

Correlating matched-filter model for analysis and optimisation of neural networks

D.R. Selviah, MA
 Prof. J.E. Midwinter, OBE, FEng, FIEE, FRS
 A.W. Rivers, MSc
 K.W. Lung, BSc

Indexing terms: Active networks, Signal processing, Pattern recognition

Abstract: A new formalism is described for modelling neural networks by means of which a clear physical understanding of the network behaviour can be gained. In essence, the neural net is represented by an equivalent network of matched filters which is then analysed by standard correlation techniques. The procedure is demonstrated on the synchronous Little-Hopfield network. It is shown how the ability of this network to discriminate between stored binary, bipolar codes is optimised if the stored codes are chosen to be orthogonal. However, such a choice will not often be possible and so a new neural network architecture is proposed which enables the same discrimination to be obtained for arbitrary stored codes. The most efficient convergence of the synchronous Little-Hopfield net is obtained when the neurons are connected to themselves with a weight equal to the number of stored codes. The processing gain is presented for this case. The paper goes on to show how this modelling technique can be extended to analyse the behaviour of both hard and soft neural threshold responses and a novel time-dependent threshold response is described.

1 Introduction

Theoretical modelling techniques based on spin glasses [1-5] have been found to be useful in assessing neural network parameters, such as memory capacity. However, these models do not easily give a clear physical insight into the neural network behaviour. In this paper we propose a new modelling formalism based on matched-filters which seems better suited to the physical processes occurring in the neural network and which allows a physically intuitive grasp to be gained of the neural network behaviour. This leads to a useful insight into the operation of auto-associative memories from a signal-processing engineering perspective. The matched-filter formalism is applied initially in this paper to model a pattern-recognition neural network architecture and is used later to design a novel associative memory. The pattern-recognition network which is considered requires distorted, partial or noisy patterns as input and ideally

Paper 6670F (E16), first received 21st April 1988 and in revised form 10th March 1989

The authors are with the Department of Electronic and Electrical Engineering, University College, Torrington Place, London WC1E 7JE, United Kingdom

IEE PROCEEDINGS, Vol. 136, Pt. F, No. 3, JUNE 1989

outputs the stored binary, bipolar (+1, -1) pattern s_i ($i = 1$ to M) which most closely resembles the input.

The neural net architecture chosen for discussion in this paper is the synchronously operated Little-Hopfield neural net [6, 7] (having threshold levels of zero), which has recently received much attention due to its suitability for electronic and optical implementation [8-12]. The network is shown in Fig. 1, where an input unknown

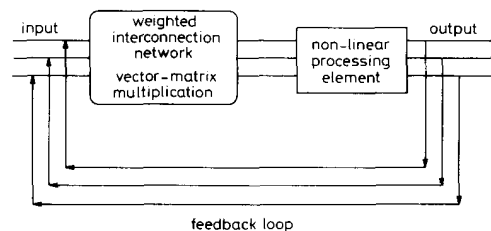


Fig. 1 Vector-matrix representation of a synchronous Little-Hopfield neural network

pattern vector is first multiplied by a 'memory' matrix. Subsequently, each bit of the resultant vector is thresholded such that if it is positive it is set to +1 and if negative to -1. These two operations are repeated until the circulating code converges onto a stable binary state, which is then output. The input vector consists of N pixels or bits, each of which may have any positive or negative analogue value. The symmetric matrix, T_{ij} , is formed from the known vector products as follows:

$$T_{\tau i} = \sum_{i=1}^M s_{i\tau} s_{i\tau}, \quad \tau \neq i \quad (1)$$

$$T_{ii} = 0 \quad (2)$$

where $s_{i\tau}$ is the τ th bit of the i th memorised vector and where M is the number of memorised vectors. In our numerical simulations we have found such a net to have several forms of behaviour:

- (a) convergence to the correct memorised state
 - (b) convergence to an incorrect memorised state
 - (c) nonconvergent oscillation between several states.
- This does not occur in the asynchronous Hopfield net in which only single bits are changed each time before thresholding, but in our formulation (similar to Little [6]) all bits can change their state on each iteration synchronously.

