Feature based Contraction of Sparse Holographic Associative Memory[1]

Javed I. Khan and D. Y. Y. Yun

University of Hawaii at Manoa
Department of Electrical Engineering
Laboratories of Intelligent and Parallel Systems
491 Holmes Hall, 2450 Dole Street, Honolulu, HI-96822
javed@wiliki.eng.hawaii.edu

Abstract

The paper presents a scheme for reducing memory space of a holographic associative memory for content based learning, searching and retrieval of sparse patterns. Holographic associative memory developed on the properties of complex valued Riemann space is one of the most promising models of associative memory. It has demonstrated 10 to 100 times speedup than most other models of associative memories in learning pattern associations with nearly arbitrary level of complexity. The correlation space of the sparse patterns, is also sparse in information, but representationally dense. Therefore, holograph of sparse patterns (such as images) becomes extremely large. In this paper we describe a holographic memory model which projects the sparse holograph on a reduced memory space along all three of its dimensions by unsupervised learning of the stimulus and response patterns. The resulting holographic model also simultaneously increases the encoding, searching and decoding speed.

1 Introduction

Associative computing is expected to play a critical role in the field of intelligent image data-base management. The applications such as content addressable retrieval, query by example, indexing by features; such as color, texture, shape, all these require some form of associative recollection.

Since the advancement of synoptic theory of signal transmission by McCulloch and Pitts (1943), and Hebb (1949) a number of models of artificial associative memories have been developed to mimic the behavior of human brain by researchers such as Marr (1969), Anderson (1989), Willshaw (1971), Kanerva (1988) and many others [Kane88, Will89]. These pioneering models were able to reproduce some of the intriguing behaviors of human brain. Two of the most serious concern with most of these associative memories are their capacity and difficulty in storing arbitrary patterns. However, for image applications, the problem becomes more acute in terms of enormous physical memory requirement. Very few work has been done to make the assosiative memory space efficient.

A typical image pattern generally consists of a large number of pixels, and by nature individually they carry small amount of significant information. As a consequence, the correlation space of the image pattern becomes physically large but sparse in information. The objective of our research is to find ways to extract useful information from this sparseness of the correlation space and to contract the space of associative memory.

Our research is specifically aimed at a model of holographic associative memory proposed by Sutherland [Suth90], which has demonstrated a major break through in speed and capacity. Experiments have revealed speedup of factors 10 to 100 times compared to other paradigms [Souc92,p8]. Multiple pattern associations at nearly arbitrary complexity, without hindered by the linear separability problem, can be enfolded with the holographic memory. Thus, it challenges the principal two limitations of earlier models. The memory operates in complex number domain, in contrast to most other models. The holographic memory itself can be considered as a 3-D volume. The first two dimensions are the stimulus and response pattern elements. The 3rd dimension or the depth represents the bits required to store each of the correlation element. In this paper we present a uniform technique based on the autonomous learning

---

of patterns to contract the holograph along all three dimensions based on the sparseness analysis of the correlation information. The next section presents a background on holographic associative memory. Section 3 presents the contraction model. Finally section 4 provides some performance results from a simulated holograph.

2 Background and Mathematical Basis

2.1 Bi-Modal Representation

A pattern S is a suit of stimulus elements in the form of $\{s_1, s_2, \ldots. s_n\}$. The individual pattern elements are represented as a complex number, where the magnitude of the complex number is representative of the *confidence* in the element and the phase of the complex number enumerates the *content*. Unlike the conventional representation schemes, our model treats information as a bi-modal (confidence:content) notion. Thus a stimulus and response patterns with respectively n and q elements are represented as:

$$S = \{s_1, s_2, \ldots. s_n\} = \left\{ \lambda_1 e^{i\theta_1}, \lambda_2 e^{i\theta_2}, \cdot\cdot\cdot, \lambda_n e^{i\theta_n} \right\} \qquad R = \{r_1, r_2, \ldots. r_q\} = \left\{ \gamma_1 e^{i\phi_1}, \gamma_2 e^{i\phi_2}, \cdot\cdot\cdot, \gamma_n e^{i\phi_q} \right\}$$

The content of each element (s_j, r_j) is transformed to the phase exponent (θ_j, ϕ_j) through some suitable transformation. A sigmoidal transformation can a generate uniform phase distribution if the stimulus elements are distributed normally. λ_j is assigned a confidence value between 0 and 1.0.

