Biol. Cybern. 92, 139–146 (2005)
DOI 10.1007/s00422-004-0523-1

Biological Cybernetics

The spatiotemporal learning rule and its efficiency in separating spatiotemporal patterns

Minoru Tsukada, Xiaochuan Pan

Department of Information-Communication Engineering/Brain Science Research Center, Research Institute, Tamagawa University, 6-1-1 Tamagawagakuen, Machida, Tokyo, 194, Japan

Received: 1 July 2004 / Accepted: 28 September 2004 / Published online: 4 February 2005

Abstract. The hippocampus plays an important role in the course of establishing long-term memory, i.e., to make short-term memory of spatially and temporally associated input information. In 1996 (Tsukada et al. 1996), the spatiotemporal learning rule was proposed based on differences observed in hippocampal long-term potentiation (LTP) induced by various spatiotemporal pattern stimuli. One essential point of this learning rule is that the change of synaptic weight depends on both spatial coincidence and the temporal summation of input pulses. We applied this rule to a single-layered neural network and compared its ability to separate spatiotemporal patterns with that of other rules, including the Hebbian learning rule and its extended rules. The simulated results showed that the spatiotemporal learning rule had the highest efficiency in discriminating spatiotemporal pattern sequences, while the Hebbian learning rule (including its extended rules) was sensitive to differences in spatial patterns.

1 Introduction

Since Scoville and Milner (1957) first reported the short-term memory storing function of the hippocampal formation, it has been considered to play an important role during the learning stage of establishing long-term memory. Later, long-term potentiation (LTP) found in the hippocampus (Bliss and Lømo 1973) provided physiological evidence that information from learning and memory is stored in the weight space of the hippocampal network by adjusting its connecting weights.

Several neurophysiological studies have reported that many hippocampal neurons respond to specific locations in the environment (place cells) (O'Keefe and Nadel 1978; Dragoi et al. 2003). In these studies, it was suggested that an important function of hippocampal neurons was to create a cognitive map that associated a "place" with objects in its surroundings. Studies of hippocampal activities during a short-term-memory task suggested that hippocampal neurons are also involved in setting the temporal context of external information (Wilson and McNaughton 1993; Muller et al. 1996). Therefore, it is believed that the hippocampus plays an important role in making short-term memory of spatial and temporal input information.

The Hebbian synaptic learning rule requires coactivity of presynaptic and postsynaptic neurons. However, under some conditions, information regarding the postsynaptic action potentials, carried by backpropagating action potentials, can be strongly degraded before it reaches the distal dendritic synapse of the hippocampal CA1 (Spruston et al. 1995; Andreasen and Ross 1995; Callaway and Ross 1995; Stuart et al. 1997; Golding et al. 2001). Yet, recent results (Golding et al. 2002) have shown that LTP can indeed occur at synapses on distal dendrites of hippocampal CA1 pyramidal neurons, even in the absence of a postsynaptic somatic spike.

Based on results observed in hippocampal LTP induced by various spatiotemporal pattern stimuli (Tsukada et al. 1990, 1994), the spatiotemporal learning rule (STLR) was proposed by Tsukada et al. (1996, 1998). The novel point of this learning rule was "cooperative plasticity without a postsynaptic spike" and its temporal summation.

The learning rule incorporated two dynamic processes: fast (10–30 ms) and slow (150–250 ms). The fast process works as a time window to detect a spatial coincidence among various inputs projected to a weight space of the hippocampal CA1 dendrites, while the slow process works as a temporal integrator of a sequence of events.

In a previous paper (Aihara et al. 2000), the decay constant of fast dynamics was identified as 17 ms by parameter fitting to the physiological data of LTP. Cell assemblies were synchronized at this time scale, which matches the period of the hippocampal gamma oscillation, and that of the slow is 169 ms, which corresponds to a theta rhythm.

In this paper, we systematically examine the functional difference between STLR and Hebbian learning rules in a single-layered neural network, computing their ability to differentiate spatiotemporal sequences.

Correspondence to: M. Tsukada
(e-mail: tsukada@eng.tamagawa.ac.jp)

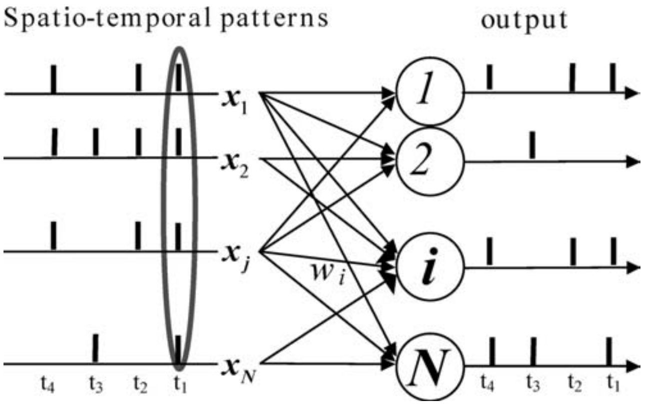


Fig. 1. Structure of the single-layered feedforward neural network and the input-output patterns. The network consists of N neurons; each neuron connects all input nodes (x_1, x_2, x_j, x_N). During the learning period, the input is the spatiotemporal pattern, the spatial snap at one moment corresponds to the spatial frame in the spatiotemporal pattern, and the output is also a spatiotemporal pattern. After learning, a spatial test pattern, which is the last frame of the learned spatiotemporal pattern, is input to the network to generate a spatial output pattern

