HOLOGRAPHIC HIGH ORDER ASSOCIATIVE MEMORY SYSTEM

Z.Q.Mao, D.R.Selviah, S.Tao and J.E.Midwinter

University College London, United Kingdom

INTRODUCTION

Associative memory systems allow vast databases to be searched rapidly to recover all of a dataset when only part of that dataset is available initially. It is also often important to be able to modify the recalled dataset and then to restore it as an updated memory. Opto-electronic technology offers the potential for high memory capacity and fast recovery time through the use of holograms for dense data storage and neural network architectures for associative recall. The neural network architectures allow all of the memories to be searched in parallel taking advantage of the high parallelism possible in optics. Electronics offers only a limited fanout and connectivity and although useful for small number of neurons, it does not scale as the number of neurons increases. This is where holographic optics has a role to play since high fanout and large connectivity are possible holographically, provided a suitable system architecture can be found which can operate within the limited dynamic range available in optical systems (typically 30dB).

For the memory element itself, Fourier transform holograms are especially useful due to their spatial invariance and information compression. There are two limits to the memory capacity of such a system:

(i) First, a practical limit imposed by the number of holograms that can be stored in the holographic medium. This may be due to coupling between the holograms when they are angularly multiplexed and superimposed in the same photorefractive crystal, for example. Or it may be due to the field of view of a lens used to read a two dimensional array of holograms in a spatially multiplexed scheme.

(ii) Secondly, a theoretical limit due to the system design. As the number of memories is increased beyond a certain limit the probability of recalling the correct dataset becomes unacceptably low.Linear associative memory systems have a low memory capacity. For example, the Hopfield model can store only about 0.15N patterns (for low values of N) where N is the number of pixels in each pattern [1]. For increasingly large values of N the memory capacity tends to saturate. In order to overcome this theoretical limit it is necessary to use higher order nets. Several optical implementations have already been demonstrated for second order, quadratic, nets [2-8]. Owechko et al. [7] made clear that the memory capacity could be increased still further by raising the order of the non-linearity.

In this paper, we describe an associative memory system based on a High Order Feedback neural NETwork (HOFNET) [9] which has a large memory capacity since it uses a high order non-linearity and is inherently tolerant to noise. The system calculates the similarity between the input and each of the stored patterns by finding the correlation between their spectra using the convolution theorem. In the next section, we outline the system for recording the holographic memory. Following this we describe the optical associative memory system design and prove that the zero order of the correlation of the spectra of two patterns is a measure of their similarity. Finally, our initial experimental results are presented demonstrating optical associative recall.

HOLOGRAPHIC MEMORY RECORDING SYSTEM

There are three possible kinds of holographic recording scheme: angularly multiplexed, spatially multiplexed and hybrid. In the angularly multiplexed recording, the holograms are superimposed on the same area of the recording medium and, to enable them to be distinguished, each hologram has a different modulation frequency so the angles between the reference beam and patterns to be stored must be set differently for each stored pattern. In the spatially multiplexed recording system, however, the modulation frequencies for all the patterns are the same, while the positions on the recording plate are different. The hybrid method involves both frequency modulation and position change. In our system, the spatial multiplexing scheme is chosen to store patterns during the training phase (Fig. 1). This scheme has the advantage that there is no cross coupling between the holograms and, if recorded in a photorefractive material, each may selectively be erased and rerecorded so updating can take place without interfering with other memories. The training exemplar pattern is illuminated by a 5mW He-Ne laser via a ground glass screen to randomise the phase. The ground glass screen causes the Fourier Transform to have an almost uniform intensity distribution instead of having an intense zero order component. This has the advantage that the beam ratio can be kept constant for all frequency components of the original data giving uniform diffraction efficiency across one hologram. Also in this way the information can be stored in a small area, from any part of which the stored information can be recovered. This reduces the effect of dust and damage on the recording plate. The recording plate is moved laterally and further training patterns are recorded side by side. In our initial experiments the recording was made in silver halide and bleached to give a phase hologram in order to increase the diffraction efficiency (but lower signal to noise ratio resulted from due to the increased scattering from the emulsion grains).