- (d) convergence to a spurious nonmemorised state.

Undesirable behaviour of types (b) and (c) are covered in Sections 2-5 and behaviour of type (d) in Section 6.

2 Correlating matched-filter formalism and its use for modelling the synchronous Little-Hopfield net

In this Section we introduce a pattern recognition network of matched-filters similar to that used by Grant [13] and then show, in the following Section, that it is formally identical to the vector-matrix multiplication operation in the synchronous Little-Hopfield neural network.

A pattern-recognition network must be able to recognise a signal when it is distorted or obscured by noise. A partial signal in which several bits are set to zero can be included by treating it as a special case of additive noise. Recognition requires the calculation of some closest-match criterion between the input signal r and each of a set of known, stored template signals s_i ($i = 1, 2, \dots, M$).

It is well known [14] from maximum entropy arguments that the best way to recognise the existence of a known signal in the presence of unknown, additive noise is to perform the correlation between the input noisy signal r and the known signal s and to see if this is greater than some threshold:

$$r * s_i = \sum_{\tau=1}^N r_{\tau} s_{i\tau} \quad (3)$$

where the t subscript indicates the bit index. The digital signals are assumed, for simplicity, to be one-dimensional time (t) dependent as, for example, in an electronic implementation of a neural network. However, it should be noted that all the reasoning is directly applicable to one- or two-dimensional signals such as may be found in an optical implementation [8–12] simply by replacing the dependent variable, time, by some spatially-dependent co-ordinate as done by Grant [13].

The correlating matched-filter model has two system components (Fig. 2) like the vector-matrix representation,

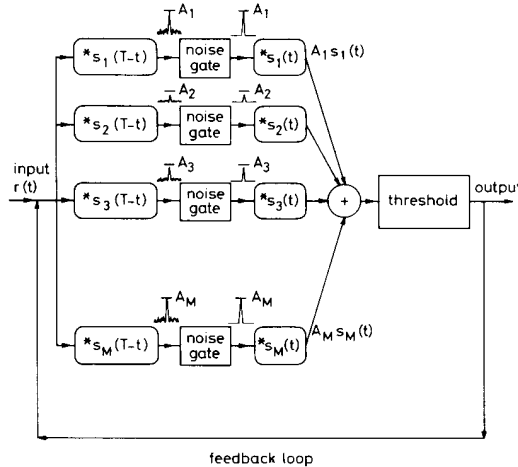


Fig. 2 Correlating matched-filter representation of a pattern-recognition synchronous Little-Hopfield neural network

but instead of representing the weighted interconnection net as a vector-matrix multiplier it is represented as a network of correlating matched-filters. When the input signal r is passed through a matched-filter with an impulse response of the signal complex conjugate $s_i^*(T-t)$ [which, since we assume real signals throughout, is equal to $s_i(T-t)$, where T is the total duration of the

stored signal] the output gives a measure of the closeness of match of r and s_i . Hence, to distinguish which code vector s_i is closest to the input vector r , it is necessary to distribute the input code in equal measure to each of a bank of matched-filters as shown in Fig. 2 and to search for the largest correlation output.

Only the magnitude of the central correlation spike is relevant for the degree of match between r and s_i since the central spike corresponds to an exact overlap of the memorised code and the input code, assuming they are both of equal length and assuming that there is no uncertainty in the time of arrival. Thus, the correlation side-lobes and much of the associated noise can be gated out. In an optical implementation this function might be performed by one pinhole for each filter as done by Abu-Mostafa and Psaltis [10]. Hence, a set of Dirac delta functions can be formed which are scaled by the magnitudes of the correlation peaks in each channel, as shown in Fig. 2. These pass into a second bank of matched-filters which have the impulse responses of the memorised codes $s_i(t)$. Each channel has, therefore, two filters having the impulse responses of the time reverse of a known code and of the known code itself. The output from each of the second filters is the corresponding code scaled by the magnitude of the correlation peak. The scaled output codes from each channel are summed together before being supplied to the nonlinear threshold. We show below that this interconnection network (Fig. 2), without its thresholding feedback, is mathematically identical (i.e. isomorphic) to the vector-matrix multiplication neural net (Fig. 1).

