Capacity of optical associative memory using a terminal attractor model

Xin Lin a, *, Junji Ohtsubo b, Masahiko Mori a

a Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba 305, Japan
b Faculty of Engineering, Shizuoka University, 3-5-1 Johoku, Hamamatsu 432, Japan

Received 26 June 1997; accepted 25 August 1997

Abstract

The capacity of terminal attractor (TA) model associative memory is investigated based on the consistency between the stored pattern and the obtained equilibrium state in statistical thermodynamics. By computer simulations, we give intuitive estimates of the capacity of the TA model associative memory. For feasibility of optical implementation of the TA associative memory, we impose some approximations to the original TA associative memory without losing the essence of the TA model. The capacity of such a modified TA model associative memory is also given by numerical simulation. The results indicate that the absolute capacity of the TA model is greater than 0.35N, which contrasts with the relative capacity of 0.15N or the theoretical absolute capacity of \( N/(4 \ln N) \) for the conventional associative memory. © 1998 Elsevier Science B.V.

Keywords: Memory capacity; Terminal attractors; Optical associative memory

1. Introduction

Optical neural network for associative memory using the Hopfield model [1] has the advantage of simplicity for its network structure. But further investigation reveals [2] that the capacity of the Hopfield model is quite limited because of the number of spurious states and oscillations. Its absolute capacity is smaller than \( N \) and its relative capacity (within 1% error) is 0.15N at most. When the rate of the number of stored patterns to that of neurons (i.e., memory rate) \( r = M/N \) is greater than about 0.15, the network state will converge to an equilibrium state which is very different from the stored patterns, i.e., a spurious memory is recalled, no matter how large the initial overlap between the input and stored patterns may be. Spurious states arise in various forms. The most common spurious states are those stable states that are not originally stored. In terms of phase-space terminology, these are false attractors trapped in the local minima in the energy landscape. Further, Montgomery and Vijaya Kumar [3] pointed out the existence of oscillating states that also affect the storage capacity.

To increase the storage capacity of a neural network associative memory, we must reduce or eliminate spurious states. Zak [4] introduced a type of attractor called terminal attractor (TA), which represents a singular solution of a neural dynamic system by eliminating the spurious states in the associative memory. Based on the concept of the TA, we have presented a TA model optical neural network associative memory for reducing the spurious states and compared it with the conventional Hopfield model [5,6]. The experimental results indicate that the TA model can reduce spurious states in the Hopfield neural network and the recalling capability can be much improved.

The purpose of this paper is to investigate the capacity of the TA model associative memory. We study the conditions which guarantee equilibrium solutions for memorized patterns in the network. Based on the concept of consistency between the stored pattern \( x_j^{(m)} \) and the obtained

* Corresponding author. E-mail: xlin@etl.go.jp.
equilibrium state $x_i$ in statistical thermodynamics, a method of intuitive estimates of the capacity for the TA model associative memory is presented. By computer simulations, we prove that the absolute capacity of the TA model is greater than $0.35N$, which is much greater than the absolute capacity $N/4 \ln N$ of the conventional model and is even larger than the relative capacity $0.15N$. These obtained results indicate that, using the TA model associative memory, the memory capacity of the conventional Hopfield model is greatly improved by introducing the terminal attractor.

2. Dynamics of TA model

First, let us consider a neural network consisting of $N$ neurons with bipolar output values. At time $t$, the output of the $i$th neuron is $x_i(t)$. Then, we assume that a set of $M$ linearly independent vectors $\{x_i^{(m)}\} (m = 1, 2, \ldots, M)$ with $N$ elements is stored as terminal attractors in the network. The dynamics of the TA model for associative memory is given by [4,6]

$$\frac{dx_i(t)}{dt} + x_i(t) = \sum_{j=1}^{N} W_{ij} f\left[ x_j(t) \right]$$

$$+ \sum_{m=1}^{M} \alpha^{(m)} \left\{ f\left[ x_i(t) \right] - x_i^{(m)} \right\}^{1/3}$$

$$\times \exp\left\{ -\beta^{(m)} \left\{ f\left[ x_i(t) \right] - x_i^{(m)} \right\}^2 \right\},$$

(1)

where $N$ and $M$ are the total numbers of neurons and stored patterns in the network, $\alpha^{(m)}$ and $\beta^{(m)}$ are positive control constants, and $W_{ij}$ is a connection weight from the $j$th neuron to the $i$th one. We assume that the Hebb law is used for $W_{ij}$,

$$W_{ij} = \begin{cases} \sum_{m=1}^{M} x_i^{(m)} x_j^{(m)}, & i,j = 1,2,3,\ldots,N, \\ 0, & i = j. \end{cases}$$

(2)
The function $f[.]$ is a neuron nonlinear function and is usually written as

$$f[x_i(t)] = \tanh[x_i(t)].$$

(3)

Here, the threshold function $\tanh[.]$ operates componentwise on vectors. The unique convergence for the expected solution and stability of equilibrium in the TA model have been mathematically proved by Zak [4].