All of the holograms were recorded under the same conditions, in particular the modulation frequency was the same for all patterns. This eliminates the problem of diffraction efficiency variation due to the modulation frequency variation of the holograms. Practically, it is difficult to realize uniform diffraction efficiency. Originally, we used ordinary slide films with a plastic substrate as input transparencies to record holograms on the holographic plate. The diffraction efficiencies were found to vary over a large range, from less than 4% to more than 14%. This was much larger than expected and far beyond the fluctuation range allowed in the experiment. This phenomenon was found to be caused by the polarisation rotation effect due to the birefringence of the plastic transparency substrate. When light passed through the slide films, the polarisation rotation angle varied across one pattern and from one pattern to another. When such films were replaced by glass holographic plates, much better results were obtained. As the efficiency also depends on exposure time, emulsion evenness, the form of the pattern itself, processing procedures as well as the modulation frequency, it is very difficult to keep them all exactly the same. In our preliminary experiment, 16 patterns were stored with a variation in diffraction efficiency from 13% to 15%. The holograms were arranged in 4x4 array with a space between two adjacent holograms of 5mm, which matched the subsector spacing of the spatial light modulator (SLM). The diameter of each hologram circular spot was 3mm.

OPTICAL ASSOCIATIVE MEMORY SYSTEM DESIGN

The optical HOFNET associative memory system (Fig. 2) is designed so that the same input pattern appears in each of the channels. This is achieved by allowing the channels to intersect at the input pattern while maintaining the individual identity of each channel by angular multiplexing. Such an arrangement has the advantage that noise in the input pattern is exactly transferred into each channel so no discrepancies between channels due to non-uniform or noisy input devices can occur. After several iterations, the noise will be averaged out. In our system, we use spectrum correlation i.e. spatial multiplication to find the similarity of the input pattern with each of the stored patterns.

The spatial correlation system is shown in Fig. 3(a) [10]. f(x) is one pattern, G(u) is the spectrum of the other pattern g(x). For real patterns, f(x)=f*(x), g(x)=g*(x). The correlation of g(x) and f(x) is calculated at the back focal plane of lens L2:

$$f(x)*g(x) = \int f(y)g(y-x)dy \qquad (1)$$

where * denotes correlation and all of the integrals are from minus infinity to plus infinity. At the centre of the correlation where x=0,

$$(f(x)*g(x))|_{x=0} = \int f(y)g(y)dy \qquad (2)$$

If the patterns are digitised with binary and bipolar values only (+1 and -1), then the above equation becomes

$$(f(i)*g(i))|_{x=0} = \sum_{i}^{N} f(i)g(i) \qquad (3)$$

If two corresponding elements are the same (both +1 or -1), their product is 1, or else the product is -1. So equation (3) reflects the similarity of patterns f(i) and g(i).

Our spectrum correlation system is similar to Fig. 3(a) and is shown in Fig. 3(b). This time F(u) is put in the front focal plane of lens L1 and g(x) is placed in the back focal plane of lens L1 or the front plane of lens L2. g(x) is superimposed on the Fourier Transform of F(u) and the two are multiplied together. At the back focal plane of lens L2, the correlation of the Fourier transforms of the two patterns is found. This is

$$F(u)*G(u) = \int F(v)G(v-u)dv \qquad (4)$$

By the convolution theorem

$$F(u)*G(u) = \int f(x)g(x)\exp(-j2\pi ux)dx \qquad (5)$$

At the centre of the correlation plane, where u=0,

$$(F(u)*G(u))|_{u=0} = \int f(x)g(x)dx \qquad (6)$$

This equation is just Parseval's Theorem for real patterns. If the patterns are digitized with binary and bipolar values only (+1 and -1), equation (6) can be rewritten as

$$(F(u)*G(u))|_{u=0} = \sum_{i}^{N} f(i)g(i) \qquad (7)$$

This is equivalent to equation (3), so we can use the zero order of spectrum correlation to replace that of pattern correlation to reflect the similarity. The advantage of this replacement is that we can record the Fourier transform holograms with the same modulation frequency in the learning procedure. This eliminates the problem of diffraction efficiency variation due to the modulation frequency variation of the holograms.