1	1	0	1	0	0	0	1	1	1	0	0
0	1	0	1	0	1	0	1	1	1	0	0
1	0	0	1	0	0	1	1	1	1	0	0
1	1	0	1	0	0	0	1	0	1	0	1
1	1	0	1	0	0	0	0	1	1	1	0
1	0	0	1	0	1	0	1	1	1	0	0
0	1	0	0	1	0	0	1	1	1	0	1
1	1	0	0	0	1	0	1	0	1	1	0
1	0	0	1	0	0	1	1	0	1	1	0
0	1	0	1	0	1	0	0	1	0	0	1

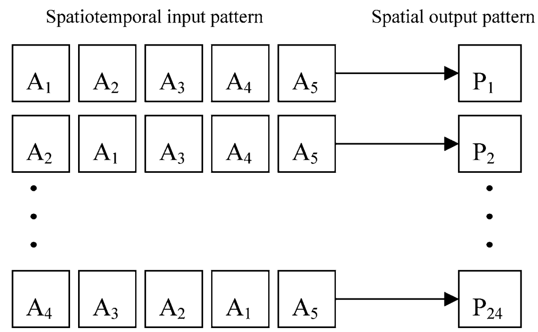


Fig. 2. A sample frame and 24 spatiotemporal patterns applied to the single-layer network. The *upper panel* shows a sample of the spatial frame, i.e., A_1, consisting of N ($N = 120$) elements. The *lower panel* shows 24 spatiotemporal patterns, each pattern having a temporal sequence of 5 frames. The Hamming distance between every two spatial pattern is 8 bits. P_i (i = 1,2,...,24) is the spatial output pattern of the trained network, which is generated by the input of a spatial test pattern to the network

2 The network and learning rules

2.1 The single-layered network

The structure of the neural network is illustrated in Fig. 1. It is a single-layered feedforward network and consists of N neurons. The elements of input patterns are connected to each neuron through a separate weight w_{ij} ($i = 1, 2, \ldots, N,\ j = 1, 2, \ldots, N$). The potential of each neuron depends on both a weighted sum of the simultaneously provided inputs (spatial summation) and the inputs that arrived in the near past (temporal summation).

The above-mentioned functions are expressed in the following equations:
Spatial summation:

$$s_i(t_n) = \sum_{j=1}^{N} w_{ij}(t_n)x_j(t_n)\,; \tag{1}$$

Temporal summation:

$$p_i(t_n) = \sum_{m=0}^{n} s_i(t_m)\exp\left(\frac{-(t_n - t_m)}{\lambda_1}\right)\,; \tag{2}$$

And the output of the neuron:

$$y_i(t_n) = f(p_i(t_n) - \theta_1)\,, \tag{3}$$

where a set of variables, $x_1, x_2, \ldots, x_N$, are inputs to neurons, $x_i(t_n)$ is an input to neuron i at time t_n; $w_{ij}(t_n)$ is the synaptic weight from neuron j to neuron i at time t_n; $p_i(t_n)$ is the potential of neuron i at time t_n, with $y_i(t_n)$ as its output; λ_1 is the time decay constant of temporal summation, which corresponds to the fast dynamic process (λ_1= 10 ms) (Aihara et al. 2000); θ_1 is threshold. The output function of neurons is defined as:

$$f(u) = \begin{cases} 1 & u > 0 \\ 0 & u \le 0 \end{cases}\,. \tag{4}$$

2.2 The learning rules

Learning rules modify network weights to produce output vectors. The most widely used learning rules have evolved from a model developed by Hebb (1949) in which synaptic weight increased if and only if both the source and destination neurons were activated simultaneously. Effective learning rules have been developed by extending this Hebbian learning concept. Recently, we (Tsukada et al. 1996, 1998) proposed a spatiotemporal learning rule, which is unique and more effective in pattern separation than the various Hebbian rules.

2.2.1 Hebb+ learning rule. This is the original Hebbian learning rule that only assumes an increase of connecting weights according to the product of the excitation levels of the source and destination neurons.

In symbols:

$$\Delta w_{ij}(t_n) = w_{ij}(t_{n+1}) - w_{ij}(t_n) = \eta x_j(t_n)y_i(t_n)\,, \tag{5}$$

where $w_{ij}(t_n)$ is the value of a weight from neuron j to neuron i prior to adjustment, $w_{ij}(t_{n+1})$ is the value of weight from neuron j to neuron i after adjustment, η is the learning rate coefficient, $x_j(t_n)$ is the level of excitation of input to neuron j, and $y_i(t_n)$ is the output of neuron i.

2.2.2 Hebb± learning rule. In the Hebb± rule, synaptic modifications of both increasing and decreasing synaptic weight are assumed.