We will assume several assumptions for our numerical simulations. At first, we assume that the neural network dynamics of Eq. (1) is of discrete time type and eliminate the time derivative term, namely, $\dot{x}_i(t) = 0$. The network changes its state at discrete times $t = 0, 1, 2, \ldots$. Next, we suggest that the parameters $\alpha^{(m)}$ and $\beta^{(m)}$ be chosen as $\beta^{(m)} = \alpha^{(m)} = 1$ for all $m$ from empirical basis. Then, Eq. (1) is written as

$$x_i(t) = \frac{1}{N} \sum_{j=1}^{N} W_{ij} f[x_j(t)] - \sum_{m=1}^{M} \left\{ f[x_i(t)] - x_i^{(m)} \right\}^{1/3} \times \exp \left\{ - \left\{ f[x_i(t)] - x_i^{(m)} \right\}^2 \right\}. \quad (4)$$

In spite of these selections of parameter values, it is proved that the stability of the dynamics in the system remains unchanged [5,6].

Optical implementation of the dynamics system described by Eq. (4) is very difficult because it contains a $1/3$ power function. For feasibility of optical implementation, we make further approximations to Eq. (4), since most optical devices used as a spatial light modulator (SLM) are of binary nature. If we assume a unipolar binary number (1,0) for the neuron-state vectors, the factor $1/3$ in the power function of Eq. (4) may be dropped and, instead of the square operation in the exponential function, the absolute value can be used without changing the value of the equation. By using this assumption, Eqs. (3) and (4) are rewritten as follows:

$$x_i(t) = \frac{1}{N} \sum_{j=1}^{N} W_{ij} f[x_j(t)] - \sum_{m=1}^{M} \left\{ f[x_i(t)] - x_i^{(m)} \right\} \times \exp \left\{ - \left\{ f[x_i(t)] - x_i^{(m)} \right\} \right\}, \quad (5)$$

$$f[x_i(t)] = 1[x_i(t)], \quad (6)$$

where $[u] = 1$ when $u > 0$ and $-1$ when $u < 0$. Eqs. (5) and (6) will be called modified TA model hereafter.

The optical neural network architecture for associative memory based on the TA model described by Eqs. (5) and (6) is shown in Fig. 1. The discussion of optical implementation can be found in Refs. [5–7].

3. Simulation method

The recalling process of the associative memory is as follows. We define that the pattern is recalled when the solution reaches an equilibrium state for a given test pattern. If we cannot obtain an equilibrium solution for a test pattern, we deem that recalling fails. The limitation of the capacity of the Hopfield associative memory is partially caused by the existence of spurious states in its recalling process. Hopfield introduced the concept of the energy function being analogous to spin-glass, and showed by computer simulation that the memory capacity of the associative memory is approximately $0.15N$ with a small error (1%), where $N$ is the number of neurons [1].

On the other hand, the TA model neural network is expected to have no spurious states, infinite stability, and large memory capacity. Several methods for the capacity of neural networks have been proposed, for example, probability way [2], vertex angle estimate [8], and so on. However, the memory capacity of the TA model cannot be given theoretically because it is influenced by the various conditions of the network, so that both mathematical analysis and experimental demonstration are difficult. We will give a rather intuitive explanation by a numerical simulation instead. The principle of the simulation method comes from the approximate equation of the mean field in statistical thermodynamics [9].

First, we introduce a parameter $M(T)^{(m)}$,

$$M(T)^{(m)} = \left\langle \frac{1}{N} \sum_{i=1}^{N} x_i x_i^{(m)} \right\rangle, \quad (7)$$

or the consistency between the stored pattern $x_i^{(m)}$ and the obtained equilibrium state $x_i$ (see Fig. 2), where $T$ is a temperature coefficient that is defined in statistical thermodynamics. Here, it corresponds to a fluctuation coefficient. $\left\langle \cdots \right\rangle$ denotes mean average. Conversely, the inconsistency between $x_i^{(m)}$ and $x_i$ is defined by $(1/2)[1-M(T)^{(m)}]$. Namely, when an equilibrium state $x_i$ is perfectly the same as the stored pattern $x_i^{(m)}$, the parameter $M(T)^{(m)}$ becomes unity and the stored pattern $x_i^{(m)}$ is recalled accurately. When $M(T)^{(m)} < 1$, recollection is unsuccessful.

Fig. 2. Concept of consistency between the stored pattern $x_i^{(m)}$ and the obtained equilibrium state.
Fig. 3. Relation between the inconsistency of recalling and the memory rate (1% of the error is permitted).