The hologram array is replayed by a parallel beam which is spatially modulated by a spatial light modulator. The SLM is divided into M "subsectors", each corresponding to a hologram recorded on the plate. According to the translation invariance property of the Fourier transform hologram and the recording condition during the learning procedure, all of the reconstructed patterns will be superimposed at the back focal plane of lens L1. If an input pattern is inserted in this plane and superimposed on the reconstructed patterns, then at the back focal plane of lens L2, the correlation peaks which reflected the similarities between the input pattern and the stored patterns are found. These correlation peaks are fed back to control the transmittance of their corresponding subsectors in the SLM. The SLM used in this system is an electrically addressed ferroelectric liquid crystal SLM and consists of 128x128 pixels in an area of 2cm x 2cm. It is divided into 4x4 subsectors each with size of 5mm x 5mm, matching the size of one hologram. Each subsector, therefore, contains 32x32 binary pixels (black and white). The SLM is used to control the transmitted light passing through it proportional to the correlation outputs. As the correlation outputs are changed continuously, we, ideally, need a grey level SLM. However, since this was not available we simulated grey levels on the binary SLM by further subdividing the subsector into 8x8 units, each containing 4x4 pixels. This resulted in 16 simulated grey level steps. This scheme gives a more uniform beam profile than if the subsector were not subdivided

On the second iteration the pattern which is most similar to the input pattern will be illuminated by a stronger beam (the corresponding subsector of the SLM having higher transmittance), thus the correlation peak will become stronger. The weaker correlations will not grow at such a rate. After several iterations, the weaker correlations fall below the noise floor assuming the maximum correlation is normalised to lie within the dynamic range on each iteration.

Suppose the images we discuss are digitized, so the set of patterns stored in the hologram can be written as

$$S(i) = \{s_1(i), \ldots, s_t(i), \ldots, s_N(i)\} \quad (i=1, \ldots, M) \qquad (8)$$

Suppose the Fourier transform of $S(i)$ is $W(i)$, then

$$W(i) = \{w_1(i), \ldots, w_k(i), \ldots, w_N(i)\} \qquad (9)$$

For simplicity, we let the interval distance between two adjacent sub-holograms be 1 and only consider one dimension. So the hologram array H can be expressed as

$$g(i) = w_k(i) \otimes \delta(x-i) \qquad (10)$$

where delta is the Kronecker delta function and x is lateral position. In the back focal plane of lens L1, we obtain the Fourier transform of $g(i)$, where an input pattern R (r_1, r_2, ..., r_t, ... , r_N) is injected, so at the back focal plane of lens L2, we have the Fourier transform of the multiplication of the input pattern R and stored patterns $S(i)$. If we put a pinhole array at this plane to threshold all the high orders of the correlation, after the pinhole array, we have the amplitude distribution on this plane

$$C(i)^{(1)} = \sum_{t}^{N} s_t(i) r_t \otimes \delta(x-i) \qquad (11)$$

$C(i)^{(1)}$ is fed back to adjust the SLM transmittance so that the transmittance of each subsector is proportional to its corresponding zero order correlation value. Now the light intensity falling on the holograms is modulated. At the end of this iteration and after the pinhole we will get

$$C(i)^{(2)} = \left(\sum_{t}^{N} s_t(i) r_t \right)^2 \otimes \delta(x - i) \tag{12}$$

Similarly, after n-th iteration, we have an output at the back focal plane of lens L2

$$C(i)^{(n)} = \left(\sum_{t}^{N} s_t(i) r_t \right)^n \otimes \delta(x - i) \tag{13}$$

At the input plane we have

$$r_t' = \sum_{i}^{M} C(i)^{(n)} s_t(i) \tag{14}$$

By comparing this equation with the high order inner product model of Owechko et al [7], it is clear that our system is equivalent to the inner product model system with a high order nonlinearity in the correlation plane. The nonlinearity function is

$$f(x) = x^n \tag{15}$$

The signal to noise (actually background of other correlations) ratio of such a system can be written as

$$SNR = \frac{|C(i)^{(n)}|}{\sqrt{\sum_{i \neq io} |C(i)^{(n)}|^2}} \tag{16}$$

where i_0 corresponds to the channel with maximum correlation.

