COMPLEX ASSOCIATIVE MEMORY NEURAL NETWORK MODEL FOR INVARIANT PATTERN RECOGNITION

Abdul Ahad S. Awwal and Farid Ahmed
Wright State University
Computer Science & Engineering Department
Dayton, OH 45435.

Abstract

A complex associative memory neural network ($\text{CAM}N^2$) model is proposed for the recognition of handwritten characters. The input and the stored patterns here are derived from the complex valued representation of the boundary of the characters. The stored vector representation is formulated based on 1-D representation of an optical pattern recognition filter. Retrieval of stored patterns from a noisy and shifted input is accomplished by using the correlation in the inverse fourier domain. An adaptive thresholding scheme is then applied to obtain a 1-step convergence. The number of convergence of patterns, usually measured as the storage capacity of the associative memory is found to increase significantly. But the major advantage obtained from the complex representation is that the recognition of patterns is invariant to translation, rotation and scaling of the input patterns.

1 Introduction

The internal storage of inner product associative memory has traditionally been represented by real discrete values [1, 3, 2, 4]. In a recent work [5], Awwal *et al.* has proposed a novel complex inner product associative memory (CIPAM) neural network model which shows that the capacity and flexibility of associative memory can be improved by representing the stored and input patterns in complex form. They used fourier descriptors as the representation for the input. Use of fourier descriptors for the invariant pattern recognition proved to be very useful [6, 7, 8], especially when the patterns are correctly identifiable by the boundary or edges. The present work is an extension of [5] that combines the efficient representation and recognition of CIPAM with the invariant property of fourier transform. Next section discusses the proposed model. Section 3 shows the simulation results, which is then followed by discussion and conclusion in section 4.

2 The Model

2.1 Representation of Input Patterns

Chain codes are used to represent the boundary information of an object by a connected sequence of line segments of given length and direction. In our model we use 8-directional chain code as shown in figure 1.

The chain coded edge image is then converted to boundary pixel data. For the 8-directional chain-coded image the consecutive pixel locations on the edges are not equidistant. For example, two neighboring pixels on the diagonal is at distance of $\sqrt{2}$ whereas those that are apart horizontally or vertically are one pixels apart. Often images of varying sizes are acquired resulting in chain code sequences of varying length. As a result the number of samples that are obtained are not equal for different sized objects. However, to make the processing faster usually a 1-D FFT algorithm is applied which requires that the sequence be either a power of two or some other number. Thus the sequence may have

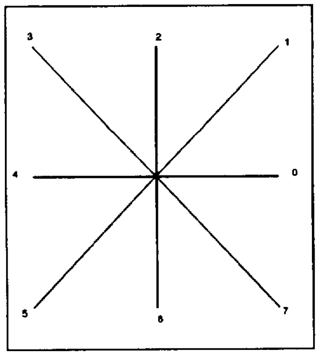

Figure 1: 8-Directional Chain code

to be resampled to make it a power of two. Wallace and Mitchell [7] have proposed sampling the contour with equal interval. This may result in fractions not representing an exact point.

In the current approach, we choose two options, depending on the size of the contour. When the size of the contour is small, we take an equal interval sample; when the size of the contour is more than the required number (N), then we propose an error diffusion technique which selects discrete points closest to the equi-interval points. The total error is minimized by the error diffusion technique. Thus a set of discrete points on the edge image of the pattern is obtained.

2.2 Getting the Fourier Descriptors

Fourier descriptor is defined as the 1-D Fourier transform of a complex sequence formed by the boundary pixels of an object outline or shape taken in a consecutive fashion. Let $f(k)$ be the complex sequence representing the boundary pixels of an edge image, then the Fourier descriptor is expressed as the following summation,

$$F_n = \sum_{k=0}^{N-1} f(k)e^{-(j2\pi/N)nk} \qquad (1)$$

where, $n = 0, 1, ..., N-1$. For the Fourier descriptors to work in our model, a zero padding of the boundary data is necessary before taking FFT, which avoids a circular convolution.

