Locally Connected BSB Neural Networks as Associative Memories Storing Grey-Scale Images

GIOVANNI COSTANTINI, MASSIMO CAROTA, DANIELE CASALI
Department of Electronic Engineering
University of Rome "Tor Vergata"
Via del Politecnico 1, 00133 - Rome
ITALY

Abstract: - In this paper, we introduce an associative memory storing grey scale images. It's based on a suitable translation of the grey scale image into a Gray-coded binary image, stored in a single BSB binary neural network. The particular BSB we are going to exploit has the property of local connectivity. The chosen learning algorithm guarantees asymptotic stability of the stored patterns, low computational cost, and control of the connection weights precision, without multiplication.

Key-Words: - BSB Neural Networks, Local Connectivity, Associative Memories, Gray Scale Images.

1 Introduction

The most renowned application of associative memories [1-3] is that of image recognition in presence of noise or, in other words, the recognition of a noisy image, even if it differs from the original one in any pixel.

An image with n pixels and L grey levels can be represented by $R = \log_2(L)$ bits for each pixel. It can be stored into a binary neural network with nR neurons; however, the number of interconnections, in case of a classical BSB, is very large, i.e. n^2R^2.

A second approach is based on a multilevel activation function with L plateaus [4-5]. The network exhibits stable equilibrium points with multivalued components, corresponding to the different grey levels. The number of neurons is n; the number of interconnections is n^2.

A third approach is based on complex-valued neural networks [6-7]. The neuron state can assume one of L complex values, with unit magnitude and phases regularly spaced in $[0,2\pi]$. Each phase angle corresponds to a different grey level of the image pixel. The number of neurons is n; the number of interconnections is n^2.

In a fourth approach, the image is decomposed into R binary images, stored into R independent *binary* neural networks. The total number of interconnections is n^2R, but each independent network has only n^2 interconnections [8].

The first approach seems to be the most disadvantageous of all. Nevertheless, it will be the one we will follow, with a trivial but deeply effective modification: the connectivity of the network will be local, that is each neuron will be connected with neurons of its neighbourhood only.

The binary patterns to be stored correspond to as many equilibrium points of the recurrent neural network. The network dynamics recovers a stored pattern starting from a noisy version of it and approaching the corresponding equilibrium point. To store a pattern, we need to find the values of the weight matrix, in order to satisfy some design requirements.

In literature, several neural models implementing binary associative memories have been proposed. Among all the Brain-State-in-a-Box (BSB) neural network is frequently used [9-11]. In addition, we will focus on a particular BSB: a locally connected BSB (LBSB), as explained in Sec. 4.

2 Review of the BSB neural model

The BSB neural model is described by the following difference equation:

$$\mathbf{x}(\mathrm{k}+1) = \mathbf{g}[\mathbf{x}(\mathrm{k}) + \alpha\mathbf{W}\,\mathbf{x}(\mathrm{k})] \qquad \mathrm{k} = 0, 1, 2 \ldots.. \quad (1)$$

$\mathbf{x}(\mathrm{k}) = [x_i(\mathrm{k})] \in [-1,+1]^m$, is the state vector at time k, $\mathbf{W} = [w_{ij}] \in \Re^{m\times m}$ is the local weight matrix. g is the piecewise-linear saturation nonlinearity

$$\begin{array}{lll} g(x) = 1 & \text{if} & x \geq 1 \\ g(x) = x & \text{if} & -1 < x < +1 \\ g(x) = -1 & \text{if} & x \leq -1. \end{array} \qquad (2)$$

Property 1. Let $w_{ii} \geq 0$ for $i = 1,\ldots,m$. Then, only the vertices of $[-1,+1]^m$ can be asymptotically stable equilibrium points of system (1). This means that only binary steady-state solutions can be observed.

Property 2. Let $\xi \in \mathrm{B}^m$, $\mathrm{B} = \{-1, +1\}$, a given binary equilibrium point of system (2). ξ is asymptotically stable if

$$\sum_{j=1}^{m} w_{ij}\xi_i\xi_j > 0 \qquad i = 1, \ldots, m. \qquad (3)$$

Given a set of desired binary equilibrium points $\xi^1 \ldots \xi^Q$ (binary images to be stored), we can find the suitable weight values, by solving constraints (3).

Property 3. Let $w_{ii} = 0$, for $i = 1,\ldots,m$. Assume that $\xi \in \mathrm{B}^m$ is an asymptotically stable equilibrium point of system (1). Then, none of the points $\xi' \in \mathrm{B}^m$ at Hamming distance one from ξ is an equilibrium point.