2.2 Holographic Memory Model

Holograph stores a large number of stimulus and response pattern associations in the form of complex correlation matrix. The encoding is performed by super-imposing individual correlations on the same holographic memory substrate. Despite the superimposition, provided with a query stimulus, the holograph can regenerate the closest associated response. An association is encoded as the correlation of the response and the transpose conjugate of the stimulus patterns. (In this paper we will use a superscript T to denote the transpose and the bar to denote transpose).

$$A = R \cdot \overline{S}^T$$

All such (k) associations are enfolded by superimposing them in the holograph.

$$H = \sum_t^k A_t \qquad \ldots(1a)$$

The learning equation (1a) has been improved to encode only that part of a new association that is new, instead of the whole. The component of a new association which is already learned is not encoded. The following equation forms the differential learning algorithm for the holograph which incorporates this modification.

$$H = H + (R - H \cdot S)\overline{S}^T \qquad \ldots(1)$$

The differential learning has demonstrate lower saturation and higher capacity of the holograph [Suth90]. To associatively retrieve a pattern, a query pattern S_Q is convolved with the holograph for target response R_Q.

$$R_Q = \frac{1}{c} \cdot H \cdot S_Q \qquad \text{where,} \quad c = \sum_{j=1}^{n} \lambda_j \qquad \ldots(2)$$

If the query is close to some priori encoded stimulus S_T in the holograph then the target response resembles the corresponding response pattern R_T.

The underlying process can be explained through the recovery of a single response element, through (1a). Let the subscripts i and j refer to the element index and t refers to the association index. According to (1a) and (2), the jth element of the query response:

$$r_{(j,Q)} = \frac{1}{c} \sum_i^n \left[\sum_t^k r_{(j,t)} \overline{S}^T_{(i,t)} \right] S_{(i,Q)}$$

If S_Q is close to some priory encoded stimulus $S_{(t=T)}$, then the above equation can be rewritten as:

$$r_{(j,Q)} = \frac{1}{c}\sum_{i}^{n} r_{(j,t=T)}\bar{s}^{T}_{(i,T)}s_{(i,T)} + \frac{1}{c}\sum_{i}^{n}\left[\sum_{(t\neq T)}^{k} r_{(j,t)}\bar{s}^{T}_{(i,t)}\right]s_{(i,Q)}$$

$$= \frac{1}{c}\sum_{i}^{n} r_{(j,t=T)} \mid \bar{s}^{T}_{(i,T)}s_{(i,T)} \mid + r_{crosstalk}$$

The phase of the first summation term here is exactly equally to the phase of $r_{(j,T)}$. Because, the product $\bar{S}^{T}_{(i,T)}S_{(i,T)}$ is always a scalar quantity. For symmetrically distributed associations, the second summation contributes as a random walk in the two dimensional vector space, The length of this path grows very slowly with the square root of the number of vectors. Thus, the resulting response phase closely resembles the phase of the correct response.

If both the stimulus and response patterns are identical, then holograph acts as a content addressable auto associative memory. More extensive analysis of this holographic process can be found in [Suth90, KhYu94]. Now we will concentrate on the reduction of physical space required by the holograph H, which is the key.

3 Dimension of the Holograph

For stimulus and response patterns respectively with n and q elements, corresponding holograph requires *qn* complex correlations requiring a large number of physical memory locations. Our objective is to exploit the sparseness of information content in the patterns and to reduce the physical memory space of the holograph. Our principal decomposition strategy is given by the following modified form of (2):

$$R = T_R^{-1}[\ddot{H}T_S[S]] \qquad \ldots(3)$$

Where, transformation $T_S[.]$ projects the n-dimensional stimulus pattern is projected on an m-dimensional space, and transformation $T_R^{-1}[.]$ projects the p-dimensional output of the holograph onto q-dimensional response pattern space. The new holograph $\ddot{H}$ has dimension $p \times m$, where both $m < n, p < q$.