3 Proof of the equivalence of the matched-filter network and the vector-matrix multiplication neural network

The magnitude of the correlation peak following the initial i th-channel filter is given by the correlation

$$A_i = \frac{1}{M^{1/2}N} \sum_{\tau=1}^N r_{\tau} s_{i\tau} \quad (4)$$

where the τ subscripts indicate which bit of the code is being considered. The $1/M^{1/2}$ amplitude factor arises from the $1/M$ power loss which occurs when the signal is divided between the M channels. A delta function scaled by the magnitude, A_i , is generated and passes through the second filter to give as output

$$A_i s_{i\tau} = \frac{1}{M^{1/2}N} \sum_{\tau=1}^N r_{\tau} s_{i\tau} s_{i\tau} \quad (5)$$

All of the M channel outputs are summed to yield

$$\sum_{i=1}^M A_i s_{i\tau} = \frac{1}{M^{1/2}N} \sum_{i=1}^M \sum_{\tau=1}^N r_{\tau} s_{i\tau} s_{i\tau} \quad (6)$$

Rearranging eqn. 6 gives the output r'_i of the weighted interconnection network before thresholding as

$$r'_i = \frac{1}{M^{1/2}N} \sum_{\tau=1}^N r_{\tau} \left[\sum_{i=1}^M s_{i\tau} s_{i\tau} \right] \quad (7)$$

$$r'_i = \frac{1}{M^{1/2}N} \sum_{\tau=1}^N r_{\tau} T_{i\tau} \quad (8)$$

where we have defined the matrix $T_{i\tau}$ to be

$$T_{i\tau} = \sum_{i=1}^M s_{i\tau} s_{i\tau} \quad (9)$$

which implies that

$$T_{ii} = M \quad (10)$$

The output from the matched-filter network after one pass (eqn. 8) is the same as that through the vector-matrix network except for an extra $1/(M^{1/2}N)$ scaling factor in the case of the matched-filter network.

The matrix T_r (eqn. 9) is identical to the autocorrelation matrix as defined by Kohonen [15, 16] for use in a vector-matrix pattern recognition memory. Matrix T_r (eqns. 9 and 10), however, differs slightly from the matrix operator proposed by Hopfield (eqns. 1 and 2) in that it has diagonal elements of magnitude M (the number of stored state vectors) for bipolar (+1, -1) codes as opposed to elements of magnitude zero in the Hopfield operator.

4 Physical behaviour of the synchronous Little-Hopfield network

4.1 Influence of the diagonal matrix elements on the convergence behaviour of the network

The matrix element T_{ii} represents the weight of the bidirectional connection between the i th neuron and the i th neuron. Diagonal elements, therefore, represent connections from the same neuron to itself. Since self connected neurons have not been observed in the brain, the diagonal elements were set to zero in the original net as in eqn. 2. However, in the last Section we saw how matrix elements of magnitude M arise from the optimum signal detection strategy based on performing correlations and so should lead to better performance. Physically, a connection of a neuron to itself with a weight M tends to stabilise the potential on the neuron (this weight is equal to, or larger than, any other weight) and thus acts as a memory between iterative cycles.

Mathematically, the diagonal elements act as an identity matrix and so are responsible for projecting the input vector through the matrix to stabilise the pattern. This has the effect of substantially reducing the oscillatory non-convergent behaviour observed in synchronous nets with zero diagonal elements. Improved behaviour has also been observed by Gindi *et al.* [17] with nonzero diagonal elements.