In symbols:

$$\Delta w_{ij}(t_n) = \eta g(x_j(t_n) y_i(t_n)) , \tag{6}$$

where

$$g(x_j(t_n) y_i(t_n)) = \begin{cases} 1 & x_j(t_n) = 1,\ y_i(t_n) = 1 \\ 0 & x_j(t_n) = 0,\ y_i(t_n) = 0 \\ -1 & \text{otherwise} \end{cases} .$$

2.2.3 Extended Hebb (local) learning rule. As we have seen, synaptic weight is modified by the temporal history of input sequences (Tsukada et al. 1996). The representational ability of the network can be improved by introducing the temporal summation of inputs from a single input neuron (local interactions) into the Hebbian learning rule. In this case, the Hebbian equation is modified into the following from:

$$\begin{aligned} &\mathrm{In}_j(t_n) = x_j(t_n) \\ &\mathrm{Out}_i(t_n) = h\left(\sum_{m=0}^{n} w_{ij}(t_m) x_j(t_m) \exp\left(\frac{-(t_n - t_m)}{\lambda_2} \right) - \theta_2 \right) \\ &\Delta w_{ij} = \begin{cases} 0 & \mathrm{In}_j(t) \le 0 \text{ and } \mathrm{Out}_i(t) \le 0 \\ \eta \mathrm{In}_j(t_n) \mathrm{Out}_i(t_n) & \text{otherwise} \end{cases} \\ &h(u) = \frac{2}{\left(1 + \exp(\frac{-u}{\varepsilon})\right)} - 1 , \end{aligned} \tag{7}$$

where $\mathrm{In}_j(t_n)$ is an input to neuron j at time t_n; $\mathrm{Out}_i(t_n)$ is the potentiation force (Aihara et al. 1997), which depends on the temporal summation of inputs to neuron i, through synaptic weight w_{ij}; $h(u)$ is a sigmoid output function of the potentiation force; θ_2 is the thresholds; λ_2 is the time decay constant of temporal summation, which is a slow dynamic process ($\lambda_2 = 223$ ms) (Aihara et al. 2000).

2.2.4 Extended Hebb (global) learning rule. This modification is extended to global interactions between neurons by introducing a spatiotemporal sum. This is expressed by the following equation:

$$\begin{aligned} &\mathrm{In}_j(t_n) = x_j(t_n) \\ &\mathrm{Out}_i(t_n) \\ &\quad = h\left(\sum_{m=0}^{n} \sum_{j=1}^{N} w_{ij}(t_m) x_j(t_m) \exp\left(\frac{-(t_n - t_m)}{\lambda_2} \right) - \theta_2 \right) \\ &\Delta w_{ij} = \begin{cases} 0 & \mathrm{In}_j(t) \le 0 \text{ and } \mathrm{Out}_i(t) \le 0 \\ \eta \mathrm{In}_j(t_n) \mathrm{Out}_i(t_n) & \text{otherwise} \end{cases} , \end{aligned} \tag{8}$$

where $\mathrm{Out}_i(t_n)$ is the spatiotemporal summation of inputs to neuron i and the other parameters are the same as described previously.

2.2.5 The spatiotemporal learning rule. To make the rule of synaptic modification more sensitive to the spatial correlation of input, we introduced the spatial coincidence factor based on the assumption that the change in quantity

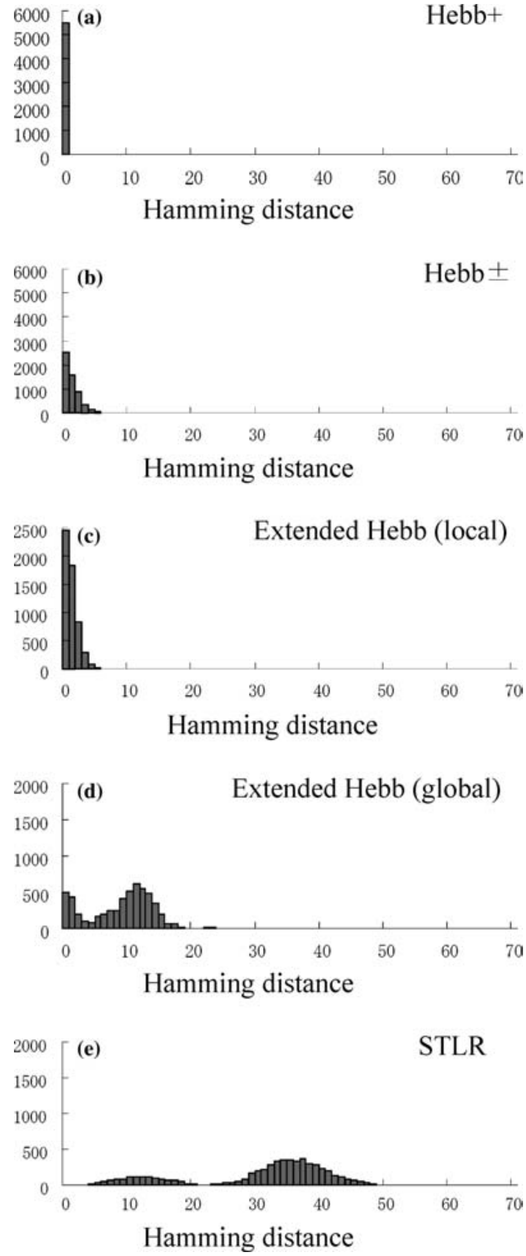


Fig. 3. Distributions of Hamming distances of output patterns for five learning rules. **a** Hebb+. **b** Hebb±. **c** Extended Hebb (local). **d** Extended Hebb (global). **e** Spatiotemporal learning rule