Suppose individual pixels of the patterns have the values +1 and -1 with equal probability, then following the approach of Owechko et al [7], we can calculate $C(i)^{(n)}$

$$C(i)^{(n)} = \left(N\delta(i - io) + \sqrt{2N/3} \right)^n \tag{17}$$

So substituting equation (17) into equation (16) and approximately, we have

$$SNR \approx \left(\sqrt{3N/2}\right)^n \Big/ \sqrt{M} \qquad (N, M >> 1) \tag{18}$$

From equation (18), we can see that the memory capacity is roughly proportional to N^n [7]. So the theoretical network storage capacity depends on the order of the nonlinearity (the value of n in equation (15)). In our system, n is equal to the number of iterations. So the nonlinearity is not fixed and increases by one on each repeated iteration. It does not require many iterations to ensure that the memory capacity limit imposed by the net architecture is greater than the memory capacity limit imposed by optical components. Another interpretation is that for the same memory capacity repeated iterations reduce the error rate since this results in a greater differential gain and more averaging to reduce noise.

EXPERIMENTAL RESULTS

In our initial experiment, a single detector preceded by a pinhole was moved to detect the zero order correlations. The electronic detector output was fed back via an IBM/XT computer to control the intensity of the light incident on each hologram independently by means of the liquid crystal SLM. For system assessment three orthogonal training patterns (Fig. 4(a)) were stored. Optically it is difficult to deal with negative amplitudes so all of the data was made positive and the number of pixels was doubled. Each 16 pixel (4x4) binary pattern was reproduced twice, side by side, one copy being made the negative of the other, so the algorithm for the multiplication of two pixels of separate patterns is

1x1+0x0=1, 1x0+0x1=0, 0x1+1x0=0, 0x0+1x1=1

where 1 and 0 represent on and off pixels respectively. The product reflects the similarity (if the two pixels are the same (black or white), the product is 1, otherwise it is 0). So the actual patterns consisted of 8x4 pixels. When one quarter of an input pattern was obscured, the net converged to the complete correct pattern in 2 iterations (Fig. 4(b)).

Recently we have stored 14 different patterns, each with 32 pixels (so the actual patterns had 64 pixels). When an input with about 1/8 obscured was inserted into the net, the net could recognise the pattern and restore the hidden part in 3 or 4 iterations, depending on the injected pattern and the system alignment.

CONCLUSIONS

We have demonstrated the use of two dimensional arrays of Fourier Transform holograms in an opto-electronic implementation of a high order feedback net (HOFNET). The nonlinearity in this net is produced by means of feedback and not by optical material non-linearities and so the net is not limited by the switching speed, non-uniformity or availability of large arrays of opto-electronic nonlinear neurons. Spatially multiplexed holograms offer the advantages of no cross coupling, convenient addressability and, if formed in a photorefractive material, can be selectively erased and updated without interfering with other memories.

REFERENCES

1. J. J. Hopfield, 1982, "Neural networks and physical systems with emergent collective computational abilities", Proc. Natl. Acad. Sci. USA, 79, 2554-2558;
2. D. Psaltis, C. H. Park and J. Hong, 1988, "High order associative memories and their optical implementation", Neural Networks, 1, 149-163;
3. S. Lin and L. Liu, 1989, "Opto-electronic implementation of a neural network with a third-order interconnection for quadratic associative memory", Opt. Commu., 73, 268-272;
4. J. Jang, S. Shin and S. Lee, 1988, "Optical implementation of quadratic associative memory with outer-product storage", Opt. Lett., 13, 693-695;
5. J. Jang, S. Shin and S. Lee, 1989, "Programmable quadratic associative memory using holographic lenslet arrays", Opt. Lett., 14, 838-840;
6. P. Horan, D. Uecker and A. Arimoto, 1990, "Optical implementation of a second-order neural network discriminator model", Japanese J. of Appl. Phys., 29, 361-365;
7. Y. Owechko, G. J. Dunning, E. Marom and B. H. Soffer, 1987, "Holographic associative memory with nonlinearities in the correlation domain", Appl. Opt., 26, 1900-1910;
8. R. A. Athale, H. H. Szu and C. B. Friedlander, 1986, "Optical implementation of associative memory with controlled nonlinearity in the correlation domain", Opt. Lett., 11, 482-484;
9. D. R. Selviah, Z. Q. Mao and J. E. Midwinter, 1990, "Opto-electronic high order feedback neural network", Electronics Letters, 26, 1954-1955;
10. R. J. Collier, C. B. Burckhardt, and L. H. Lin, 1971, "Optical Holography", Academic Press.