4.2 Relationship between Hopfield energy and net input/output crosscorrelation

In the spin glass model of neural networks an energy function, or Hamiltonian, is defined which plays a central role in neural net behavioural analysis. This function is defined [7] to be

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{r=1}^N r_i T_{ii} r_i \quad (11)$$

where matrix T has its diagonal elements set to zero. This is almost identical to the correlation of the output of the matched-filter network (before thresholding) with the input

$$C = \frac{1}{M^{1/2}N^2} \sum_{i=1}^N \sum_{r=1}^N r_i T_{ii} r_i \quad (12)$$

apart from the fact that the diagonal elements of the memory matrix are not zero in the matched-filter case (Section 4.1) and that the proportionality constant is $1/M^{1/2}N^2$ as opposed to $-1/2$ in the conventional expression. The physical interpretation is that the output of the matched-filter net (before thresholding) starts off

being quite different to the input (i.e. small correlation, high energy) and gradually begins to resemble the feedback input until they become identical (correlation peak, energy minimum).

4.3 Optimum memorised codes for signal discrimination in noise

Consider the input to the net to be one of the memorised codes plus noise. Clearly, by minimising the cross-correlation between memorised codes one maximises the difference in correlation magnitudes emerging from the first bank of filters, so improving the ability to discriminate the largest peak in the greatest amount of noise. This is optimised by choosing the stored code set to be mutually orthogonal. Confirmation is provided by the fact that our matrix (eqns. 9 and 10) reduces to Kohonen's optimum memory matrix [15, 16] (based on the generalised inverse matrix) when the memorised codes are chosen to be orthogonal. This limitation is overcome by the novel net proposed in Section 5. A net storing orthogonal codes cannot converge onto an incorrect memorised code.

4.4 Network filtering behaviour, upper limit on memory capacity and processing gain

In this Section we assume that the stored codes are orthogonal for the reasons given in the previous Section. The maximum number of mutually orthogonal N -bit bipolar (+1, -1) codes that can be found is N and they only exist when $N = 2^n$, where n is an integer. Numerous sets of N mutually orthogonal codes exist.

Consider the case when N , N -bit orthogonal codes are stored in a synchronous Little-Hopfield neural net. An arbitrary digital (bits having any analogue magnitude) code is input. This can be expressed as a series expansion in terms of the complete set of orthogonal codes:

$$r = \sum_{i=1}^N \left[\sum_{r=1}^N r_i o_{ir} \right] o_i \quad (13)$$

This is the digital equivalent (in terms of codes on a closed interval) of a Fourier transform (in terms of sine waves on an infinite interval).

The intermediate A_i correlation peak coefficients after the first bank of filters are effectively the 'spectrum' in terms of orthogonal digital codes and contain all the information of the original code.

The second filter bank and the summation perform the equivalent of an inverse digital orthogonal code transform and so reconstruct exactly the original input code apart from a $1/M^{1/2}$ scaling-loss factor. This is equivalent to performing the inverse Fourier transform. Obviously, such a net is then equivalent to a straight parallel highway and performs no useful function. This, therefore, sets an upper limit on the number of codes that can be stored, which is N . Abu-Mostafa and St. Jacques [18] recently obtained the same result by information theory. Earlier, Cover [19] obtained an upper bounded capacity which was a factor of just two higher.

When one of the N stored orthogonal code channels is removed or when an infinite attenuator is employed instead of the noise gate, the net acts as a filter preventing any such code contribution from passing through. For M ($< N$) stored codes the output from the matched-filter net before thresholding only consists of the sum of varying amounts of the stored codes. If a known code plus white noise, which contains equal RMS (time averaged) powers of each of the orthogonal codes, were supplied as input,

the net output noise would be reduced by a factor of M/N compared to the output code. This results in an overall processing gain for the interconnection net of

$$PG = 10 \log_{10} (N/M)$$

5 Design of a novel associative memory neural network using matched-filters

Once the matched-filter modelling concept has been grasped, numerous new neural nets can be designed. Here, we present one particularly useful associative memory which can be used to associate arbitrary input and output, digital or analogue, stored patterns. It consists (Fig. 3) of an optimum synchronous Little-Hopfield pattern-recognition neural net storing orthogonal binary, bipolar codes sandwiched between two pattern-association nets. The first pattern-association net associates arbitrary analogue or digital patterns, s_1 , stored in the first matched-filter bank with orthogonal code 'labels'. The central pattern-recognition net converges

with the strongest stored code in the input). The final pattern-association net maps the strongest code to an arbitrary analogue or digital code s_3 stored in the final matched-filter bank. In the language of neural nets this net consists of three layers with internal feedback in the central layer, which has its weights preprogrammed prior to training.