A zero-diagonal weight matrix guarantees the absence of two or more equilibrium points in a neighborhood and leads to large basins of attraction for the stored patterns [11].

Property 4. System (1) is completely stable if the weight matrix **W** is symmetric.

This means that the system evolution, starting at every initial point **x**(0), always converges to an equilibrium point, without limit cycles or chaotic behavior. Lots of simulations showed that system (1), when properly designed, is completely stable, even if Property 4 is not satisfied. So, stability conditions will not be taken into account.

3 Design of associative memories with finite precision

Starting from Properties 2 and 3, the design of a binary associative memory, based on model (1), can be formulated as follows. Find a local connection matrix **W** such that:

- a given set of Q bipolar patterns to be stored $\xi^{(1)}\ldots\xi^{(Q)} \in B^m$ represent as many asymptotically stable equilibrium points of system (1);
- the basins of attraction of desired equilibrium points are as large as possible;
- the number of undesired stable equilibrium points is as small as possible.
- the following set of constraints be satisfied

$$\sum_{\substack{j=1\\ j\neq i}}^{m} w_{ij}\xi_i^{(q)}\xi_j^{(q)} \geq \delta > 0 \quad i = 1,\ldots,n;\ q = 1,\ldots,Q \quad (4)$$

The last point can be formulated as an unconstrained optimization problem, solved by minimizing the following convex cost function [11]

$$\psi(\mathbf{W}) = \frac{1}{2}\sum_{q=1}^{Q}\sum_{i=1}^{m}\left[P(\Delta_i^{(q)})\right]^2 \quad (5)$$

where P(x) = 1, if x < 0; P(x) = 0 if x ≥ 0, and

$$\Delta_i^{(q)} = \sum_{\substack{j=1\\ j\neq i}}^{m} w_{ij}\xi_i^{(q)}\xi_j^{(q)} - \delta \quad (6)$$

Gradient descent applied to (6) gives [10]

$$w_{ij}(t+1) = w_{ij}(t) + \eta\sum_{q=1}^{Q}\xi_i^{(q)}\xi_j^{(q)}P(\Delta_i^{(q)}(t))$$

$$t = 0,1,2,\ldots \quad i,j = 1,\ldots,m \quad i \neq j \quad \eta > 0 \quad (7)$$

If there's a solution, related to a global minimum of the cost function (5), iterative algorithm (7) converges to it [11].
Each addendum in (7) can be +1, -1 or zero. Consequently, the learning algorithm exhibits three main features:

1. Only additions are required for its implementation.
2. Starting from $w_{ij}(0) = 0$, weights assume the values $w_{ij}(t) = \pm\,\eta\, N_{ij}(t)$, where $N_{ij}(t)$ is a positive integer. Hence, all the weights will have finite precision; the required number of bits is $\log_2(N_{max})+1$, where N_{max} is the maximum value of N_{ij}.
3. The algorithm can be implemented on digital hardware without numerical errors, provided that a sufficient number of bits are used. No rounding or truncation is required to represent the weights. A digital implementation of the algorithm is discussed in [12].

Asymptotic convergence of (7) to a solution of (4) is not guaranteed; the iteration can approach a limit cycle in the solution space [12]. In our experiments, the algorithm is stopped when all the terms $\Delta_i^{(q)}$ become non-negative, for every *i*, *l*. If this condition is not met within a given number of iterations, we say that the desired patterns cannot be stored with the stability margin δ. By choosing η small enough, satisfactory performances can be obtained.

4 Locally connected BSB (LBSB) and decomposition of gray scale images

Consider an M×N LBSB, with M×N cells arranged in M rows and N columns. The basic unit of a LBSB is the cell. Our BSB is locally connected, because any cell in a LBSB directly interacts with only its neighbour cells. In Fig. 1 the case of first order neighborhood is shown: a cell and its neighbourhood are highlighted. In the architecture proposed in this paper, we adopt a neighborhood order such that the total number *m* of neighborhood neurons is much less than M×N. The local connection weights of a LBSB can be computed by the iterative algorithm (7).
Each of the *n* pixels of an *L* grey level image can be represented by *R* bits, $b_1,\ldots,b_R$, with $R = \log_2 L$. In this paper, we consider *L* = 16 grey levels (*R* = 4). One image can be decomposed into *R* binary images, *n* pixels each. The binary images can be stored into a binary LBSB associative memory, mapping the bits of each pixel over the cells, as depicted in Fig. 1. Four contiguous are used to store a single pixel. With this particular binary pattern, bits from different pixels can fall in the same neighbourhood.

b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4

Fig. 1. LBSB neighborhood and pixel bit distribution

We reconstruct the original image by recalling a stored binary pattern and applying the inverse mapping to extract the bits of each pixel from the stable state vector.