of glutamate binding depends on the coincidence between inputs from a fiber and its surroundings. Combining the above-mentioned spatial and temporal summations, the spatiotemporal learning rule (STLR) can be expressed by the following equation:

$$\begin{aligned} &\begin{cases} I_{ij}(t_n) & = w_{ij}(t_n) x_j(t_n) \sum_{\substack{k \ne j \\ k=1}}^{N} w_{ik}(t_n) x_k(t_n) \\ I_{\mathrm{add}}(t_n) & = \sum_{m=0}^{n} I_{ij}(t_m) \exp\left(-\frac{(t_n - t_m)}{\lambda_2} \right) \end{cases} \\ &\Delta w_{ij}(t_n) = \begin{cases} \eta h(I_{\mathrm{add}}(t_n) - \theta_2) & I_{\mathrm{add}} > \theta_2 \\ 0 & \theta_2 \ge I_{\mathrm{add}} \ge \theta_3 \\ \eta h(I_{\mathrm{add}}(t_n) - \theta_3) & I_{\mathrm{add}} < \theta_3 \end{cases} , \end{aligned} \tag{9}$$

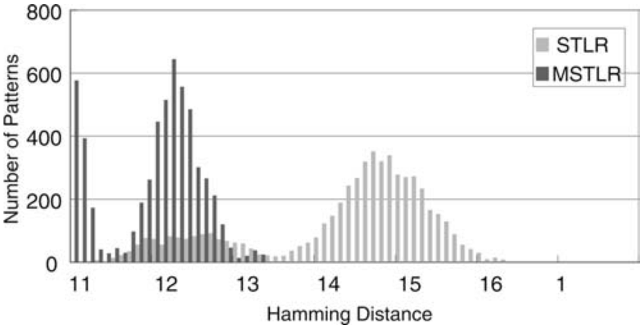


Fig. 4. Distributions of Hamming distances of output patterns for two learning rules. The *gray columns* represent the distribution for the spatiotemporal learning rule (STLR) with the spatial coincidence, while the *black columns* are the distribution for the modified spatiotemporal learning rule (MSTLR) without coincidence. The STLR produces the bimodal distribution on Hamming distance, but the MSTLR only one-modal distribution. The main cluster of the MSTLR corresponds to the cluster with short HD in STLR

where $I_{ij}(t_n)$ is the value of spatial coincidence from neuron j to neuron i and $I_{\text{add}}(t_n)$ is the temporal summation of $I_{ij}(t_n)$.

To demonstrate the significance of spatial coincidence, we introduced a learning rule similar to STLR but that does not incorporate a spatial coincidence factor. This rule is expressed as follows:

$$\begin{cases} I_{ij}(t_n) & = w_{ij}(t_n)x_j(t_n) + \sum\limits_{\substack{k \neq j \\ k=1}}^{N} w_{ik}(t_n)x_k(t_n) \\ I_{\text{add}}(t_n) & = \sum\limits_{m=0}^{n} I_{ij}(t_m)\exp\left(-\frac{(t_m - t_n)}{\lambda_2}\right) \end{cases} \tag{10}$$

$$\Delta w_{ij}(t_n) = \begin{cases} \eta h(I_{\text{add}}(t_n) - \theta_2) & I_{\text{add}} > \theta_2 \\ 0 & \theta_2 \geq I_{\text{add}} \geq \theta_3 \\ \eta h(I_{\text{add}}(t_n) - \theta_3) & I_{\text{add}} < \theta_3 \end{cases}.$$

The two learning rules given by (9) and (10) look very similar. The only difference is that the input from channel j multiplies the sum of inputs from the other channels in (9), while in (10) the two values are simply summed together. The latter rule is called the modified spatiotemporal learning rule (MSTLR).

2.3 The spatiotemporal pattern

The spatiotemporal pattern used in this simulation consists of five frames of spatial patterns (Fig. 2), i.e., A_1, A_2, A_3, A_4, A_5 (A_i is a spatial frame).

Every frame consists of N elements ($N = 120$), and each element is chosen as "1" or "0" randomly, but the total number of "1"s activity is maintained throughout the various spatial patterns (in this simulation, half of the elements in one spatial frame are "1" and half are "0"). The Hamming distance (HD) between every two spatial patterns is 8 bits (if not specified in the simulation). In some cases it is 2 or 24 bits (mentioned). Calculating all of the permutations of four spatial patterns, 24 spatiotemporal patterns were grouped as a training set. The last frame of each spatiotemporal pattern is the same (A_5). During the learning process, the 24 spatiotemporal patterns in the training set were learned by each neural network under the same initial conditions. The spatiotemporal pattern is mapped onto the synaptic weight space of a single neural network. This stored information in the weight space is read by a test pattern (which is given by the last frame of the learned spatiotemporal pattern). For each learning rule, the threshold of neurons, θ_1, is set so that about half of the elements in the output pattern are "1." We compared HDs between output patterns for each learning rule. The averaged HD is often adopted to compare the ability of discriminating spatiotemporal patterns, which is defined as:

$$\text{averaged HD} = \frac{\sum(\text{number of pairs} \times \text{HD of this pair})}{\sum \text{number of pairs}}. \tag{11}$$

