

Neural Network Model of Context-Dependent Neuronal Activity in Inferotemporal Cortex

Atsuo Suemitsu

Department of Mathematics and Computer Science
Interdisciplinary Faculty of Science and Engineering
Shimane University
Shimane 690-8504, JAPAN
E-mail: sue@cis.shimane-u.ac.jp

Masahiko Morita

Graduate School of Systems and Information Engineering
University of Tsukuba
Ibaraki 305-8573, JAPAN
E-mail: mor@bcl.esys.tsukuba.ac.jp

Abstract—Neuronal activities related to context-dependent recall have been found in the monkey inferotemporal cortex. If we set the same task for an artificial neural network, however, a serious computational difficulty arises. In the present paper, we overcome this difficulty by implementing a novel method of contextual modulation, termed selective desensitization, and construct a neural network model that performs the same context-dependent memory task as that assigned to the monkey. The model, being consistent with the anatomical structure of the inferotemporal lobe, as well as with physiological findings, not only reproduces the empirical data well but also gives a clear account for a phenomenon that had not been explicable to date. This strongly suggests that the brain implements context-dependent recall based on the same principle as adopted in the model.

I. INTRODUCTION

Naya et al. [1] reported neuronal activities related to context-dependent recall in the inferotemporal cortex (IT) of monkeys. They used two kinds of visual memory tasks and found neurons that changed activity when the task was switched. This finding is interesting because it reflects the recall process of the target that depends on the context.

If we try to model such context-dependent recall using artificial neural networks, however, we face a serious computational difficulty caused by many-to-many correspondence between the input patterns and the target patterns. Accordingly, no existing model explains the above neuronal activities sufficiently.

It has recently been shown, however, that the above difficulty can be resolved by implementing a novel method of contextual modulation, termed selective desensitization [2], [3]. In the present study, we construct a neural network model of context-dependent recall in IT on the hypothesis that the selective desensitization method is used in the brain, and compare simulation results with the empirical data.

II. CONTEXT-DEPENDENT ACTIVITY IN INFEROTEMPORAL CORTEX

To investigate memory mechanisms in the primate brain, the delayed matching-to-sample (DMS) and the delayed pair-association (DPA) tasks have often been used. In these tasks, a cue picture is presented for a short period, and the monkey is required to judge whether a test picture presented after a delay

period is the target or not. The two tasks are different in that the target is the same picture as the cue in the former whereas it is the paired associate of the cue in the latter. Combining these, Naya et al. [1] developed a task termed the pair-association with color switch (PACS) task.

In their experiment (Fig. 1), 12 pairs $(G_1, C_1), \dots, (G_{12}, C_{12})$ of computer-generated pictures were prepared, each pair containing a green picture and a cyan picture. One of them was used as a cue, and the target was the same picture in trials for the DMS task or its paired associate in trials for the PACS task. In a DMS trial, delay period 1 (d1) after cue presentation was 5 s, during which a stimulus (color signal) of the same color as the cue was shown on a display. After d1, the color signal changed to gray and delay period 3 (d3) of 1 s began. A test picture was presented at the end of d3. In a PACS trial, d1 lasted only 2 s and delay period 2 (d2) started at the switching of the color signal to a different (the same as the target) color; d2 lasted 3 s and the color signal became gray in d3.

Naya et al. recorded neuronal activities in the anterior IT while two well-trained monkeys were performing these tasks. Fig. 2 shows the results of analyzing 15 neurons which exhibit selective delay responses, where firing rates of each neuron during d1 and d3 of the DMS task and the PACS task are plotted; (a) and (b) are cases where the best picture, in which the neuron exhibits the highest cue response, is presented as a cue and where the paired associate of the best picture is presented, respectively. There is little difference in d1 firing rates between the two tasks, but there is a significant difference in d3. This means that subsequent neuronal activity changes significantly after the color switch.

III. SELECTIVE DESENSITIZATION METHOD

In the previous study [4], we constructed a computational model for the DPA task (referred to as the DPA model). In this model, trajectory attractors connecting from the cue-coding to target-coding states are formed by learning, and when a cue pattern is given, the target pattern is recalled by continuous state transitions along the corresponding trajectory attractor. Although the DPA model agrees well with some empirical results, this model cannot perform the above DMS/PACS task because it always recalls the same target pattern from the same