The line segments in the chain coded characters of our storage matrix are not fixed, the average size being 48. We then took 128 point FFT on the stored border pixels data after padding it with zeros. Wallace et al. [7] have provided a normalization procedure which scales, rotates the edge data and then fixes the same starting point so that when the input is compared with the template they start at the same reference point. Next the comparison is made using a mean square error method. Since in their approach mean square error is used, slight variation of scale, rotation or starting point will introduce large comparison error. In our case, we propose to compare the input with the stored templates using a newly proposed associative memory.

2.3 The Storage and Matching prescription

The complex associative memory model which is an extension of our previously proposed trinary associative memory [1, 2, 4] model consists of a collection of stored (learned) vectors represented in complex form. The storage prescription is a 1-D variant of our earlier proposed amplitude modulated phase only filter [9, 10]. The input vector V^{in}, of length L is the Fourier descriptor of the edge data of the input character.

The edge data, as obtained from sampling the chain coded image, is zero padded to make a double length sequence. If $F(n)$ be the discrete Fourier transform of this sequence and $exp(j\theta(n))$ be the phase part of $F(n)$, then the kth stored template is calculated as,

$$V_n^{st,k} = \frac{exp(-j\theta(n))}{|F_k(n)| + a} \qquad (2)$$

Here $0 < a$ and $|.|$ denotes a modulus operation. The parameter a is useful in (i) overcoming the indeterminate condition and (ii) reducing noise.

In order to search for a stored vector in the associative memory, the memory is addressed by an input vector, which could either be a complete or a distorted (rotated, scaled or shifted) representation of one of the stored vectors. The input vector, V^{in}, forms M complex multiplications with the M stored

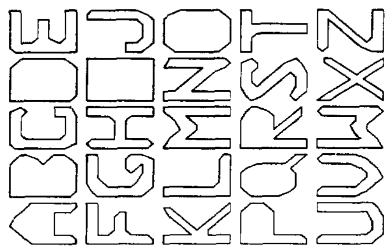

Figure 2: Training Patterns

complex vectors, $V^{st,k}$. These are then normalized. We then take the inverse fourier transforms of these normalized vectors. Depending upon the amount of rotation or shifting of the input pattern the correlation peaks of these output vectors will shift from the first component of the fourier transform. We find the peaks from all of them and choose the maximum. This maximum value is then passed to a thresholding criterion, as illustrated below.

The threshold operation, T_τ, which selects the match or unmatch is given by:

$$\begin{aligned} T_\tau[w] &= 1(\text{match}) \quad \text{if} \quad w - \tau \geq a\tau, \\ &= 0(\text{no match}) \qquad \text{otherwise} \end{aligned} \tag{3}$$

here a is a constant. Note that the above threshold does not simply select the maximum, but selects only if it exceeds the noise margin by certain amount which is again proportional to the noise level.

3 Computer Simulation

Simulation of the model for handwritten character recognition problem is done to illustrate the usefulness of the proposed research. An adaptive thresholding scheme is developed which permits one step

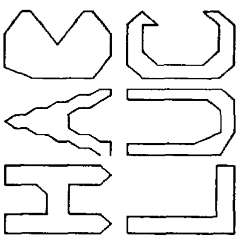

Figure 3: Testing Patterns

matching. The edge images for the stored characters are given in Fig. 2. Some of the test patterns are shown in Fig 3. We used error-diffusion technique to encode the chain coded data. We ran the program with 32 and 64 point FFT. With higher dimension of FFT we got slight improvement.

In the training phase of the simulator, we tuned up the parameters of the model. A known scaling, rotation and shifting of an input resulted in the desired location and magnitude of the correlation. We then tested our model with distorted inputs for both the inner product model and the inverse fourier model. In terms of convergence the inverse fourier model resulted in better results. The following table illustrates this fact, when the test patterns are one code shofted from the training patterns. Note that the inner product model could not recognize 4 such shifted characters, while the proposed model failed to recognize only one character. Of course, both the models succeeded in recognizing these as 2nd best selection, as seen from the 3rd and 5th column of the table. Note that the numbers within the parenthesis in column 1 identifies the corresponding character.