3 The simulated results

3.1 Spatiotemporal pattern separation

Five learning algorithms were used to train each of 24 spatiotemporal input patterns in single-layer neural network models. Each of the neural networks had the same initial condition. The differentiation of output patterns represented in learned networks was analyzed by their Hamming distances (Fig. 3). Hebb+ produced the same output pattern, with a Hamming distance of zero, for all of the different spatiotemporal input patterns (Fig. 3a). This proves that the Hebbian learning rule cannot discriminate different spatiotemporal input patterns. Hebb(±) and extended Hebb (local) showed a slight improvement in their pattern separation ability (Fig. 3b, c). Extended Hebb (global) produced a wider histogram with a high peak at 12 Hamming distance (Fig. 3d). These results indicate that the global interaction contributes to an improvement in pattern separation. Finally, the network trained by the spatiotemporal learning rule produced the widest bimodal distribution of Hamming distance (Fig. 3e), which shows that it has the highest efficiency in pattern separation.

The two factors responsible for the high efficiency in pattern separation are spatial coincidence and temporal summation. The network trained by the learning rule without spatial coincidence from (10) produced the one-model distribution, corresponding to the histogram produced by extended Hebb (global), shown in Fig. 4. From this fact we can conclude that the distribution in the longer range of bimodal distribution (Fig. 3e) in the histogram is generated by the spatial coincidence factor in (9), while the distribution in the short range is generated by the spatiotemporal summation. Thus, the ability to separate patterns in the network can be improved by introducing two factors: spatiotemporal summation and spatial coincidence, but the latter is more important.

3.2 Spatiotemporal pattern convergence

From the results shown in Fig. 4 it can be noted that the output patterns of the network trained by the spatiotemporal learning rule were separated into two groups by

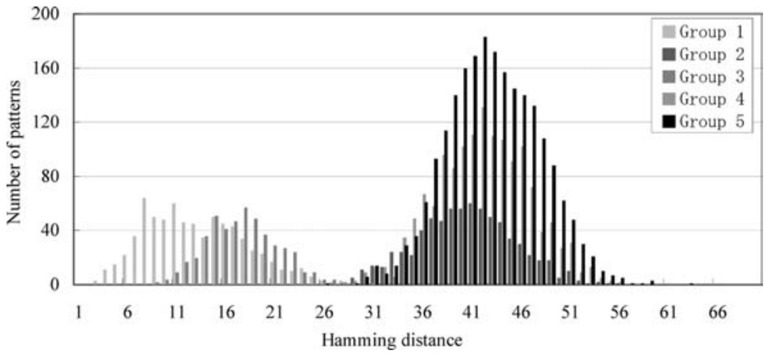


Fig. 5. Separated distributions of output patterns on the Hamming distance under the spatiotemporal learning rule (STLR). The Hamming distance between spatial patterns is 8 bits. The spatiotemporal input patterns that are classified into group 1 and group 2 have 16 bits of HD. Those in group 3 and group 4 have 24 bits of HD; and patterns in group 5 have 32 bits of HD. The first frames of the input patterns in group 1 and group 3 are the same, while the first frames of the input patterns in group 2, group 4, and group 5 are different. The spatiotemporal patterns in group 1 and group 3 are mapped into the cluster with short HD, and those in group 2, group 4, and group 5 are mapped into the second cluster with long HD

Hamming distance. One is the distribution in the short range of HD, and the other is that in the long range of HD. We classified the trained patterns according to Hamming distance and the first frame, i.e., A_1 (shown below). Given a random spatiotemporal pattern, i.e., SP1 ($A_1A_2A_3A_4A_5$), other spatiotemporal patterns can be divided into five groups according to the given pattern SP1. The first group (group 1) includes those input patterns that have 16 bits HD from pattern SP1 and whose first frames are the same as that in SP1. The input patterns in group 2 have the same HD as in group 1 but a different spatial pattern than the SP1 in the first frame. Groups 3 and 4 are divided in the same way as groups 1 and 2, but the patterns in these two groups have a 24-bit HD with pattern SP1. Group 5 patterns have a 32-bit HD between pattern SP1, and, in this case, their first frames must be different. Here are some examples of each group. If the given spatiotemporal pattern in the training set is SP1 – $A_1A_2A_3A_4A_5$ – then:

Group 1 includes patterns such as:

$A_1A_3A_2A_4A_5$, $A_1A_2A_4A_3A_5$; ...

Group 2 includes patterns:

$A_2A_1A_3A_4A_5$, $A_3A_2A_1A_4A_5$; ...

Group 3 includes patterns

$A_1A_3A_4A_2A_5$, $A_1A_4A_2A_3A_5$; ...