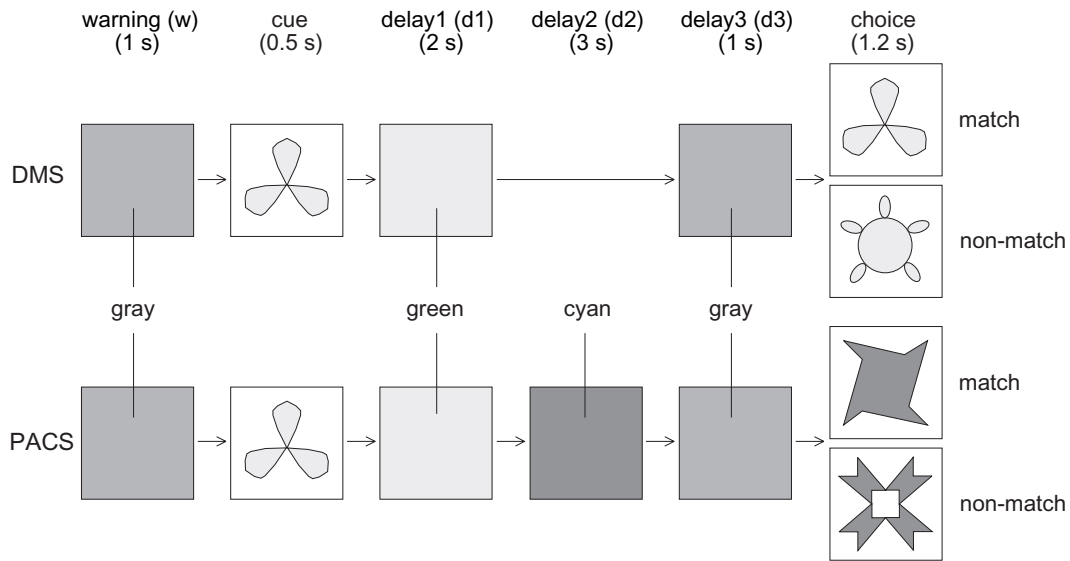


Fig. 1. Illustration of DMS and PACS tasks. Warning, gray square (1 s in both tasks); cue, 1 of 24 pictures presented as a cue stimulus (0.5 s); delay period 1, the same color as the cue stimulus (5 s in the DMS task; 2 s in the PACS task); delay period 2, the same color as the paired associate of the cue stimulus (3 s in the PACS task); delay period 3, gray square (1 s in both tasks); choice, a choice of two stimuli (1.2 s in both tasks), the same picture as the cue (matching) and the others (nonmatching) in the DMS task, or the paired associate of the cue (matching) and the others (nonmatching) in the PACS task

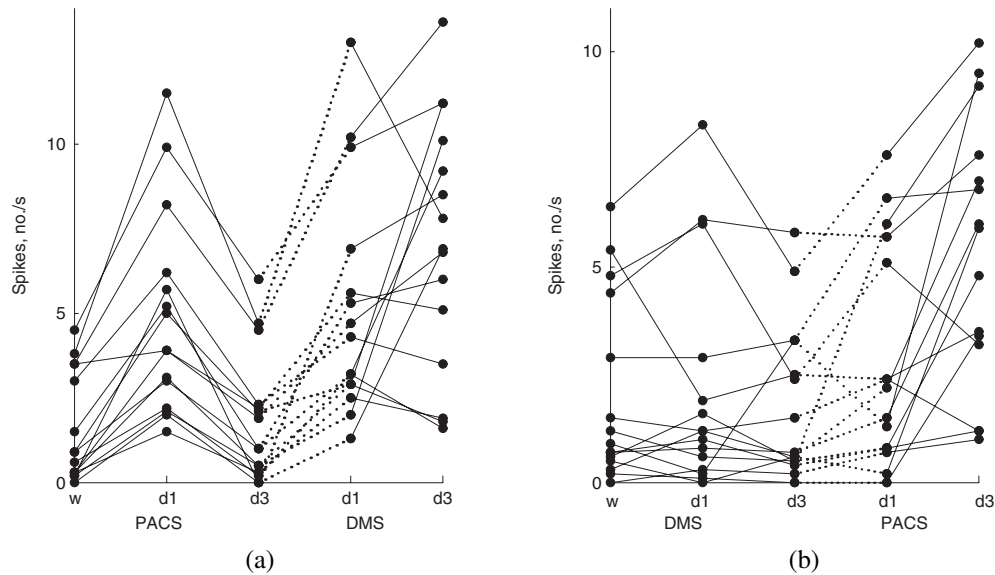


Fig. 2. Average firing rates of each neuron during delay periods of the DMS and PACS tasks. (a) Trials in which the best pictures are presented as a cue. (b) Trials in which the best picture's associates are presented as a cue. (Reproduced from [1])

cue pattern. Accordingly, we must modify this model so that it can recall different targets from the same cue depending on the context, namely, whether or not the color signal changes during the delay period.