Fig. 3 shows the relation between the inconsistency \((1/2)[1 - M(T)^{\alpha m}]\) of the recalling and the memory rate \(r = M/N\) at the fluctuation coefficient \(T = 0\) for conventional Hopfield associative memory. We can see that if the memory rate \(r\) is smaller than a critical value \(r_c\), the inconsistency of recalling increases with increasing memory rate. Namely, the equilibrium point \(x_i\) retreats from the stored pattern \(x_i^{(m)}\), but recollection is successful within 1% error. When the memory rate comes rather close to the critical value \(r_c\) (for the Hopfield neural network, \(r_c \approx 0.145\)), the inconsistency is suddenly increased. Then the network state will converge to an equilibrium state which is very different from the stored patterns.

Fig. 4 shows the memory rate versus with the temperature coefficient. From Fig. 4, we can see that the critical value of the memory rate decreases with increasing temperature coefficient. Namely, the memory capacity deteriorates if there exist fluctuations in the networks, i.e., high temperature.

Numerical simulations for the memory capacity have been performed using a \(10 \times 10\) neuron network model based on the above principle. The Hamming distance between the inputs and the stored patterns is chosen to be larger or equal to 5. The memory rate is defined by \(r = M/N\) in the same way as for the Hopfield model. We only investigate the case \(T = 0\). Namely, the fluctuations of the network are dropped in our numerical simulations.

4. Results and discussions

The results of the computer simulations for the memory capacity in the Hopfield and TA (original and modified models) neural network associative memories are shown in Figs. 5–7, respectively. The abscissa is the iteration time of the network and the ordinate is the Hamming distance of the recalled pattern from a stored pattern "Y", i.e., the consistency of recalling. The Hamming distance of an initial imperfect input from the stored pattern "Y" is 5.
Fig. 7. Hamming distance $H_0$ between output patterns and stored patterns versus iteration time $t$ in the modified TA associative memory. Memory rates are 0.15, 0.25, and 0.35.

We tested three memory rates of 0.15, 0.25, and 0.35. Of course we can set a greater value for the memory rate, but the rate is limited by the capacity of the computer used in the present simulations.

The implication of the results shown in Figs. 5–7 are as follows:

1. When the memory rate is 0.15, the Hamming distance between the recalled pattern and the stored pattern will become small with increasing iteration time and the correct pattern is recalled both for the Hopfield and TA models. We can see that the Hopfield model has faster speed of convergence because of the simplicity of its structure.

2. When the memory rates are 0.25 and 0.35, the Hamming distances of the recalled pattern are respectively 3 and 6 in the Hopfield model (see Fig. 5). Namely, the inconsistency of recalling is always much greater than 1%. This means that the relative capacity is smaller than 0.25 in the Hopfield associative memory. On the contrary, for the TA associative memories both in the original and modified models, the consistency of recalling is 100%. These results indicate that the absolute memory rate of the TA model is greater than 0.35.

Figs. 8–10 show some instant states in the evolutions of the neural networks for the Hopfield and the TA (original and modified models) neural network associative memories, respectively. When the memory rate is $r = 0.15$,
the networks correctly converged to "Y" starting from an imperfect version (the Hamming distance is 5) in all models. But, when \( r > 0.15 \), the neural network state of

the Hopfield model converges to an equilibrium state which is very different from the stored patterns, i.e., a spurious memory is recalled as shown in Fig. 8.

Figs. 11 and 12 show the relation between the memory rate and the iteration time for correct recollection by numerical simulation. Fig. 11 gives the results for the original TA model described by Eqs. (4) and (3), while Fig. 12 for those for the modified TA model of Eqs. (5) and (6). We can see that the speed of correct recalling decreases with increasing memory rate and the Hamming distance of the input from the stored pattern, but a stable association characteristic is shown for both models.

5. Conclusion

We have proposed a method of intuitive estimates of the memory capacity for a TA associative memory based on the concept of consistency between the stored pattern \( x_i^{(m)} \) and the obtained equilibrium state \( x_i \) in statistical thermodynamics. We have proved that its absolute capacity is greater than \( 0.35N \) by using a \( 10 \times 10 \) neuron network model. We also have shown that, by the TA model dynamics, the conventional model is greatly improved in terms of recollection ability and memory capacity. For the feasibility of optical implementation of the TA associative memory, we make approximations to the original TA associative memory without losing the essence of the TA model. A numerical simulation for the memory capacity of such a modified TA associative memory has been also performed. The absolute capacity of the modified TA model is also of the order of \( N \) and is greater than the relative capacity of the conventional associative memory.

However, there still remain some important problems such as mathematical analysis and experimental demonstration of the memory capacity as well as the upper bound of the absolute capacity in the TA model. Further study on the dynamics of neural networks should be conducted.

References