Group 4 includes patterns:

$A_2A_3A_1A_4A_5$, $A_3A_1A_2A_4A_5$; ...

Group 5 includes patterns:

$A_2A_3A_4A_1A_5$, $A_3A_1A_4A_2A_5$. ...

Figure 5 demonstrates the HD distributions of the outputs for these five groups of input patterns after being learned by the STLR. Comparing Fig. 5 (component histogram) and Fig. 4 (whole histogram), it is very clear that patterns in groups 1 and 3 are mapped into a cluster with short HD, and those in groups 2, 4, and 5 are mapped into a second cluster with long HD. The similar property of patterns in groups 1 and 3 is that their first spatial frames are the same, while the first frames of patterns in groups 2, 4, and 5 are different. Although the HDs in the latter three groups are very different (16, 24, and 32, respectively), their distributions of output patterns are very similar and are all mapped into one cluster. This suggests that the first frame in the spatiotemporal pattern is more effective than the other frames in contributing to the STLRs ability to separate input. It also shows that differences in temporal sequence are more important than differences in spatial pattern in separating input patterns. Other learning rules have not been able to produce any such clear difference in the mapping of input and output pattern spaces.

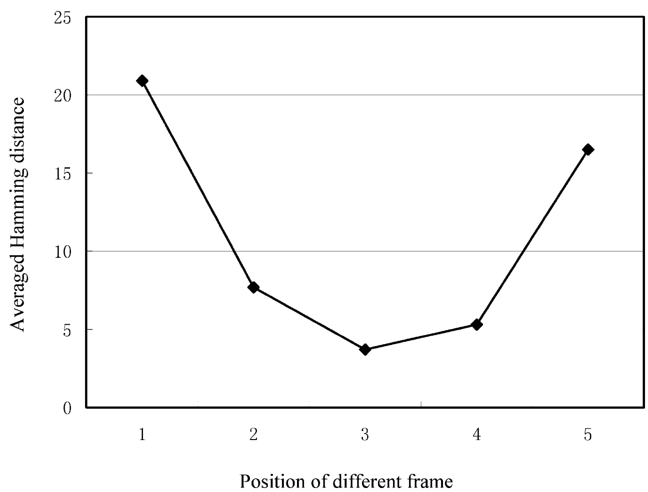


Fig. 6. The effect of the position of different spatial patterns on averaged HD. The network learned the spatiotemporal patterns using the spatiotemporal learning rule (STLR). The spatiotemporal input patterns consist of five spatial frames. There are five input pattern sets, with each set having 24 spatiotemporal patterns. In the first set, the first frames of all input pattern are different, while other frames are same, for example, $P_1A_2A_3A_4A_5$, $P_2A_2A_3A_4A_5$, and so on. The other four sets are produced by the same method. After learning, the last frames of the input pattern are input to the network to generate the output pattern. The averaged HD is calculated based on these output patterns

It has been shown that the first frame is essential in differentiating spatiotemporal patterns. Next, to compare

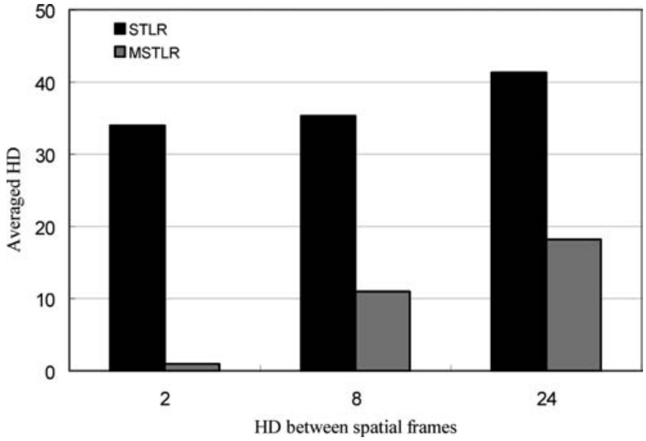


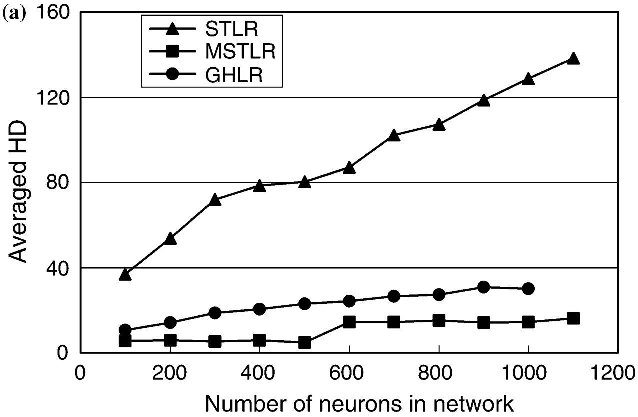
Fig. 7. The comparison of averaged HDs between two learning rules to learn three kinds of spatiotemporal patterns. The HDs between spatial patterns is 2 bits, 8 bits, and 24 bits. The *black columns* represent the spatiotemporal learning rule (STLR), and the *gray columns* are the learning rule without coincidence (MSTLR)

the importance of temporal frames, a new set of spatiotemporal patterns were generated. A spatiotemporal pattern in this set consists of five frames of spatial patterns and the HD between every two spatial patterns is 8 bits. Patterns in the set were divided into five groups. In the first group, the first frame of each of the spatiotemporal patterns is different, while in the other frames they are the same. In the second group, only the second frames are different and other frames are the same. Similarly, we produced patterns labeled group 3, group 4, and group 5 in which only the third, fourth, and fifth frames are different, respectively.

