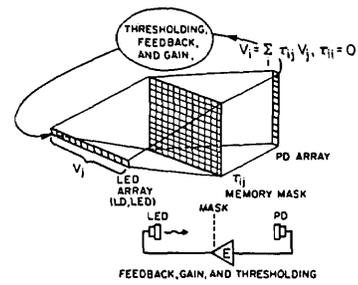
If N is sufficiently larger than M , then with high probability $TH[\hat{v}_i^{(m_0)}] = v_i^{(m_0)}$, where $TH[\hat{v}_i^{(m_0)}] = 1$ if $\hat{v}_i^{(m_0)} > 0$ and zero otherwise. Thus the vector-matrix product in Eq. (2) combined with the thresholding operation results in a pseudoeigensystem in that the output vector equals the input. Now suppose that the full vector $v_i^{(m_0)}$ is in fact such a pseudoeigenvector of the system but that only N_1 of N bits ($N_1 \leq N$) of $v_i^{(m_0)}$ are known. In this case we define an input vector consisting of the N_1 known bits, and the rest are set equal to zero. When this vector is multiplied by the matrix T_{ij} , an estimate of the complete vector $[2v_i^{(m_0)} - 1]$ is obtained. The SNR of the estimate is now $\text{SNR} = [N_1/2(M - 1)]^{1/2}$. If N_1 becomes sufficiently small, then, with high probability, $TH[\hat{v}_i^{(m_0)}] \neq v_i^{(m_0)}$ for some of the values of i . Let N_2 be the number of correct bits in $TH[\hat{v}_i^{(m_0)}]$. If $N_2 > N_1$, we can multiply this thresholded estimate by T_{ij} and obtain a new estimate with a higher SNR. This procedure can be continued until the number of correct bits in the thresholded vector is equal to N . The crucial issue is under what conditions N_2 will be higher than N_1 . If the SNR of the initial

estimate $\{[N_1/(2M - 1)]^{1/2}\}$ is sufficiently large, then the probability of N_2 being bigger than N_1 will be high; in this case the nonlinear, iterative procedure described above will be likely to converge to the correct vector $v_i^{(m)}$. Ideally, each of the M stored binary words is a pseudoeigenvector of the nonlinear system. Notice that each pseudoeigenstate is a stable state of the system, whereas any other input vector (state) will cause a change to occur in the next cycle. In general, the system converges to the stable state that is at the shortest Hamming distance away from the initial state.

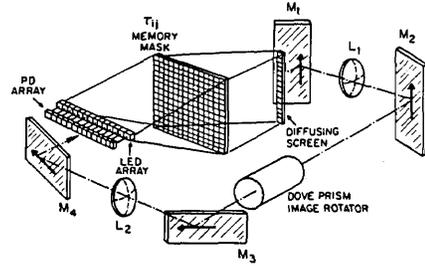
This model has been studied computationally by Hopfield.² In simulations, correct convergence was obtained reliably for $M \leq 0.15N$ and $N_1 \approx 0.75N$, taking $N = 30$. At present there is no (adequate) theoretical prediction of the maximum number of words that can be stored or the maximum Hamming distance between the input vector and one of the stored words that is required for convergence. Several interesting properties were observed. The model does not require synchronism. Convergence can be obtained if the output vector is fed back to the input as a whole or, randomly, one element at a time. There is some evidence that asynchronous operation is actually preferable. The system is quite insensitive to imperfections such as nonuniformities, the exact form of the threshold operation, and errors in the T_{ij} matrix. Convergence to the correct vector was obtained even when the T_{ij} matrix was thresholded. Such properties are most desirable when an optical implementation is considered.

One possible optical implementation of the Hopfield model is through the arrangements shown in Fig. 1, in which the array of light-emitting diodes (LED's) represents N logic elements with binary states $v_j = 0, 1, j = 0, 1, \dots, N$ (LED on or off), which are to be interconnected in accordance with the model. This is achieved by the addition of nonlinear feedback (feedback, thresholding, and gain) to the well-known optical vector-matrix multiplier.³ Gain is included in the feedback loop to compensate for losses. Two possible feedback schemes are shown. One uses electronic wiring and the other is optical, with the thresholding (point nonlinearity) and the gain concentrated between the photodiode (PD) array and the LED array, which can be fabricated monolithically on GaAs. Furthermore, with the accelerating pace of research in thin-film nonlinear light amplifiers⁴ and optical bistable devices,⁵ it can be possible to substitute a single distributed bistable light-amplifier device for the PD/LED arrays and the intervening thresholding and amplifying electronics.