A zero mean additive Gaussian noise is one of the principal cause of image corruption. This kind of noise can provoke a jump from a quantization level to an adjacent one and, as a consequence, the reversing of several bits. Thus, the Hamming distance between a stored pattern and its noisy version is in each neighbourhood amplified. To circumvent this problem, instead of the usual binary coding, we adopt the reflected binary *Gray coding*: moving from one quantization level to an adjacent one, only one bit can change.
Using the Gray code, zero-mean additive Gaussian noise results in the minimal Hamming distance, in each neighbourhood, between the stored pattern and its noisy version. So, the probability of correct recall is improved.

5 Experimental tests

To show the effectiveness of the proposed method, many experiments have been accomplished. More noticeable is the following. We store two images with 200x200 pixels and L=16 grey levels. One of these images is shown in Fig. 2a). Due to computer memory limitations, we partition each image into 16 parts, 50x50 pixels each. This way we obtain 32 50x50 images that can be stored into a LBSB with 10000 cells.
Then, we tried to recall the stored images starting from a corrupted version of their. Noisy initial states were generated by adding to the stored images a zero mean Gaussian noise with standard deviation equal to 1.7. An example of noisy image is shown in Fig. 2b). In Fig. 2c) we show the same image after mapping in the LBSB. All the 32 images were correctly recalled.

a) **b)**

c)

Fig. 2 a) One stored image
b) Noisy version of image a)
c) Image a) mapped on the LBSB

7 Conclusions

In this paper a LBSB neural network implementation for associative memories has been proposed. This approach is a variation of the one proposed in [8]. The main differences are: the mapping procedure of the images in the BSB and the local connectivity of the network.
The experimental results show the good behaviour of the structure in term of images recall, when the original images are corrupted by gaussian noise.

References:

[1] J.A. Anderson, J.W. Silverstein, S.A. Ritz, R.S. Jones, *Distinctive features, categorical perception and probability learning: some applications of a neural model*, Psychol. Rev., No. 84, pp. 413-451, 1977

[2] Y. Kamp, M. Hasler, *Recursive Neural Networks for Associative Memory*, Wiley, Chichester, NY, 1990

[3] J.M. Zurada, *Introduction to Artificial Neural Systems*, West Publishing Company, 1992

[4] J. Si, A.N. Michel, *Analysis and syntehsis of discrete-time neural networks with multilevel threshold functions*, Proc. ISCAS 1991, pp. 1461-1464

[5] J.M. Zurada, I. Cloete, E. van der Poel, *Generalized Hopfield Networks for Associative Memories with Multi-valued Stable States*, Neurocomputing (13), pp. 135-149, 1996

[6] N.N. Aizenberg, I. N. Aizenberg, *CNN based on multi-valued neuron as a model of associative memory for grey-scale images*, Proc. IEEE Int. Workshop on Cellular Neural Networks and Applications, CNNA92, Munich, Germany, pp. 36-41, 1992

[7] S. Jankowski, A. Lozowski, J.M. Zurada, *Complex-valued multistate neural associative memory*, IEEE Trans. Neural Networks, vol. 7, No. 6, pp. 1491-1496, 1996

[8] G. Costantini, D. Casali, R. Perfetti, *Neural associative memory storing Gray-coded gray-scale images*, IEEE Trans. Neural Networks, vol. 14, no. 3, pp. 703-707, May 2003

[9] S. Hui, S.H. Zak, *Dynamical analysis of the brain-state-in-a-box (BSB) neural models*, IEEE Trans. Neural Networks, No. 3, 1992, pp. 86-100

[10] W.E. Lillo, D.C. Miller, S. Hui, S.H. Zak, *Synthesis of brain-state-in-a-box (BSB) based associative memories*, IEEE Trans. Neural Networks, No. 5, 1994, pp. 730-737

[11] R. Perfetti, *A synthesis procedure for Brain-State-in-a-Box neural networks*, IEEE Trans. On Neural Networks, vol. 6, N. 5, 1995, pp. 1071-1080

[12] R. Perfetti, G. Costantini, *Multiplierless digital learning algorithm for cellular neural networks*, IEEE Trans. on Circuits and Systems – Part I, vol. CAS-48, No. 5, pp. 630-635, 2001.