These spatiotemporal patterns in the five groups were applied to the network trained by the STLR. After training, a spatial test pattern was applied to the network to obtain its output. The test pattern was the last spatial frame of the trained pattern. The averaged HD of the five groups is shown in Fig. 6, where it is clear that the averaged HD in group 1 is the largest. Thus, it can be interpreted that the first frame is the most important in separating spatiotemporal patterns. But noting that the averaged HD in group 5 is larger than those in groups 2, 3, and 4, it is possible that the network incorporating the STLR remembers the last and first frames better than the other frames.

It has been shown that temporal sequence is important in separating input patterns. We then investigated the question of how the HD between spatial patterns affects output separation. Two new groups of spatial patterns were introduced. One group had spatial patterns with a HD of 2 bits, and the other group had a HD of 24 bits. We compared outputs for three different input patterns with HDs of 2 bits, 8 bits and 24 bits, under the STLR and MSTLR. These three input patterns gave similar outputs under the STLR, but produced very different outputs when using the MSTLR. The averaged HD of output patterns is illustrated in Fig. 7. When the HD between spatial patterns varied from 2 to 24 bits, the averaged HD increased greatly for the MSTLR but only changed slightly for the STLR. These results indicate that the STLR pays more attention to the difference in temporal sequences than to that of

(a)



(b)

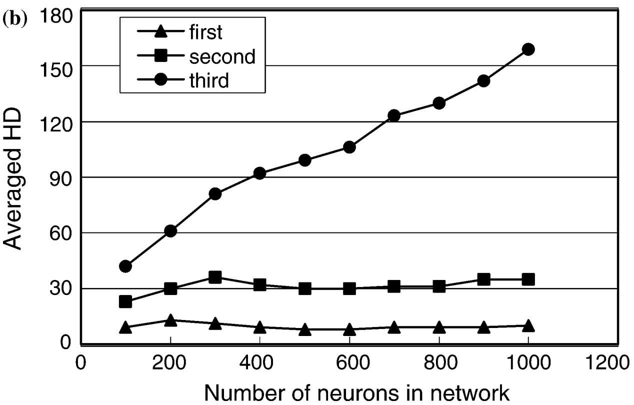


Fig. 8. The relation of averaged HD to the number of neurons in the network. **a** Dependence of the averaged HD on the number of neurons in the network for three learning rules. The *triangle-line* is the spatiotemporal learning rule (STLR), the *square-line* is the learning rule without coincidence (MSTLR), and the *circle-line* represents the global Hebbian learning rule (GHLR). **b** Positions of three peaks with the number of neurons in the network for the STLR. The *triangle-line* is the first peak (nearest to the origin), the *square-line* is the second peak (the middle one), and the *circle-line* is the third peak (furthest to the origin)

spatial patterns. For the MSTLR, on the other hand, the HD of spatial patterns is more important than the difference in temporal sequences for separating spatiotemporal patterns.

3.3 Effect of neuron quantity in the network

During the simulation it was found that the STLR's ability to perform pattern separation increases with the number of neurons in the network, as shown in Fig. 8a. In this figure, the triangle-curve represents the STLR, the square-line represents the MSTLR, and the circle-line is the global Hebbian learning rule (GHLR). During the simulation, the HD between spatial frames is always kept at 8 bits. With the increasing number of neurons in the network, the averaged HD increases significantly for the STLR. For the other two rules, however, their averaged HDs remain unchanged. Furthermore, the peaks of the distributions for GHLR and MSTLR are not affected by the number of neurons. On the other hand, the shape of the STLR is divided into three groups (three peaks): first peak (nearest to the origin), second peak (middle), and third peak (furthest from the origin). When the number of neurons increases, the position of the third peak moves further away, leaving the position of the first and second peaks unchanged, as illustrated in Fig. 8b. If the number of neurons is less than a particular value (about 80), the STLR generates similar results as the GHLR and MSTLR in which the three peaks are merged together. These results indicate that the STLR is sensitive to the scale of the network; the more neurons, the greater the STLR's ability to discriminate spatiotemporal patterns.

4 Conclusion and discussion

Hebbian learning is characterized by coincident pre- and postsynaptic activity; the interconnected weights that contribute to fire a postsynaptic neuron are strengthened according to the delta rule. In the training process, once a neuron is fired, the same neuron tends to fire and the similar weights are strengthened because the succeeding spatial pattern is slightly different (only 8/128 bits) from the previous one.

The Hebb+ learning rule, therefore, cannot separate different spatiotemporal sequences with identical mean rates of firing and shows a tendency to draw these sequences into a fixed spatial pattern (Fig. 3a). Hebb± and local Hebb rules show a slight improvement in their ability to separate patterns (Fig. 3b, c). This improvement observed in Fig. 3b depends on the effect of the decreasing weight and that of Fig. 3c on the temporal summation from a single input neuron. Further improvement seen in Fig. 3d was dependent on the effect of spatiotemporal summation.

