R.J. Marks II, L.E. Atlas, J.J. Choi, S. Oh and D.C. Park, "Nonlinearity requirements for correlation based associative memories", Proceedings of O-E/LASE '88 Conference on Optical Computing and Nonlinear Materials, Los Angeles, January 1988, SPIE volume 881, pp 179-183.

PROCEEDINGS

Of SPIE-The International Society for Optical Engineering

Volume 881

Optical Computing and Nonlinear Materials

Nasser Peyghambarian
Chair/Editor

Sponsored by
SPIE—The International Society for Optical Engineering

Cooperating Organizations
American Academy of Otolaryngology—Head and Neck Surgery
American Society for Laser Medicine and Surgery
Applied Optics Laboratory/New Mexico State University
Beckman Laser Institute and Medical Clinic, University of California/Irvine
Center for Applied Optics Studies/Rose-Hulman Institute of Technology
Center for Applied Optics/University of Alabama in Huntsville
Center for Devices and Radiological Health, FDA
Center for Electro-Optics/University of Dayton
Center for Laser Studies/University of Southern California
Center for Optical Data Processing/Carnegie Mellon University
European Laser Association
Georgia Institute of Technology
Institute of Optics/University of Rochester
Optical Sciences Center/University of Arizona
SPSE—The Society for Imaging Science and Technology

11-13 January 1988
Los Angeles, California

Nonlinearity Requirements for Correlation Based Associative Memories

R.J.Marks II, L.E.Atlas, J.J.Choi, S.Oh, K.F.Cheung, D.C.Park

University of Washington, Interactive System Design Lab
FT-10, Seattle, WA. 98195

ABSTRACT

A matched filter based architecture for associative memories (**MFAM**) has been proposed by many researchers[1-12]. The correlation from a leg of a matched filter bank, after being altered nonlinearly, weights its corresponding library vector. The weighted vectors are summed and clipped to give an estimate of the library vector closest to the input. We analyze the performance of such architectures for binary and/or bipolar inputs and libraries. Sufficient conditions are derived for the correlation nonlinearity so that the MFAM outputs the correct result. If, for example, N bipolar library vectors are stored, then the correlation nonlinearity $Z(x) = N^{x/2}$ will always result in that library vector closest to the input in the Hamming sense.

1. INTRODUCTION

In detection theory, the matched filter is optimum in many scenarios [13-15]. Thus, under similar conditions, a matched filter followed by a table lookup would be the optimal architecture for an associative memory [2]. In terms of implementation, however, it is more straightforward to use the correlations of the matched filter output to weight each corresponding library object [6,7,11,12]. The weighted terms are summed to give an estimate of the memory output. If the library is known to be binary, this output can be clipped to hopefully improve the result. The output can then, in turn, be iteratively fed into the memory. Certain neural network associative memories [1,4,5] are algorithmically identical to this procedure [2,3].

In this paper, we show that for binary objects, matched filter architectures are optimum for finding that library object closest in the Hamming sense to the input. The use of nonlinearities in the correlation domain is also considered. Sufficient constraints for the matched filter associative memory (MFAM) to operate successfully are explored for a number of cases.

2. A MATCHED FILTER ASSOCIATIVE MEMORY(MFAM)

In this section, we develop a generalized matched filter approach to associative memory architectures. Let $\mathcal{F} = \{\vec{f}_n | 1 \le n \le N\}$ denote a set of bipolar (+1,-1) library vectors of length L and $\vec{g}$ a perturbed bipolar version of one of the library vectors. An associative memory architecture for finding that vector is shown in Fig.I.

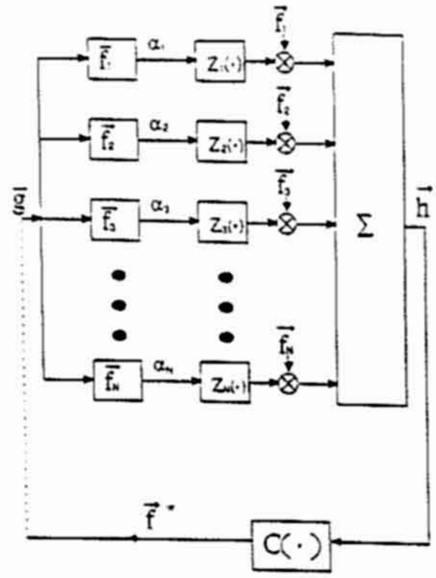

Figure I: Block diagram of a matched filter based associative memory (MFAM) with nonlinearities in the correlation domain.