3.1 Optimum Transformations and Feature Construction

The optimality criterion for the transformations is the faithful reconstruction of the patterns. Now we will concentrate on the computation of the transformations, which can satisfy minimum mean square error (MMSE) transformation criterion.

For a given sparse pattern space, the dense features are be selected from the feature space of the following format, where each of the feature elements is a product of the original pattern elements, each raised to an exponent.

$$y_j = \prod_{i}^{n}[s_i]^{d(j,i)} \qquad \ldots(4)$$

An infinite number of features can be constructed from this feature space for each unique set of exponents. However, we will choose only m such feature elements. So that the dense pattern will be of the form:

$$Y = \{y_1, y_2, \ldots y_m\} = \left(\alpha_1 e^{i\beta_1}, \alpha_2 e^{i\beta_2}, \ldots \alpha_m e^{i\beta_m}\right\}$$

Where α_j is the confidence of feature j, and β_j is the content of feature j. This feature space is a generalization of the higher order feature space previously proposed by Sutherland to attack the reverse problem of dense pattern [Suth90], where the pattern space is small in comparison to the number of features present in the stimulus pattern. Such situation arises in cases such as decoder problem, x-or problem etc.

The exponents $d(j,i)$, should be chosen to suit the MMSE optimum reconstruction criterion of the information content. For each feature, there should be a set of n such exponents corresponding to each of the pattern elements. There should be m such sets to construct the complete feature set Y. We will denote the exponents in the form of matrix D. Now, we would like to compute D which will satisfy the criterion of optimum linear reconstruction.

Let, the vector Β defines a vector with the exponents (phase components) of Y, and vector Θ defines a vector with the exponents (phase components) of S. Then the following iterative equations provide the autonomous learning equation to compute D.

$$D = D + \mu \cdot (\mathrm{B} \cdot \Theta^T - LT[\mathrm{B} \cdot \mathrm{B}^T] \cdot D \qquad \ldots(5)$$

Where, μ is the learning constant for the encoder which is decreasing with time. Operator LT[] refers to the lower triangularization. The reverse transformation is given by (5), which is symmetric to (4) and uses the transposed form D^T as its transformation matrix.

$$\tilde{s}_i = \prod_j^m y_j d(j,i) \qquad \ldots.(6)$$

The following two theorems explain the mathematical basis of (4), (5) and (6).

Theorem 1: If D is assigned random values at time zero, then with probability 1.0 equation (5) will converge and D and D^T *will approach to a transformation pair between* S and Y *planes. The pair ensures MMSE reconstruction of the phase components (content) of the suit* S.

Proof: Using the bi-modal representation, a feature element can be re-written as:

$$y_j = \prod_i^n (\lambda_i)^{d_{(j,i)}} \cdot \exp\left(\mathbf{i} \sum_i^n d_{(j,i)} \cdot \theta_i\right)$$

$$= \alpha_j \exp(\mathbf{i}\beta_j) \qquad \ldots(7)$$

Thus, the transformation (4) in the phase plane (content) can be written in the matrix form as:

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_m \end{bmatrix} = \begin{bmatrix} d_{(1,1)} & d_{(1,2)} & \cdots & d_{(1,n)} \\ d_{(2,1)} & d_{(2,2)} & \cdots & d_{(2,n)} \\ . & & & \\ . & & & \\ d_{(m,1)} & d_{(m,2)} & \cdots & d_{(m,n)} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ . \\ . \\ \theta_n \end{bmatrix}, \quad \text{or } \mathrm{B} = D\Theta$$

