Memory capacity of terminal attractor optical associative memory

Xin Lin and Junji Ohtsubo

Faculty of Engineering, Shizuoka University, Japan
3-5-1 Johoku, Hamamatsu, 432 Japan

ABSTRACT

The memory capacity of terminal attractor (TA) model associative memory is investigated based on the consistency between the stored pattern $x_{\text{in}}$ and the obtained equilibrium state $x_e$ in statistical thermodynamics. By the computer simulations, we give intuitive estimates of the memory capacity of the TA model associative memory. For the feasibility of the 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 memory capacity of such a modified TA model associative memory is also given by the numerical simulation. In this simulation, a $10 \times 10$ neuron network model is used and Hamming distances among inputs and the stored patterns are chosen to be equal to 5 or more both in the original and modified TA models. The results indicate that the absolute memory capacity of the TA model is greater than $0.35N$, which contrasts with the relative capacity of $0.15N$ or the theoretical absolute capacity $N/(4\ln N)$ for the conventional associative memory.

Keywords: terminal attractors, memory capacity, associative memory, optical neural network

1. INTRODUCTION

Optical neural network for associative memory using the Hopfield model\(^1\) has the advantage of the simplicity for its structure of the network. But further investigation reveals\(^2\) that the memory capacity of the Hopfield model is quite limited because of the number of spurious states and oscillations. Its absolute capacity of $N/(4\ln N)$ is smaller than $N$ in the order and its relative capacity (with in 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. The 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 et. al.\(^3\) 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 introduced a type of attractor called a terminal attractor (TA),\(^4\) which represents a singular solution of a neural dynamic system by the elimination of the spurious states in the associative memory. Based on the concept of the TA, we have been 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 memory capacity for the TA model associative memory. We study the conditions which guarantee equilibrium solutions for memorized patterns in the network. Based on the concept of the consistency between the stored pattern $x_{\text{in}}$ and the obtained equilibrium state $x_e$ in statistical thermodynamics, a method of intuitive estimates of the memory capacity for the TA model associative memory is presented. By the computer simulations, we prove that the absolute memory 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 larger than even the relative capacity $0.15N$. These obtained results indicate that, by the TA model associative memory, the memory capacity of the conventional Hopfield model is greatly improved by the introduction of the terminal attractor.
2. DYNAMICS OF TA MODEL

2.1 TA model for associative memory

First, let us consider a discrete time type of neural networks 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 \( \{x^{(m)}_i\} (m = 1,2,\ldots,M) \) with \( N \) elements is stored as terminal attractors in the neural network. The neural dynamics of the TA model for associative memory is given by

\[
dx(t)/dt = x(t) - \sum_{m=1}^{M} \alpha^{(m)} \{f[x_i(t)] - x^{(m)}_i\}^3 \times \exp\{\beta^{(m)} \{f[x_i(t)] - x^{(m)}_i\}^2\},
\]

(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’s law is used for \( W_{ij} \).

\[
W_{ij} = \begin{cases} \sum_{m=1}^{M} x^{(m)}_i x^{(m)}_j, & i,j = 1,2,3,K,N, \\ 0, & i = j \end{cases}
\]

(2)

The exponential multipliers are introduced into Eq. (1) in order to localize the effects of the terminal attractors and to provide a Gaussian distribution peaked at \( f[x_i(t)] = x^{(m)}_i \). When \( f[x_i(t)] \) moves away from \( x^{(m)}_i \), the exponential function decays to zero. This term maximizes the influence of the terminal attractor for \( f[x_i(t)] \) that approaches \( x^{(m)}_i \). \( f[\cdot] \) is a output threshold function

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

(3)

That is, the network changes its state at discrete times \( t = 0,1,2,\ldots \) according to Eqs. (1) and (3). Here, the threshold function \( \tanh[\cdot] \) operates componentwise on vectors. The unique convergence for the expected solution and stability of the equilibrium in the TA model have been mathematically proved by Zak.

We will assume several assumptions for our numerical simulations. At first, we assume a steady state solution for Eq. (1) and eliminate the time derivative term, namely, \( \Delta x_i(t) = 0 \). Next, we suggest that the parameter \( \alpha^{(m)} \) and \( \beta^{(m)} \) be chosen as \( \beta^{(m)} = \alpha^{(m)} = 1 \) for all \( m \) from empirical basis. Then, Eq. (1) is written by

\[
x_i(t) = \sum_{j=1}^{N} W_{ij} f[x_j(t)] - \sum_{m=1}^{M} \{f[x_i(t)] - x^{(m)}_i\}^3 \times \exp\{\beta^{(m)} \{f[x_i(t)] - x^{(m)}_i\}^2\},
\]

(4)

Inspire of these selections of parameter values, it is proved that the stability of the dynamics in the system remains unchanged.

2.2 TA model for optical associative memory

Optical implementation of dynamics system described by Eq. (4) is very difficult because it contains a 1/3 power function. For the feasibility of the optical implementation, we make further approximations to Eq. (4), since most of optical devices used as a spatial light modulator (SLM) are binary nature. If we assume unipolar binary number \((1,0)\) for the neuron-state vectors, the factor of 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. (4) and (3) are rewritten as follows:
\[ x_i(t) = \sum_{j=1}^{N} W_{ij} f[x_j(t)] - \sum_{m=1}^{M} \left| f[x_j(t)] - x_j^{(m)} \right| \times \exp\left[-\left| f[x_j(t)] - x_j^{(m)} \right|\right], \tag{5} \]

\[ f[x_j(t)] = \mathbb{1}[x_j(t)], \tag{6} \]

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

---

Fig. 1 Optical TA neural network. P1-P6: polarizers.