Multiplication of the vector v_j by the T_{ij} matrix in these schemes is accomplished by horizontal imaging and vertical smearing of v_j using anamorphic optics (omitted from Fig. 1 for simplicity). A bipolar T_{ij} can be realized optoelectronically with incoherent light by assigning its negative and positive values to adjacent rows. Light passing through each row is focused onto adjacent pairs of photodiodes of the PD array that are electronically connected in opposition, as shown in Fig. 2. Here the positive and negative elements of each row of the T_{ij} matrix are separated into two subrows, one for positive values and one for negative. The light transmitted through the two subrows is integrated horizon-



(a) ELECTRONIC FEEDBACK SCHEME



(b) OPTICAL FEEDBACK SCHEME

Fig. 1. Two schemes for adding nonlinear feedback to an optical vector-matrix multiplier utilizing (a) electronic feedback and (b) optical feedback.

tally with the aid of another set of anamorphic lenses (omitted from Figs. 1 and 2) and brought to focus on two adjacent photodiodes of the PD array connected in opposition. The output of the first diode-pair circuit will be proportional, $v_1 = \sum_j T_{1j} v_j$. This output is applied through an electronic thresholding circuit to the first element of the LED array, as shown in Fig. 1. Similar connections are made between other detector pairs of the photodiode array and corresponding elements in the LED array. Thus each LED assesses the state of its input $v_i = \sum_j T_{ij} v_j$ and fires according to whether v_i exceeds the threshold or not.

We now consider the possibility of optically storing two-dimensional (2-D) functions (images). Let $v^{(m)}(i, i')$ be the bipolar binary (1, -1) images to be stored. If we directly extend the Hopfield model to two dimensions, then these images must be stored in a four-dimensional function in the following general form:

$$T(i, i', j, j') = \sum_m^M v^{(m)}(i, i') v^{(m)}(j, j'). \quad (5)$$

In order to implement a 2-D Hopfield memory optically, we need to realize a 2-D, linear optical system whose spatial impulse response is the four-dimensional function defined in Eq. (5). Since we have only two spatial coordinates to work with in an optical system, it is difficult to implement such a system directly for the nonseparable, shift-variant kernel defined in Eq. (5). One possible solution is the use of wavelength multiplexing and/or time-domain processing to obtain additional independent variables. Another solution is based on holographic associative memories, as we have discussed earlier.⁶ Here we present an implementation based on spatial-frequency multiplexing.

The entire optical system, including nonlinear feedback, is shown in Fig. 3. The system accepts a 2-D

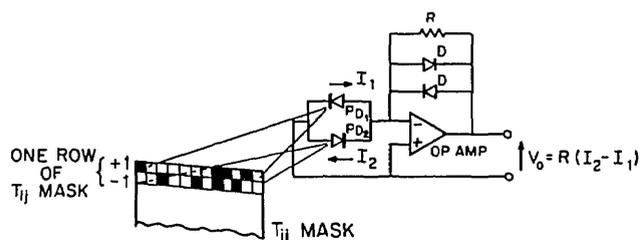


Fig. 2. Scheme for realizing bipolar mask transmittance with incoherent light.

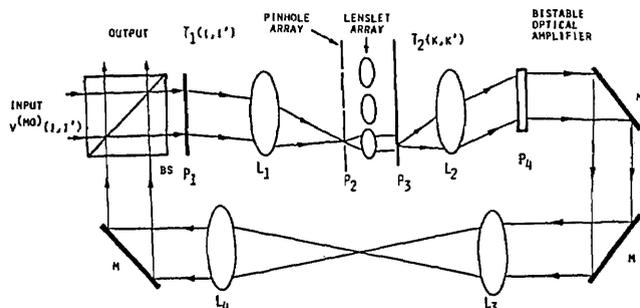


Fig. 3. Coherent optical implementation for 2-D inputs.