β_j is a linear combination of the individual pattern elements, where $d(j,i)$ are the coefficients. The optimum coefficients for linear reconstruction of the content of S can be found by determining the eigenvectors of the space defined by the auto correlation matrix $Q = E[\Theta\Theta^T]$ of the patterns. Let, the eigenvectors of Q are ordered in the decreasing order of their corresponding eigenvalues. The patterns are spanned in an n dimensional space. Then the subspace spanned by the first m eigenvectors will retain maximum reconstruction information for the suit of *Thata* for a given m. Thus, it will allow reconstruction of the input suit Ss from Ys with minimum mean square error. Sanger [Sang89] has proved for scalar case that the learning algorithm of (5) is doing exactly that. For, bounded magnitude of the elements of D, and decreasing μ, irrespective of the initial values, (7) converges in such a manner that the rows of D converges to the eigenvectors of Q in sorted order. Since, the matrix D is an eigenvector matrix, therefore, the reverse transformation is given by transpose of D.*O*

Theorem 2: The transformation pair specified by (4) and (6), using the transformation matrix D *computed through (5) can also reconstructs the magnitude component (confidence) of the pattern suit* S.

Now, we will proof the reconstruction criterion for the magnitude component of S. The forward and reverse transformations are respectively specified by (4) and (6), which can be re-written in the following forms:

$$|y_k| = \prod_i [\lambda_i]^{d(k,i)} \quad \text{and} \quad \tilde{\lambda}_p = \prod_k |y_k|^{d(k,p)}$$

Where, $\tilde{\lambda}_p$ is the reconstructed pattern element. Expanding the right hand side of the reverse transformation,

$$\tilde{\lambda}_p = \prod_k \left[\prod_i [\lambda_i]^{d(k,i)} \right]^{d(k,p)} = \prod_k \prod_i [\lambda_i]^{d(k,i)d(k,p)}$$

$$= \prod_i \prod_k [\lambda_i]^{d(k,i)d^T(p,k)} = \prod_i [\lambda_i]^{\sum_k d^T(p,k)d(k,i)}$$

Since, rows of the matrix D are the eigenvectors, therefore the product of its transpose with itself is an identity matrix of size n.

$$D^T \cdot D = I_{n \times n}, \text{ thus, } I(p,i) = \sum_k d(k,p)d(k,i) = 1 \textit{ when, } p=i$$

$$= 0 \textit{ otherwise.}$$

Thus,

$$\tilde{\lambda}_p = \prod_i (\lambda_i)^{I(p,i)} = (\lambda_p)^1 \prod_{i \neq p} (\lambda_i)^0 = \lambda_p \qquad \ldots(8)$$

Thus, the above transformation pair also reconstructs the magnitude component of the patterns.*O*

In fact, Sanger's proof can be further generalized for complex valued numbers. The basic assumptions of Oja [Oja83] and Ljung's [Ljun77] theorems (which are the foundation of Sanger's proof) are also valid for complex numbers. Thus, an equation analogous to (5), where B and Θ are replaced by corresponding Y and S vectors, can directly compute the eigenvector matrix.

There are several other algorithms for computing multiple eigenvectors such as by Brockett (89), Karhunen & Oja (82), Kuusela & Oja (82) [Oja83, Sang89]. However, the advantage of (5) is that it also sorts the eigenvectors, and the computation in (5) can be performed using smaller final and intermediate memory.

3.2 Feature Based Associative Memory Model

The following set of equations provide the operational principle of the contracted holograph. First transformation matrices, D_S and D_R are computed from the autonomous learning algorithm given by (5) using the stimulus and response pattern suits. Then (9) is used to enfold the pattern associations onto the contracted holograph $\ddot{H}$. The query into the holograph is performed by (10).

$$\ddot{H} = \ddot{H} + \gamma \cdot T_R(R - T_R^{-1}(\ddot{H}Y))\overline{Y}^T \qquad \ldots(9)$$

$$R = T_R^{-1}[\ddot{H} T_S[S]] \qquad \ldots(10)$$

In (9), Y is computed from (4), using $T_S[\cdot]$. The transformations $T_R[\cdot]$, $T_R^{-1}[\cdot]$, and $T_S[\cdot]$ all are computed from (4) using D_R, D_R^T, and D_S respectively.

3.3 Size of the Contracted Holograph

From the storage perspective, a holograph can be thought as a 3 dimensional volume. Where the first two dimensions respectively represent the stimulus (n) and pattern (q) lengths. The depth dimension represents the number of bits required to store each of the correlation elements. If b bits are used to encode each of the correlation element, then the size of the original holograph is qnb bits.