On the other hand, when the STLR is applied, the weights are strengthened according to the magnitude of correlation between input spatial pattern X and weight vector W in a neuron, and so the weight change produced by the spatial pattern is distributed among neurons in the layer depending on the magnitude of correlation. Therefore, the distributed weights resulting from the STLR are strongly influenced by initially distributed weight vectors. The network weights are set to initial values before training starts by randomizing the weights to small numbers. The pattern separability corresponds to the second peak shown in Fig. 3e, while the time history with exponential decay constant λ_2 in the STLR corresponds to the first peak in Fig. 3e.

From these simulated results it can be concluded that the STLR has the highest ability to discriminate spatiotemporal patterns. It was also demonstrated that spatial coincidence is essential for this ability of the learning rule. One of the interesting properties of the STLR is that its ability to separate input patterns depends on the number of neurons in the network. Although more neurons can increase its separation ability, if the number of neurons decreases below a certain value, the STLR will lose this ability. It may also be noted that the STLR is more effective than the Hebbian learning rule (including its extended rules) in differentiating between spatiotemporal sequences with identical mean rates. While the Hebbian learning rule is suited for memorizing abstract global context, the STLR is better for memorizing concrete, local context. This suggests that the STLR and the Hebbian learning rule may have different functional roles in hippocampal memory.

Although the STLR considers the coincidence between multiple inputs, it does not consider the coincidence between inputs and outputs. The latter plays an important role in memory formation. In the present study, the STLR was implemented only in a single-layered feedforward neural network to discriminate patterns. The next step will be to propose a more realistic model of the hippocampus consisting of excitatory and inhibitory neurons and feedforward and feedback circuits to test the spatiotemporal learning rule. These results will be compared to experimental data (such as: LTP, LTD) and help approach a closer understanding of the memory-formation processes of the hippocampus.

Acknowledgements. This study was supported by a Grant-in-Aid (12210017) for Scientific Research on Priority Areas Advanced Brain Science Project from Ministry of Education, Culture, Sports, Science and Technology, Japan and by The 21st Century Center of Excellence Program (Integrative Human Science Program, Tamagawa Univ.)

References

Aihara T, Tsukada M, Crair MC, Shinomoto S (1997) Stimulus-dependent induction of long-term potentiation in CA1 area of the hippocampus: experiment and model. Hippocampus 7:416–426

Aihara T, Tsukada M, Matsda H (2000) Two dynamic processes for the induction of long-term potentiation in hippocampal CA1 neurons. Biol Cybern 82:189–195

Andreason M, Lambert JD (1995) Regenerative properties of pyramidal cell dendrites in area CA1 of the rat hippocampus. J Physiol 483:421–441

Bliss TV, Lømo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J Physiol 232:331–356

Callaway JC, Ross WN (1995) Frequency-dependant propagation of sodium action potentials in dendrites of hippocampal CA1 pyramidal neurons. J Neurophysiol 74:1395–1403

Dragoi G, Harris KD, Buzsaki G (2003) Place representation within hippocampal networks is modified by long-term potentiation. Neuron 39:843–853

Golding NL et al (2001) Dichotomy of action-potential backpropagation in CA1 pyramidal neuron dendrites. J Neuropyshiol 86:2998–3010

Golding NL et al (2002) Dendritic spikes as a mechanism for cooperative long-term potentiation. Nature 418:326–331

Hebb DO (1949) The organization of behavior. Wiley, New York

Muller RU, Stead M, Pach J (1996) The hippocampus as a cognitive graph. J Gen Pysiol 107:663–694

O'Keefe J, Nadel L (1978) The hippocampus as a cognitive map. Oxford University Press, Oxford

Spruston N et al (1995) Activity-dependent action potential invasion and calcium influx into hippocampal CA1 dendrites. Science 268:297–300

Stuart G et al (1997) Action potential initiation and backpropagation in neurons of the mammalian CNS. Trends Neurosci 20:125–131

Tsukada M, Aihara T, Mizuro M, Kato H, Ito K (1990) Long-term potentiation on temporal pattern stimuli in hippocampal slices. In: Iwai E, Mishikin M, (eds) Vision, memory, and the temporal lobe. Elsevier, New York pp 251–254

Tsukada M, Aihara T, Mizuro M, Kato H, Ito K (1994) Temporal pattern sensitivity of long-term potentiation in hippocampal CA1 neurons. Biol Cybern 70:495–503

Tsukada M, Aihara T, Saito H, Kato H (1996) Hippocampal LTP depends on spatial and temporal correlation of inputs. Neural Netw 9:1357–1365

Tsukada M, Hagiwara H, Aihara K (1998) A high ability of proposed spatio-temporal learning rule to differentiate spatio-temporal sequences demonstrated in a simple model of short-long memory. International symposium on nonlinear theory and its applications, Crans-Montana, Switzerland, pp 1313–1315

Wilson MA, McNaughton BL (1993) Dynamics of the hippocampal ensemble code for space. Science 261:1055–1058