Table 1 : Comparison of Recognition

Input Test Pattern	Converged Pattern			
	Inner Product Model		CAMN^2 Model	
	C1	C2	C1	C2
A(0)	0	1	0	15
B(1)	1	3	3	1
C(2)	2	8	2	6
D(3)	3	22	3	1
E(4)	4	6	4	6
F(5)	1	5	5	9
G(6)	6	2	6	2
H(7)	7	23	7	10
I(8)	8	23	8	18
J(9)	9	23	9	15
K(10)	10	16	10	16
L(11)	11	14	11	16
M(12)	1	12	12	0
N(13)	13	16	13	5
O(14)	14	0	14	19
P(15)	15	1	15	9
Q(16)	14	16	16	14
R(17)	17	3	17	3
S(18)	18	14	18	24
T(19)	19	4	19	9
U(20)	20	21	20	21
V(21)	21	20	21	20
W(22)	22	23	22	20
X(23)	23	8	23	10
Z(24)	14	24	24	18

C1 - 1st Selection
C2 - 2nd Selection

4 Conclusion

The CAMN2 neural network scheme with an adaptive thresholding measure is proposed and its application for the real-time hand-written character recognition is presented. The proposed memory model results in a highly parallelizable neural net architecture. Consequently, a possible parallel opto-electronic implementation may be identified which will result in a high speed automatic character recognition. In the proposed scheme there is no need for training, since learning is achieved simply when the new complex vector is added to the associative memory. To increase the effectiveness, topological feature extraction criterion could be coupled with the system to increase the degree of certainity of recognition.

5 References

1. A. A. S. Awwal, M. A. Karim, and H. K. Liu, "Machine parts Recognition Using a Trinary Associative Memory" Opt. Engg., Vol. 28, no. 5, pp. 537-543, 1989.

2. M. A. Karim and A. A. S. Awwal, "Convergence Mechanism of Associative Memory" Proc. NAECON Conf, Dayton, Ohio, pp. 698-704, 1989.

3. H. K. Liu, S. Y. Kung, and J. A. Davis, "Real-Time optical Associative retrieval technique," Opt. Engg. Vol. 25, no. 7, 853-856, 1986.

4. F. Ahmed, A. A. S. Awwal, and P. Chen, "Character Recognition Using a Trinary Inner Product Associative Memory," IEEE 1992 NAECON Conference Proc., pp. 905-909, Dayton, Ohio, May 1992.

5. A. A. S. Awwal and G. Power, " Three Dimensional Pattern recognition using an opto-electronic inner product complex neural network," to be presented in SPIE meeting on Optical Pattern Recognition, vol. 15, Apr 1993.

6. Timothy P. Wallace and Paul A. Wintz,"An efficient Three-Dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors," Computer Graphics and Image Processing, 13, 99-126, 1980.

7. T. P. Wallace and O. R. Mitchell, "Analysis of Three-Dimensional Movement Using Fourier Descriptors" IEEE PAMI, pp. 583-588, 1980.

8. M. Shridhar and A. Badreldin, "High Accuracy Character Recognition Algorithm Using Fourier and Topological Descriptors", Pattern Recognition, Vol. 17, no. 5, pp. 515-524, 1984.

9. M. A. Karim, and A. A. S. Awwal *Optical Computing: An Introduction.* Wiley, New York, 1992.

10. A. A. S. Awwal, M. A. Karim, and S. R. jahan, "Improved correlation discrimination using an amplitude modulated phase only filter," Applied Optics, Vol. 29, pp. 233-236, 1990.

11. J. J. Hopfield, "Neural networks and physical systems with emergent Collective computational abilities," Proc. Natl. Acad. Sci. USA, Vol. 79, no. 8, 2554-2558, 1982.