6 Threshold behaviour

We initially consider the behaviour of hard thresholds (Fig. 4) and go on to look at soft and time-dependent thresholds later. The output of the interconnection net (i.e. the input to the array of nonlinear thresholding elements) was shown by the use of the matched-filter model to consist of the sum of only the known memorised codes. The magnitudes of each are weighted according to how closely they resemble the input. We show that the hard threshold acts together with the summation at the output of the interconnection network to emphasise the stronger memorised codes and to suppress the weaker on repeated iterations.

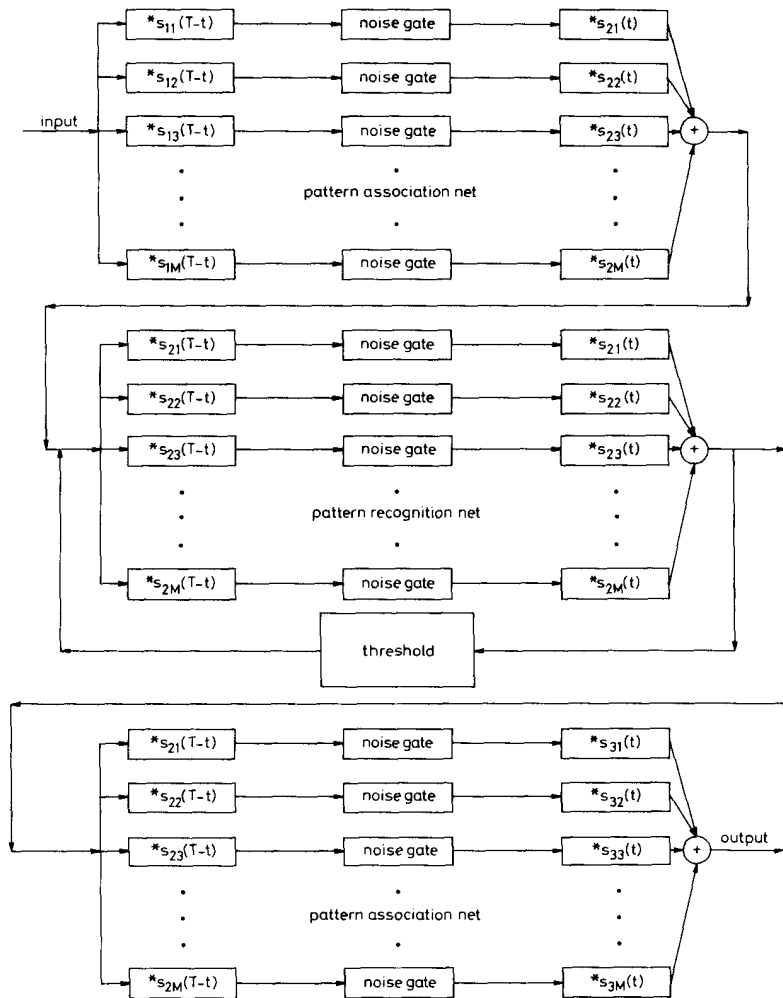


Fig. 3 Pattern-association net with three layers of interconnections for arbitrary analogue or digital pattern association

When several binary, bipolar codes are added, if all the i th bits have the same polarity then the sum code i th bit has the largest amplitude (assumed greater than unity) of