The input vector is correlated with each library vector to form:

$$\alpha_n = \vec{g}^T \vec{f}_n$$

or, in vector form:

$$\vec{\alpha} = \underline{F}^T \vec{g}$$

where the library matrix is defined by

$$\underline{F} = [\vec{f}_1 | \vec{f}_2 | \vec{f}_3 | \ldots\ldots | \vec{f}_N]$$

Each correlation coefficient is then operated on by a point non-linearity $Z_n(.)$ which, in turn, weights the corresponding library vector. The weighted vectors are summed to obtain

$$\vec{h} = \sum_{n=1}^{N} z_n(\alpha_n) \vec{f}_n$$

Each element of $\vec{h}$ is fed through a hard limiter

$$C[\cdot] = \mathrm{sgn}[\cdot]$$

The output is thus

$$\vec{f}^* = C\left[\sum_{n=1}^{N} z_n(\alpha_n) \vec{f}_n\right]$$

Ideally, $\vec{f}^*$ should be that library vector closest to $\vec{g}$ in some sense.

3. EXTENSIONS

1. The output, $\vec{f}'$,could be again fed into the processor to yield a hopefully better result. We refer to such a processor as an iterative matched filter associative memory (**IMFAM**).

2. We will also want to examine the MFAM performance when the input or matched filter bank is binary(0,1). As such, define

$$\hat{\vec{f}}_n = \frac{1}{2} (\vec{f}_n + \vec{1})$$

$$\hat{\vec{h}}_n = \frac{1}{2} (\vec{h}_n + \vec{1})$$

where $\vec{1}$ is a vector of 1's. Clearly, the hatted vectors are binary when the unhatted vectors are bipolar (-1,+1). For binary inputs, the hard limiter performs a unit step operation.

4. SPECIAL CASES

1. When operated iteratively with a binary input and bipolar matched filter, IMFAM performs algorithmically similar to Hopfield's content addressable memory (**CAM**) when $Z_n(x) = x$ [1,5].

2. Soffer *et. al.*[7] implemented a nonlinear holographic associative memory with feedback using a hologram and phase-conjugate mirrors (PCM's). The hologram acts the memory element. The PCM's were used for feedback, thresholding and amplification to improve the correlation between the input object and desired image.

3. Athale *et. al.*[10] suggested library vector-dependent weights for 'attentive' associative memory. Specifically, $Z_n(x) = a_n x$ where a_n is an attentive factor.

4. Psaltis *et. al.*[6] used square law and unit step nonlinearities in a MFAM optical architecture.

5. We justify a form of MFAM from a receiver design point of view. Let $\vec{g}$ be a library vector perturbed by additive white Gaussian noise. The corresponding decision function is [13]

$$d(\vec{f}_n) = \frac{Q_n}{(2\pi\sigma^2)^{L/2}} e^{-\frac{\| \vec{g} - \vec{f}_n \|^2}{2\sigma^2}}$$

where Q_n is a priori probability, i.e., $Q_n = Pr\ \{\vec{f}_n\}$. Hence if the decision function for $\vec{f}_n$ is comparatively large, it should dominate the expression

$$\vec{u} = \sum_{n=1}^{N} d(\vec{f}_n)\ \vec{f}_n \qquad (1)$$

If the library vectors are bipolar, then $\|\vec{f}_n\|^2 = L$ and we can write

$$\vec{u} = a \sum_{n=1}^{N} Q_n\ e^{\frac{\vec{f}_n^T \vec{g}}{\sigma^2}}\ \vec{f}_n$$

where

$$a = \frac{1}{(2\pi\sigma^2)^{L/2}}\ e^{\frac{(L + \| \vec{g} \|^2)}{2\sigma^2}}$$

is simply a proportionality constant. Thus, $\vec{h}$ is proportional to $\vec{U}$ when we use

$$Z_n(x) = Q_n\ e^{x/\sigma^2} \qquad (2)$$

If all priors are equal (i.e. $Q_n = 1/N$) then they may be absorbed into the constant.

5. HAMMING DISTANCE AND CORRELATION

Matched filters are used largely due to their optimal performance in certain scenarios. If $\vec{g}$ is a library vector perturbed by white Gaussian noise, then choosing the library vector corresponding to the largest correlation coefficient results in minimum probability of error and maximum SNR [13,14]. We show in four cases below where the matched filter also minimizes the Hamming distance.