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 the optical implement can be found in refs. 5, 6 and 7.

### 3. SIMULATION METHOD OF MEMORY CAPACITY

The recalling process of the associative memory is as follows. We define that the pattern is recalled when the solution reaches to an equilibrium state for a given test pattern. If we cannot obtain an equilibrium solution for a test pattern, we deem that the recalling is failed. 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 with spin-glass, and showed by the computer simulation that the memory capacity of the associative memory is approximately 0.15 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 the large memory capacity. 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 the mathematical analysis and the experimental demonstration are difficult. We will give a rather intuitive explanation by a numerical simulation instead. The principle of the simulation method is from the approximate equation of the mean field in statistical thermodynamics.\(^8\)

---

Fig. 2 Concept of the consistency between the stored pattern \( x_j^{(m)} \) and the obtained equilibrium state \( x_j \).
First, we introduce a parameter $M(T)^{(m)}$

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

(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 correspond to a fluctuation coefficient. $\langle \rangle$ denotes a mean operation. Conversely, the inconsistency between $x_i^{(m)}$ and $x_i$ is defined by $\frac{1}{2}[1 - M(T)^{(m)}]$. Namely, when an equilibrium state $x_i$ is the same as the stored pattern $x_i^{(m)}$ perfectly, the parameter $M(T)^{(m)}$ becomes unity and the stored pattern $x_i^{(m)}$ is recalled accurately. When $M(T)^{(m)} < 1$, the recollection is unsuccessful.

![Graph 3](image3.png)  
**Fig. 3** Relation between the inconsistency of the recalling and the memory rate (1% of the error is permitted).

![Graph 4](image4.png)  
**Fig. 4** Dependence of the memory rate $r(T)$ on the temperature coefficient $T$.

Figure 3 shows the relation between the inconsistency $\frac{1}{2}[1 - M(T)^{(m)}]$ of the recalling and the memory rate $r = \frac{M}{N}$ 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 the recalling increases with increasing the memory rate. Namely, the equilibrium point $x_i$ retreats from the stored pattern $x_i^{(m)}$, but the recollection is successful in area of 1% error. When the memory rate once 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.

Figure 4 shows how the memory rate changes with the temperature coefficient. From Fig. 4, the critical value of the memory rate decreases with increasing the temperature coefficient. Namely, the memory capacity deteriorates if there exists fluctuations in networks, i.e. high temperature.

The numerical simulations for the memory capacity have been performed by using a $10 \times 10$ neuron network model based on the above principle. The Hamming distance between inputs and the stored patterns are chosen to be equal to $5$ or more. The memory rate is defined by $r = \frac{M}{N}$ in the same way as the Hopfield model. We only investigate the case of
$T = 0$. 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, 6, and 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 the recalling. The Hamming distances of an initial imperfect input from stored pattern “Y” is 5. We tested for 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.

**Fig. 5** Hamming distance $H_r$ between output patterns and stored patterns for the variations of the iteration time $t$ in the Hopfield associative memory. Memory rates are 0.15, 0.25, and 0.35.

**Fig. 6** Hamming distance $H_r$ between output patterns and stored patterns for the variations of the iteration time $t$ in the TA associative memory. Memory rates are 0.15, 0.25, and 0.35.
Fig. 7 Hamming distance $H_o$ between output patterns and stored patterns for the variations of the iteration time $t$ in the modified TA associative memory. Memory rates are 0.15, 0.25, and 0.35.

The implication of the results shown in Figs. 5, 6, and 7 are as follows:
1. When memory rate is 0.15, the Hamming distances of the recalled pattern from the stored pattern will become small with increasing the 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 the 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 the 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 the recalling is 100%. These results indicate that the absolute memory rate of the TA model is greater than 0.35.

Fig. 8 Examples of the recalling process in the Hopfield model associative memory. Memory rates are 0.15, 0.25, and 0.35. Hamming distance of the input from the stored pattern “Y” is 5. $t$ is the iteration time.
Fig. 9 Examples of the recalling process in the TA associative memory. Memory rates are 0.15, 0.25, and 0.35. Hamming distance of the input from the stored pattern “Y” is 5. 

Fig. 10 Examples of the recalling process in the modified TA associative memory. Memory rates are 0.15, 0.25, and 0.35. Hamming distance of the input from the stored pattern “Y” is 5.

Figures 8, 9, and 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 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.

Fig. 11 Relation between the memory rate and the iteration time in the TA associative memory. Hamming distances of the input from the stored pattern “Y” are 5, 10, and 12.

Fig. 12 Relation between the memory rate and the iteration time in the modified TA associative memory. Hamming distances of the input from the stored pattern “Y” are 5, 10, and 12.
Figures 11 and 12 show the relation between the memory rate and the iteration time for the correct recollection by the numerical simulation. Fig. 11 is the results for the original TA model described by Eqs. (4) and (3), while Fig. 12 is those for the modified TA model of Eqs. (5) and (6). We can see that the speed of correct recalling decreases with increasing the memory rate and the Hamming distance of the input from the stored pattern, but a stable association characteristic is shown for the 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 the consistency between the stored pattern and the obtained equilibrium state 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 the optical implementation of the TA associative memory, we make approximations to the original TA associative memory without loosing the essence of the TA model. The numerical simulation for the memory capacity of such a modified TA associative memory have been also performed. The absolute capacity of the modified TA model is also of order $N$ and is greater than the relative capacity of the conventional associative memory.

However, there still remains 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 for the dynamics of the neural networks should be conducted.

6. REFERENCES