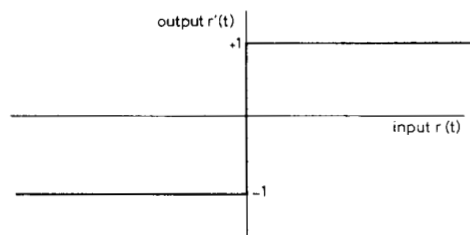


Fig. 4 Hard threshold response

all the bits and is suppressed by the hard threshold. Conversely, if all the i th bits differ in polarity then the sum code i th bit will have a small amplitude (assumed less than unity) and will be amplified by the hard threshold. The greatest amplification takes place if the polarity of the two strongest codes differ and if the next two strongest codes differ, and so on. In summary, the summation and hard threshold have the effect of emphasising the differences between the codes and suppressing the similarities so that they act as a feature extractor. In this way the code sum entering the interconnection net on the second iteration consists of the sum of nonbinary, bit-weighted feature codes with the same polarity sequence as the corresponding memorised codes. These correlate with their corresponding memorised codes in the matched-filters to give a large output and with the others to give a low output. The difference between the outputs of the correlation filters is larger than before due to the feature extraction performed by the summation and threshold. Thus, repeated iterations select the strongest memorised code component.

However, if too many of the memorised code components of the input are large and similar in magnitude, convergence to a spurious *nonmemorised* code can occur. In this case, the hard threshold, which is an odd function, generates from the input codes new output codes which are products of odd numbers of the input codes according to a power series expansion of the threshold function. Spurious codes only occur when the input (to the threshold) consists of three or more memorised codes since it is impossible to form a binary, bipolar output code using less. The probability of converging to a spurious code increases as the fractional memory capacity M/N , for a given number of neurons, increases and this limits the available capacity for a given error rate.

This undesirable behaviour is due to the strength of the threshold nonlinearity, that is, the coefficients of the higher order odd terms (third, fifth, etc.) in a series expansion of the threshold. This can be overcome by using a soft threshold (Fig. 5) with weaker higher order terms since this reduces the number of stable spurious states at the expense of the increased number of iterations which are required before convergence due to the lower threshold gain. To reduce the convergence time we propose a novel method reminiscent of simulated annealing in a Boltzmann machine, namely the use of a threshold with a time-dependent response. Start with a hard threshold which causes fast convergence to a probably spurious state. Gradually reduce the threshold gain so that it becomes softer on repeated iterations. This causes the spurious state to become unstable and the net to home in on the next stable state, which may also be spurious.

After repeated iterations the net ends up in the lowest energy minimum. To obtain the fastest overall convergence time the threshold gain must be decreased

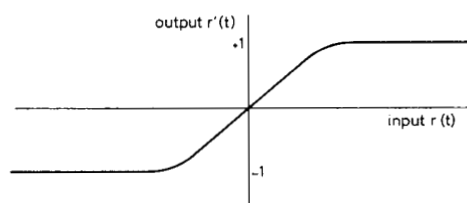


Fig. 5 Soft threshold response

slowly enough to ensure that convergence is almost completed before the spurious state becomes unstable. This procedure differs from that of the Boltzmann machine in that the energy surface is initially full of local minima which gradually disappear until only one remains whereas in the Boltzmann machine the energy surface is fixed but the convergence path is guided towards the lowest minimum. By further reducing the rate of threshold 'softening', higher fractional memory capacity can be obtained.

7 Conclusions

A new neural network modelling technique has been described in which the neural net is represented by an equivalent net of matched-filters. This enables a physically intuitive understanding of the network behaviour to be gained. The synchronous Little-Hopfield net was analysed and optimised in this way. The use of orthogonal stored codes prevented convergence to incorrect stored codes. The connection of the neurons back to themselves with a weight equal to the number of stored codes reduced the possibility of oscillatory nonconvergence. The use of a novel time-dependent hard and soft threshold response avoided the chance of converging to a spurious nonstored state and so increased the memory capacity. A novel three-layer neural net architecture was proposed incorporating the optimised synchronous Little-Hopfield net as a building block. This net can perform arbitrary analogue or digital pattern association in only a few iterations with low error rate and high memory capacity. Finally, it is worth noting that the equivalent matched-filter net can be implemented directly using electronic or optical matched-filters; this offers a new way of implementing neural nets using standard building blocks.