The proposed contraction method not only contracts the holograph in the first two dimensions, but also it can be used to save space in the 3rd dimension. The eigenvalues for the auto-correlation matrices of stimulus and response patterns can be computed by:

$$D_S E[\Theta\Theta^T] D_S^T = \Delta_S \qquad D_R E[\Phi\Phi^T] D_R^T = \Delta_R \qquad \ldots(11)$$

Where, Θ and Φ are the vectors spanned by the phase components of stimulus and response patterns. Δ are real valued diagonal matrices. E is expectation. Each of the elements in the diagonal represents the eigenvalue corresponding to the eigenvectors arranged in D. The eigenvalues of Q_Θ and Q_Φ correspond to the estimate of the variance of the patterns along the dimensions spanned by their corresponding first m eigenvectors. Each of the feature contents, therefore need not to be encoded with same number of bits. Each of them can be quantized according to their variance along corresponding dimensions. One way of assignment is to allot bits proportional to the log of variance. In the holograph, each of the correlation terms requires bits equal to the sum of the bits required by the component product terms. Therefore, if $e_1, e_2, \ldots e_m$ are the bits assigned to the stimulus feature elements, and $f_1, f_2, \ldots f_p$ are the bits assigned to the response feature elements, then the bits to quantize the entire holograph is:

$$= p(e_1 + e_2 + \ldots + e_m) + m(f_1 + f_2 + \ldots . + f_p) = a \cdot \{p \cdot \log(trace(\Delta_S)) + m \cdot \log(trace(\Delta_R))\} \text{ bits}$$

Where a is some proportionality constant. Physically, as a result of this contraction along the depth $\ddot{H}$ will resemble a rectangular box with four trapezoidal sides and tapered bottom. A Typical holograph for storing 64x64 frame images, requires about 128 megabytes. While the contraction reduces the size to 1.5 megabytes.

3.4 Reduction in Computation

Search and Retrieval: The holograph query is performed by (10). If we consider, right parenthetical computation in the order as shown in (12):

$$\overset{(qx1)}{R}_{target} = \left[\overset{(qxp)}{T_R^T} \left[\overset{(pxm)}{\ddot{H}} \left[\overset{(mxn)}{T_S} \left[\overset{(nx1)}{S}_{query} \right] \right] \right] \right] \qquad \ldots(12)$$

The dimensions are shown over each element. Then, evaluation of (12) can be performed by $(pq + mn + 2pm)$ multiplications, $(pq + mn + 2pm)$ additions, and $(pq + mn)$ logarithm evaluation. For, a typical image pattern with 256x256 pixels, $p = m$, $q = n$, and a ratio $n/m = 8$, it means approximately 2^{31} multiplications and 2^{30} logarithms. On the other hand, the regular uncontracted holograph requires $4qn$ multiplications and $2qn$ additions. For the same typical case it means 2^{34} multiplications and half as much additions.

Holographic Learning: Similar saving can be attained in the case of encoding too. If we consider a single step evaluation of equation-7 in the parenthetical order shown below:

$$\overset{(pxm)}{\ddot{H}} = \left[\overset{(pxm)}{\ddot{H}} + \gamma \cdot \left[\left[\left[\overset{(pxq)}{T}_R \left[\overset{(qx1)}{R} - \left[\overset{(qxp)}{T_R^{-1}} \left[\overset{(pxm)}{\ddot{H}} \overset{(mx1)}{Y} \right] \right] \right] \right] \overset{(1xm)}{\overline{Y}^T} \right] \right] \qquad \ldots(13)$$

A single step iteration including the stimulus pattern transformation requires $2(mn + 2pq + 4pm)$ multiplications, $(mn + 2pq + 5pm)$ additions and $(mn + 2pq)$ logarithms. In contrast, the uncontracted holograph requires $8qn$ multiplications and $(6qn + 2q)$ additions. For the typical case, the contracted case means approximately 3.5×2^{30} multiplications and additions, and 3×2^{29} logarithms. The uncontracted holograph requires 2^{35} multiplications and 3×2^{33} additions.