1. Let $\vec{g}$ be a bipolar library vector at a Hamming distance of h_n from the bipolar vector $\vec{f}_n$. Then

$$\alpha_n = L - 2h_n \qquad (3)$$

where L is the length of the library vectors. Thus, maximizing correlation is the same as minimizing the Hamming distance.

2. In some situations (e.g. Hopfield's model) the binary vectors are stored at the memory as bipolar forms. To observe this binary input, bipolar matched filter case, the correlation coefficient is

$$\beta_n = \frac{1}{2} (\vec{g} + \vec{1})^T \vec{f}_n$$

$$= \frac{1}{2} \alpha_n + \frac{1}{2} S_n$$

where S_n is the sum of the elements in $\vec{f}_n$. Substituting Eq.(3) gives

$$\beta_n = \frac{1}{2} (L + S_n) - h_n \qquad (4)$$

If we assume $S_n = \sum_{m=1}^{L} f_{nm} = S$ for all n,** Eq.(4) becomes

$$\beta_n = \frac{1}{2} (L + S) - h_n$$

Then maximizing the binary-bipolar correlation is the same as minimizing the Hamming distance.

** This constraint has a similar flavor to that of requiring each library vector has the same energy. Such constraints are common in matched filter detectors [13].

3. If both input and library vectors are binary, then the correlation coefficients are

$$\gamma_n = \frac{1}{4} (\vec{g} + \vec{1})^T (\vec{f}_n + \vec{1}) \qquad (5)$$

$$= \frac{1}{2} L + \frac{1}{4} (S_g + S_n) - \frac{1}{2} h_n$$

where S_g and S_n are the total number of 1's in $\vec{g}$ and $\vec{f}_n$ respectively. Again, if $S_n = S$, Eq.(5) can be maximized by minimizing h_n.

4. Finally, the bipolar input-binary filter correlation coefficient is

$$\lambda_n = \frac{1}{2} \vec{g}^T (\vec{f}_n + \vec{1})$$

$$= \frac{1}{2} (S_g + L) - h_n$$

Thus, we have maximum correlation by choosing minimum Hamming distance.

Above four cases are summarized in table I.

Input Vector \ Library Element	Bipolar	Binary
Bipolar	$\alpha_n = L - 2h_n$	$\lambda_n = (1/2)\ (S_g+1) - h_n$
Binary	$\beta_n = (1/2)\ (L+S_n) - h_n$	$\gamma_n = (1/2)\ (L-h_n) + (1/4)\ (S_g+S_n)$

Table 1: Correlation coefficients for the four possible combinations of bipolar-binary / input-library. For bipolar inputs, maximizing the correlation minimizes the Hamming distance. This is also true for binary inputs if the sum of the elements of each library vector, in bipolar form, is the same.

6. MFAM PERFORMANCE

Under what conditions will the MFAM output the library vector closest in Hamming distance to the input? We consider three cases, all of which assume homogeneous nonlinearities (i.e. $Z_n(\alpha) = Z(\alpha)$ for all n) and bipolar inputs and libraries. Similar results can be obtained for the binary input-bipolar library and bipolar input-binary library cases. The remaining binary input - binary filter case will not work on the MFAM without modification of the sgn clip (i.e. h is always non-negative).

6.1. Strictly Increasing Nonlinearities.

Here, we require that $Z(x)$ be strictly increasing function and that the library vectors are separated sufficiently such that

$$\vec{f}_p^{\,T} \vec{f}_q \le \alpha \qquad ; p \ne q \qquad (6)$$

where α represents the magnitude of largest correlation coefficient among the library elements. Note that, if $\vec{f}_p^{\,T} \vec{f}_q = L \cdot \delta(p - q)$, the library vectors are orthogonal. Let α_{max} be the maximum correlation between the input and library vectors:

$$\alpha_{max} = \max_{1 \le n \le N} (\vec{g}^T \vec{f}_n) = L - 2k \qquad (7)$$

where k is the minimum Hamming distance between $\vec{g}$ and the library. We will assume there is a unique library vector, $\vec{f}_{max}$, that results in this maximum correlation.

