Higher Order Associative Memories and their Optical Implementations

Demetri Psaltis   Cheol Hoon Park   John Hong

Presented By

Sarthak Bagaria

IIT Bombay

Introduction

- An associative memory is a system which stores a mapping of N-dimensional input vectors to N_0-dimensional output vectors
- The system should be capable of implementing all possible mappings.It's effectiveness depends on
  - Capacity
  - Learning
  - Generalization
- For binary outputs, capacity is given by $C \leq \frac{D \log_2 K}{N_0}$,where D is the number of independent variables and K is the separate values each can assume.

Linear discriminant functions

- It maps an input vector to +1 or -1 by

$$y = sgn\{w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_N x_N\}$$

- This function dichotomizes the set of input vectors (partitions into two space, separated by a hyperplane).
- All the dichotomies are possible only when the input vectors are less than N+1, so we get the capacity as : $C = N + 1$
- Since this capacity is low, we expand the vectors to their r^{th} order to get $L = \binom{N+r}{r}$ number of independent terms.

$$z_j(x) = x_{p_1(j)}^{n_1} x_{p_2(j)}^{n_2} \cdots x_{p_r(j)}^{n_r}$$

$$y = sgn\{w_0 + w_1 z_1(x) + w_2 z_2(x) + \cdots + w_L z_L(x)\}$$

Here the number of weights used to describe the mapping is L+1, and so is the capacity of the memory.

Binary Vectors

- There are 2^N non redundant terms in a complete polynomial expansion of a binary vector. Which is equal to the total number of possible input vectors. hence, we can say that this memory is capable of implementing all the possible mappings.

x_1	x_2	x_3	1	x_1	x_2	x_3	x_1x_2	x_2x_3	x_3x_1	$x_1x_2x_3$
1	1	1	1	1	1	1	1	1	1	1
1	1	−1	1	1	1	−1	1	−1	−1	−1
1	−1	1	1	1	−1	1	−1	−1	1	−1
1	−1	−1	1	1	−1	−1	−1	1	−1	1
−1	1	1	1	−1	1	1	−1	1	−1	−1
−1	1	−1	1	−1	1	−1	−1	−1	1	1
−1	−1	1	1	−1	−1	1	1	−1	−1	1
−1	−1	−1	1	−1	−1	−1	1	1	1	−1

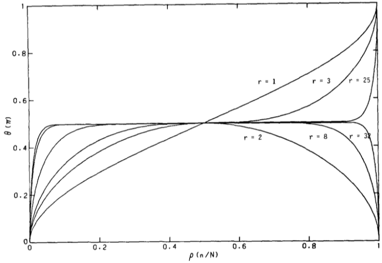
- The orthogonalization property of the full expansion is interesting because it shows that higher order memories provide a complete framework that takes us from the simplest "neuron", the linear discriminant function to the full capacity of a Boolean truth table.

R^{th} order expansions

- For a large N, the full expansion is too long. Hence, we only take r^{th} order expansions which gives us large enough capacity to learn.
- Angle between two expanded vectors is given by

$$\cos\theta_r \approx 2\sqrt{r\rho}$$

where ρ is n/N and n is the Hamming distance between the original input vector.



Training

$$W_{lj_1j_2\cdots jr} = \sum_{m=1}^{M} y_l^m x_{j1}^m x_{j2}^m \cdots x_{jr}^m$$

$$\begin{aligned} y_l &= \text{sgn}\{N^r Y_l^n + \sum_{m \neq n} y_l^m (\sum_{j=1}^{N} x_j^m x_j^n)^r + w_l^0\} \\ &= \text{sgn}\{N^r y_l^n + n_l(x^n)\} \end{aligned}$$

- The first term is identified with the desired signal and the second term with the noise. Thus we can calculate the signal to noise ratio.