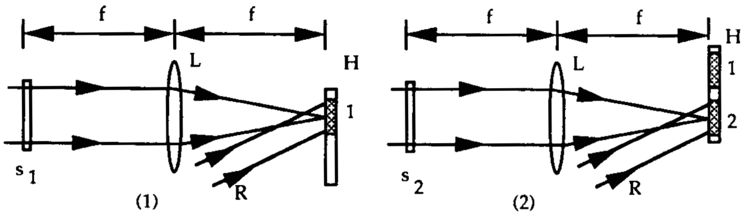


Fig. 1 Spatial Multiplexing Holographic Recording

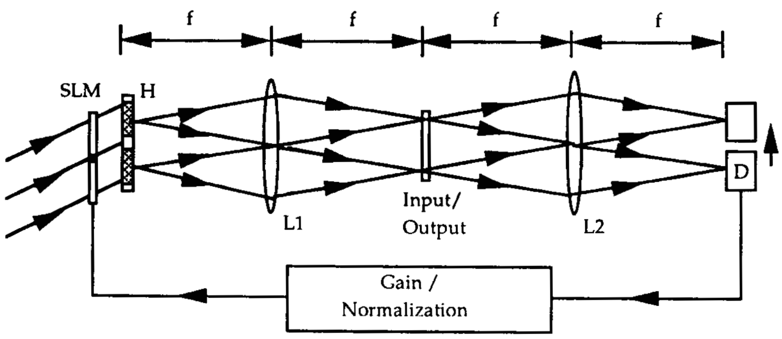
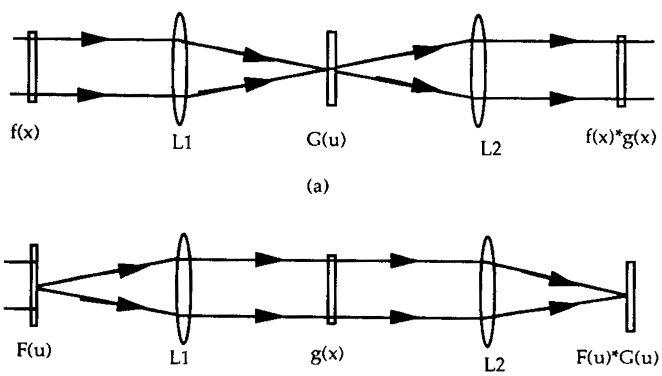


Fig. 2 Optical HOFNET Associative Recall System



(b)

Fig. 3 Correlation Systems
(a) Spatial correlation; (b) Spectrum correlation

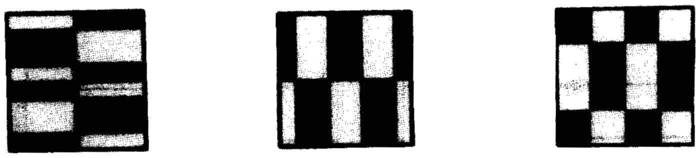
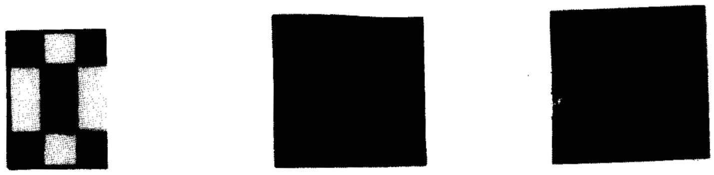


Fig. 4(a) Patterns Stored in the System



(1) Partial input (1/4 obscured)

(2) Output before iteration

(3) Output after two iterations

Fig. 4(b) Pattern Recognition Procedure