As a consequence of our assumptions, it follows that

$$\alpha_n \le 2k + \alpha \qquad ; n \ne max \qquad (8)$$

The input to the clipper is

$$\vec{h} = \sum_{n=1}^{N} Z(\alpha_n)\ \vec{f}_n$$

or, in vector component form,

$$h_\ell = \sum_{n=1}^{N} Z(\alpha_n)\ f_{n\ell}$$

$$= Z(\alpha_{max})\ f_{max,\ell} + \sum_{n \ne max}^{N} Z(\alpha_n)\ f_{n\ell}$$

In order for $\underline{sgn}(h_\ell) = f_{max,\ell}$, we require that

$$h_\ell \begin{array}{c} f_{max,\ell} = 1 \\ \gtrless \\ f_{max,\ell} = -1 \end{array} 0$$

or equivalently, that

$$h_\ell\ f_{max,\ell} > 0$$

Since a<b implies that $Z(a) < Z(b)$ and $f_{n\ell}\ f_{max,\ell} = \pm 1$, we write

$$h_\ell\ f_{max,\ell} = Z(\alpha_{max}) + \sum_{n \ne max}^{N} Z(\alpha_n) f_{n\ell}\ f_{max,\ell}$$

$$\ge Z(L - 2k) - \sum_{n \ne max}^{N} Z(\alpha_n)$$

$$\ge Z(L - 2k) - (N-1)\ Z(2k + \alpha)$$

where we have used Eq.(7) and Eq.(8). Thus, if

$$Z(L-2k) > (N-1)\ Z(2k + \alpha) \qquad (9)$$

then the (noniterative) associative memory in Fig.1 will give the desired result. For orthogonal library elements, the constraint becomes

$$Z(L-2k) > (N-1)\ Z(2k)$$

We consider two special cases:

(a) Odd Power Nonlinearities

$Z(x) = x^{2Q+1}$; $Q=0,1,2,\ldots$

From Eq.(9), for a given L, N, α and Q, the MFAM will produce the correct output if the input Hamming distance obeys,

$$k < \frac{L - \alpha (N-1)^{\frac{1}{2Q+1}}}{2[(N-1)^{\frac{1}{2Q+1}} + 1]}$$

or, for orthogonal library vectors,

$$k < \frac{L}{2[(N-1)^{\frac{1}{2Q+1}} + 1]}$$

For $Q=0$ we are at the equivalent of Hopfield's associative memory neural net for bipolar inputs and library. Here, $k < L/2N$ assures one step convergence [2].

We also can compute the minimum Q required for proper MFAM operation by solving Eq.(9) for a given L, k and α The result is

$$2Q + 1 > \frac{\ln(N-1)}{\ln\left(\frac{L-2k}{2k+\alpha}\right)}$$

for nonorthogonal library elements and

$$2Q + 1 > \frac{\ln(N-1)}{\ln\left(\frac{L}{2k} - 1\right)}$$

for orthogonal elements.

(b) An Exponential Nonlinearity

$$Z(x) = e^{ax} \quad ; \quad a \geq 0 \qquad (10)$$

This nonlinearity is motivated from Eq.(2). Using Eq.(9), we can find a proper condition for k which guarantees convergence.

$$k < \frac{1}{4}\left(L - \alpha - \frac{\ln(N-1)}{a} \right)$$

From a design point of view, we would require

$$a > \frac{\ln(N-1)}{L - 4k - \alpha} \qquad (11)$$

6.2. Exponential Nonlinearities.

The exponential nonlinearity in Eq.(10) has an additional convergence property that warrants special attention. If there is only one library vector, $\vec{f}_{max}$, that produces α_{max}, then

$$h(\vec{f}_n, \vec{g}) - h(\vec{f}_{max}, \vec{g}) \geq 1$$

where h denotes the Hamming distance. Equivalently

$$\alpha_{max} - \alpha_n \geq 2 \qquad ; \ n \neq max$$

From Eq.(7) and (8), it follows that

$$(L - 2k) - (2k + \alpha) \geq 2$$

Thus, from Eq.(11), we have the weaker convergence criterion

$$a > \frac{1}{2} \ln(N-1) \qquad (12)$$

This constraint is not parameterized by α or k. Thus, assuming equality in Eq.(12) is sufficient, the nonlinearity

$$Z(x) = (N-1)^{x/2}$$

will always produce the correct output if there is a unique solution.