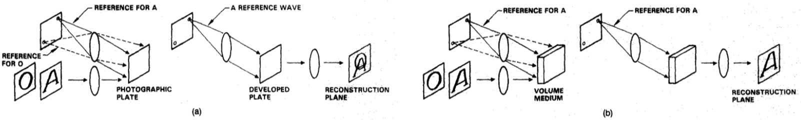
$$SNR \approx \left(\frac{N^r 2^r r!}{M(2r)!}\right)^{\frac{1}{2}}$$

- Equating the signal to noise ratio of linear and r^{th} order we get the r^{th} order capacity as

$$\frac{M_r}{M_1} = N^{r-1} \frac{2^r r!}{(2r)!}$$

Optical Implementations of Quadratic associative Memories

- The holographic process consists of recording and reconstruction. In the recording step, the interference between the reference plane wave and the wave originating from object "A" is recorded on a photographic plate. This plate illuminated with thereference fied, to give a virtual projection of "A".
- The weight of each interconnection is given by the interference pattern
- The volume hologram is similar, but it records the pattern on a three dimensional medium.

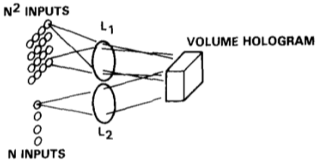


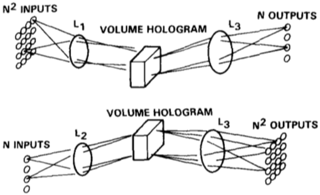
Volume Hologram Systems

- To implement the quadratic memory, volume holograms are used to connect the input and output patterns. They are implementations of the weight tensor.
- $N^2 \rightarrow N$ schemes use the weight tensor

$$W_{ijk} = \sum_{m=1}^{M} y_i^m x_j^m x_k^m N^2$$

- This allows for error driven learning where interconnections are developed by an iterative training process.





Volume Hologram Systems

- $N \to N^2$ is the inverse of the one described previously.
- Input dependent weights enable quadratic memories in an $M \to N$ scheme

$$w_{ij} = \sum_{k=1}^{N} w_{ijk} x_k$$

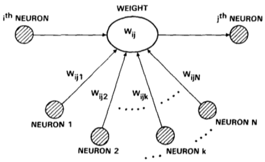


FIGURE 10. Quadratic mappings implemented as nonlinear interconnections.

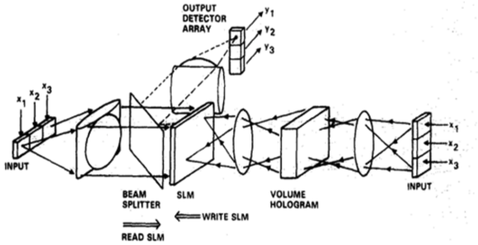
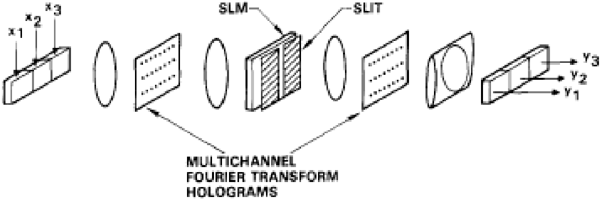


FIGURE 11. Optical architecture for the implementation of the nonlinear interconnections of Figure 10.

Planar Hologram Systems

- They do not have the extra dimension to directly implement the weight tensor, nevertheless they can be used in a similar way.
- The first part of the system is a multichannel correlator, which correlates input vectors to the stored vectors. The correlation functions are then sampled at the slit and squared by the SLM. Second hologram produces the weighted of vectors, which are Fourier transformed to obtain output vectors.
- One can incorporate shift invariance by lengthening the input SLM and removing slits.



References

- Psaltis D., Park C.H., Hong J. (1988). *Higher order associative memories and their optical implementations*. Neural Networks, Volume 1, Issue 2, 149-163.
- Cover T. M. (1965). *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*. IEEE Transactions on Electronic Computers, EC-14, 326.
- Psaltis D., Hong J. (1987). *Shift-invariant optical associative memories*. Optical Engineering. 26(1), 10.
- Paek E. G., Psaltis D. (1987). *Optical associative memory using Fourier transform holograms*. Optical Engineering, 26(5), 428.