input $v^{(m0)}(i, i')$, which illuminates the system from the left in Fig. 3. The 2-D interconnection pattern between the planes P_1 and P_4 that is prescribed by Eq. (5) is stored on two separate optical transparencies, denoted T_1 and T_2 in Fig. 3. Each image $v^{(m)}(i, i')$ is placed on a separate spatial-frequency carrier, and the images are added to form the first transparency T_1 . When T_1 is illuminated with an input image, the products between the input and all the stored images are formed. The lens L_1 produces the Fourier transforms of all the products and displays them spatially separated at the back focal plane of L_1 (plane P_2). An array of pinholes is placed at P_2 , the position of each pinhole being at the spatial frequency of each carrier used in the recording of transparency T_1 . Therefore the amplitude of the light transmitted through each individual pinhole is proportional to the integral of the product of the input image and the corresponding image stored in T_1 . In the discrete notation used in this Letter, the amplitude of the light transmitted through the m th pinhole is proportional to $\sum \sum v^{(m0)}(i, j)v^{(m)}(i, j)$. The second transparency, $T_2(k, k')$, consists of a 2-D array of Fourier-transform holograms of all the stored images. Each hologram is formed at a separate position on the hologram, the transform of the m th image being centered at the location of the corresponding m th pinhole. Light emerging from each pinhole illuminates only the corresponding hologram. The lens L_2 takes the Fourier transform between planes P_3 and P_4 , thereby reconstructing all the images stored in T_2 . The light amplitude at P_4 is a weighted sum of all the stored images, the weights being proportional to the inner product between the input and stored images. This is precisely the desired output that is produced by the interconnection prescription given in Eq. (5). The modulation of the light at plane P_4 will in general be bipolar, and it is interferometrically detected by the nonlinear optical amplifier at P_4 , which performs the thresholding op-

eration. If interferometric detection proves to be too cumbersome, it is possible to modify the interconnection pattern such that the output is always positive. This results in a loss of storage capacity by a factor of 2, but it may be a welcome trade-off. The thresholded image is fed back to the input through mirrors and imaging optics (lenses L_3 and L_4). The output of the system is taken at the beam splitter. Optical gain must be included in the cavity (preferably through the bistable optical element) to compensate for losses that are due to the passive components. The requirement that $T(i, i', j, j') = 0$ for $i = i', j = j'$ must also be satisfied because otherwise the diagonal elements always become equal to M , whereas the off-diagonal elements have an average value of \sqrt{M} . The result is that for large M the system becomes an imaging system; any input is replicated at the output. This is avoided by forming each of the Fourier-transform holograms in P_3 with a randomly chosen, uniform phase.

We have described several specific optical implementations of the Hopfield model; undoubtedly others are also possible. The most important feature of all such implementations is the robustness of a system that utilizes nonlinear feedback. The systems that we have described behave basically as associative memories (the whole is retrieved from a partial input), even with open-loop operation. However, the nonlinear feedback can correct errors of the open-loop system since it forces the state of the system to change continuously until a stable condition is reached. The nonlinearity plays a crucial role; if linear feedback were used, the system would either be unstable or converge to the eigenstate of the open-loop system with the highest eigenvalue, independently of the initial condition.

This error-correcting capability can provide the accuracy that is lacking from analog optical processors without, however, sacrificing the processing power that can be derived from the global processing capability of optics; the class of processors that we described are fully interconnected optical systems and hence utilize fully the parallelism and the interconnectivity capability of optics. In general, there is an excellent match between the global, linear operations and local, point nonlinearities that are required for the implementation of the Hopfield model, and the capabilities and limitations of optical techniques.

The authors thank John Hong and Yaser Abu-Mostafa for many helpful discussions on this subject.

* On scholarly leave from the University of Pennsylvania, Philadelphia, Pennsylvania 19104.

References

1. D. Gabor, IBM J. Res. Dev. **13**, 156 (1969).
2. J. J. Hopfield, Proc. Nat. Acad. Sci. USA **79**, 2554 (1982).
3. J. Goodman, A. R. Dias, and I. M. Woody, Opt. Lett. **2**, 1 (1978).
4. Z. Porada, Thin Solid Films **109**, 213 (1983).
5. H. M. Gibbs, S. L. McCall, and T. N. C. Venkatesan, Opt. News **5**(3), 6 (1979).
6. D. Psaltis and N. Farhat, presented at the Meeting of the International Commission for Optics, ICO-13, Sapporo, Japan, August 1984.