6.3. Even Power Nonlinearities

Physics many times dictates nonlinearities of the form

$Z(x) = x^{2Q}$; $Q=1,2,3,\ldots$

(e.g., square law detectors). Such nonlinearities have the disadvantage of emphasizing negative and positive correlations equally. The analysis of the strictly increasing nonlinearity, however, is applicable here if, in lieu of Eq.(6), we require

$$|\vec{f}_p^{\,T} \vec{f}_q| \leq \alpha \qquad ; \ p \neq q$$

Large negative correlations are then taken into account. Following the analysis, we find that Eq.(9) is still applicable. It then follows that

$$2Q > \frac{\ln(N-1)}{\ln\left(\frac{L-2k}{2k+\alpha}\right)}$$

The technique is also directly applicable to other even nonlinearities that are strictly increasing for positive argument such as $Z(x) = \exp(a|x|)$ or polynomials in $|x|$ with positive coefficients.

7. EXTENSION TO CONTINUOUS BINARY OBJECTS

An object, $g(x,y)$, is said to be bipolar if it takes on values of only ± 1. For such objects, the Hamming distance can be defined as the area over which one object from the other. If $\mathcal{A}$ denotes this area, then, Eq.(3) is applicable if we let

$$L = \int_{\mathcal{A}} dxdy$$

and

$$\alpha_n = \int_{\mathcal{A}} g(x,y) f_n(x,y) \, dxdy$$

where g and f_n are, respectively, the bipolar input and n^{th} bipolar library object. There are similar expressions for when the input and/or library is binary. Further application of this generalization to results in this paper is obvious.

8. CONCLUSION

We have shown that, for binary (bipolar) objects, the matched filter indicates that library element closest to the input in the Hamming sense. The performance of non iterative matched filter associative memories was also analyzed. Sufficient conditions for desired performance were derived for a number of cases.

9. ACKNOWLEDGEMENTS

This work was supported by the SDIO/IST's program in Ultra High Speed Computing administered through ONR in conjunction with Texas Tech University. Les.E. Atlas was also supported by an NSF PYI award. The authors gratefully acknowledge the inputs of Dziem Nguyen, Fred Holt and Don Wunsch of the Boeing High Technology Center, and Jim Ritcey of the Interactive Systems Design Lab.

10. REFERENCES

1. J.J. Hopfield, " Neural Networks and Physical Systems with Emergent Collective Computational Abilities" Proc. Natl. Acad. Sci. USA. 79, pp 2554-2558, (1982).

2. R.J.Marks II, L.E.Atlas,"Content address memories: a relationship between Hopfield's neural net and an iterative matched filter", submitted for publication.

3. L.E.Atlas, T.Homma, R.J.MarksII, "A neural network model for vowel classification ", Presented at Int.Conf. Acoust.,Speech and Sig. Proc., Dallas, TX, Apr.1987.

4. N.H.Farhat, D.Psaltis, Aluizio, E.Paek," Optical implementation of the Hopfield model", Appl. Opt. 24 pp 1469-1475 (1985).

5. D.Psaltis, N.H.Farhat," Optical information processing based on an associative memory model of neural nets with thresholding and feedback", Opt. Lett.10,pp98-100 (1985).

6. D.Psaltis, J.Hong," Shift - invariant Optical associative memories ", Opt.Eng.26, pp10-15 (1987).

7. B.H. Soffer, G.J.Dunning, Y.Owechko and E.Marom, " Associative holographic memory with feedback using phase-conjugate mirrors.", Opt.Lett 11, (1986).

8. G.J.Dunning, E.Marom, Y.Owechko, B.H.Soffer," Optical holographic associative memory using a phase conjugate resonator",SPIE Opt. Computing 625, pp205-213 (1986).

9. G.Dunning, E.Marom, Y.Owechko, B.Soffer," All-optical associative memory with shift invariance and multiple image recall", Opt.Lett. 12, pp346-348 (1987).

10. R.Athale, H.Szu,C.Freidlander," Optical implementation of associative memory with controlled nonlinearity in the correlation domain", Opt.Lett.11, pp482-484 (1986).

11. E.Paek, D.Psaltis, "Optical associative memory using Fourier Transform holograms", Opt. Eng.26, pp428-433 (1987).

12. Y. Owechko, G.J.Dunning,E.Marom B.H.Soffer , " Holographic Associative Memory with Nonlinearities in the Correlation Domain", Appl. Opt. 26, pp1900-1910 (1987).

13. J.M.Wozencraft, I.R.Jacobs, Principles of Communication Engineering, Wiley, N.Y., 1965.

14. H.L.Van Trees, Detection, Estimation and Modulation Theory, Part I, Wiley, N.Y., 1968.

15. J.Goodman, Introduction to Fourier Optics, Mc-Graw Hill, N.Y. 1968.