As shown above, despite the additional input and output processing, the contracted holograph model requires less overall computation both, in the enfolding and query process. The saving is approximately given by the n/m ratio. The computations can be performed with high degree of parallelism with pipelined stages. The matrix nature of the computations makes the holographic assosiative model highly scalable of conventional high performance parallel machines. The parallel execution time also reduces by n/m ratio.

4 Experiment

The feasibility of the contracted holograph has been demonstrate by implementing a contracted holographic associative memory. Fig-1(a) shows the original image that was stored in the holograph. Fig-1(b) shows the retrieved image from the holograph for 100% frame query. Fig-1(c) shows the retrieved pattern using query with 50% of the full frame. Unlike most other associative memories, holograph has the unique capability to focus at any arbitrary region of the query frame, with small degeneration of the retrieved pattern. Fig-2(a),(b) and (c) shows similar frames for another image stored in the same holograph. The 50% frame window is shown in Fig-5. The excitation of the holograph can be seen at the distribution presented in Fig-4, which has 32 images encoded in it. The contraction mask set is shown in Fig-3.

The following table provides the signal to noise (SNR) characteristics among these images. The loss only due to transformation is characterized by the comparison between the original image and retransformed image (18.65 db). The loss due to holographic encoding is shown by the comparison between the retransformed and 100% frame based retrieved images (31.37 db). The loss due to partial frame query is expressed by the comparison of 100% frame with 75% and 25% frame based recovered images (20.39 db and 16.62 db respectively). The table provides a comparative picture of the trade-off, for which our methods provides a mechanism. The result of this table corresponds to 64 times contraction in holograph size, and almost 8 times faster encoding and search speed.

SNR Characteristics

SNR between:	Retranformed Image	Retrieved Image (100% Frame)	Retrieved Image (75% Frame)	Retrieved Image 50% Frame
Original	18.65 db	18.43 db	16.47 db	14.54 db
Retranformed	x	31.37 db	20.29 db	16.52 db
(100% Frame)	x	x	20.39 db	16.62 db

5 Conclusion

The proposed work provides a formal mechanism to perform trade-off between size of space and quality of space for holographic associative memory. As for other search problems, the reduction of search space also simultaneously reduces the search speed. The more is the sparseness of the pattern information, the less is the loss of quality in storage due to this contraction and vice verse. Sutherland has previously proposed a method to increase the feature space for dense patterns [Suth90], however without any optimality consideration.

One of the principal significance of this work is to develop an effective means to construct reversible optimum features for complex valued patterns. We have shown the process of constructing compressed holographs based on autonomous learning of features from the pattern space. The features are not only optimum in MMSE reconstruction sense, but also warrants classification. Because, the m feature dimensions are selected in order of pattern variance along each of them.

6 References

[Kane88] Kanerva, Pentti., Sparse Distributed Memory, MIT Press, Cambridge, 1988.

[KhYu94] Javed I. Khan & D. Y. Y. Yun, Chaotic Vectors and A Proposal for Multi-Dimensional Associative Network, Proceedings of the IS&T/SPIE 1994 International Symposium of Electronic Imaging, (to be published), San Jose, February 1994.

[Ljun77] Ljung, L, Analysis of Recursive Stochastic Algorithms, IEEE Trans. on Automatic Control, AC-22, pp551-575, 1977.

[Oja83] Oja, E., Subspace Methods of Pattern Recognition, John Wiley & Sons, NY, 1983.

[Sang89] Sanger, T. D., Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network, Neural Networks, v2, pp459-473, 1989.

[Souc92] Soucek, B. and The IRIS Group, Fuzzy, Holographic, and Parallel Intelligence, John Wiley & Sons, NY, 1992.

[Suth90] Sutherland, J. G., A Holographic Model of Memory, Learning and Expression, Journal of Neural Systems, v1, pp259-267, 1990.

[Will89] Willshaw, D., Holography, Associative memory and Inductive Generalization, Parallel Models of Associative Memory (updated edition), ed- G. E. Hinton, Lawrence Erlbaum Associates, New Jersey, 1989.