8 Acknowledgment

The authors wish to express their grateful thanks to Dr. M. Shirivani, Prof. P.M. Grant, Prof. I. Aleksander, Prof. P. Treleaven, Prof. W. Stuart, and Prof. A. Tanguay for most helpful discussions and to the Digital Optics Joint Optoelectronic Research Scheme of Great Britain for financial support.

9 References

- PERSONNEZ, L., GUYON, I., and DREYFUS, G.: 'Information storage and retrieval in spin-glass like neural networks', *J. Phys. Lett.*, 1985, **46**, pp. L359-L365
- AMIT, D.J., GUTFREUND, H., and SOMPOLINSKY, H.: 'Spin-glass models of neural networks', *Phys. Rev. A*, 1985, **32**, (2), pp. 1007-1018
- AMIT, D.J., GUTFREUND, H., and SOMPOLINSKY, H.: 'Storing infinite numbers of patterns in a spin-glass model of neural networks', *Phys. Rev. Lett.*, 1985, **55**, (14), pp. 1530-1533

- 4 KANTER, I., and SOMPOLINSKY, H.: 'Associative recall of memory without errors', *Phys. Rev. A*, 1987, **35**, (1), pp. 380-392
- 5 AMIT, D.J., GUTFREUND, H., and SOMPOLINSKY, H.: 'Information storage in neural networks with low levels of activity', *Phys. Rev. A*, 1987, **35**, (5), pp. 2293-2303
- 6 LITTLE, W.A.: 'Existence of persistent states in the brain', *Math. Biosci.*, 1974, **19**, pp. 101-120
- 7 HOPFIELD, J.J.: 'Neural networks and physical systems with emergent collective computational abilities', *Proc. Natl. Acad. Sci. USA*, 1982, **79**, pp. 2554-2558
- 8 PSALTIS, D., and FARHAT, N.: 'Optical information processing based on an associative-memory model of neural nets with thresholding and feedback', *Opt. Lett.*, 1985, **10**, (2), pp. 98-100
- 9 FARHAT, N.H., PSALTIS, D., PRATA, A., and PAEK, E.: 'Optical implementation of the Hopfield model', *Appl. Opt.*, 1985, **24**, (10), pp. 1469-1475
- 10 ABU-MOSTAFA, Y.S., and PSALTIS, D.: 'Optical neural computers', *Sci. Am.*, March 1987, pp. 66-73
- 11 SELVIAH, D.R., and MIDWINTER, J.E.: 'Pattern recognition using opto-electronic neural networks'. IEE Colloquium Digest on Optical techniques for information processing, No. 1987/105, 1987, pp. 6/1-6/4
- 12 MIDWINTER, J.E., and SELVIAH, D.R.: 'Digital neural networks, matched-filters and optical implementations'. Chap. 13 in ALEXANDER, I. (Ed.): 'Neural computing architectures' (Kogan Page, 1989), pp. 258-278
- 13 GRANT, P.M., and SAGE, J.P.: 'A comparison of neural network and matched-filter processing for detecting lines in images'. American Inst. of Phys. Conf. Proc. No. 151 on Neural networks for computing, 1986, pp. 194-200
- 14 CHAMBERS, W.G.: 'Basics of communications and coding' (Clarendon Press Publishers, 1985), pp. 101-103
- 15 KOHONEN, T.: 'Correlation matrix memories', *IEEE Trans.*, 1972, **C-21**, (4), pp. 353-359
- 16 KOHONEN, T.: 'Self organisation and associative memory' (Springer-Verlag, 1984), pp. 49-52, 96, 111-113, 187
- 17 GINDI, G.R., GMITRO, A.F., and PARTHASARATHY, K.: 'Hopfield model associative memory with nonzero-diagonal terms in memory matrix', *Appl. Opt.*, 1988, **27**, (1), pp. 129-134
- 18 ABU-MOSTAFA, Y.S., and ST. JACQUES, J.-M.: 'Information capacity of the Hopfield Model', *IEEE Trans.*, 1985, **IT-31**, (4), pp. 461-464
- 19 COVER, T.M.: 'Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition', *IEEE Trans.*, 1965, **EC-14**, p. 326