The simplest method for associating plural targets with a single cue is to concatenate the cue pattern and the context pattern to be the input pattern. Although this method has generally been used, it causes a computational difficulty when the number of the targets corresponding to one cue is increased. Let us briefly explain the issue taking the case of

the DMS/PACS task as an example (refer to [3] for detailed discussions).

Assume the cue picture to be G_1 and thus the initial color signal is “green”; then the target is G_1 if the color signal remains “green” and C_1 if the color signal switches to “cyan”. It seems easy to recall the correct target by associating G_1 with “green” and C_1 with “cyan”. However, “green” and “cyan” should be associated also with other pictures G_2, \dots, G_{12} and C_2, \dots, C_{12} , respectively. Consequently, the direct connections between the units representing the color signal and those

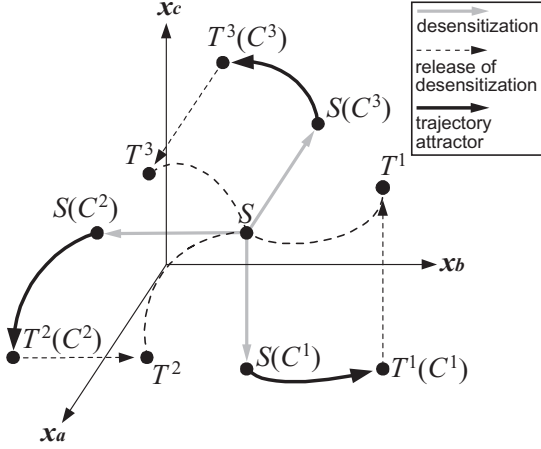


Fig. 3. Schematic process of context-dependent recall. The n-dimensional state space of the network is expressed three dimensionally.

representing the target are diluted because of the averaging effect. This means that the color signal does not aid the network to recall a particular target pattern.

This difficulty, called the problem of averaging caused by one-to-many correspondence, is deep-rooted and is unavoidable for conventional models unless they involve local representations, or hidden units each of which codes a particular combination of cue and context patterns. Such local representations, however, do not accord with physiological findings about IT. Thus we use the selective desensitization method, which was devised to resolve the problem of averaging without using local representations [2], [3].

In this method, some units are desensitized depending on the context, and the desensitization pattern or the configuration of desensitized units corresponds to the context pattern. Here, “desensitize” means to fix the output of units at a neutral value regardless of the input, which is implemented by setting the output gain to zero. In mathematical terms, assuming that the output x_i of the i -th cell ($i = 1, \dots, n$) is given by

$$x_i = g_i \cdot (f(u_i) - \bar{x}_i) + \bar{x}_i, \quad (1)$$

where u_i denotes the potential of the unit, $f(u)$ is the activation function and \bar{x}_i is the mean level of x_i , the output gain $g_i = 0$ for desensitized units and $g_i = 1$ otherwise; accordingly, desensitized units emit a constant value of \bar{x}_i , whereas valid (not desensitized) units emit $f(u_i)$. For simplicity, we regard the gain vector (g_1, \dots, g_n) in the same light as the context pattern.

Although desensitization (setting $x_i = \bar{x}_i$) is a different concept from inactivation (decreasing x_i to near 0), the two are approximately equivalent if $\bar{x}_i \simeq 0$, namely, the mean output level is low. Since the above IT neurons responded to only a few pictures, and also for simplicity of the model, we use the mechanism of sending strong inhibitory signal to some units to implement desensitization. It should be noted that there exist other possible biological mechanisms such as presynaptic inhibition and neuromodulators.

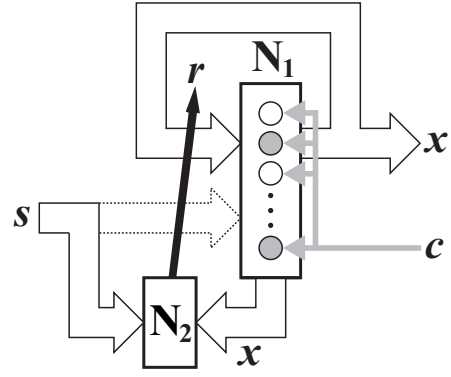


Fig. 4. Block diagram of the model. Network N_1 receives the output r of N_2 and stores memories, and its units are desensitized depending on the color signal c . N_2 receives the output x of N_1 and the external input s to generate the learning signal for N_1 .

Fig. 3 schematically shows how different targets can be recalled depending on the context, where the n-dimensional state space of a fully recurrent neural network is expressed three dimensionally. When the network at state S is modified by a context pattern $C^1 = (c_1^1, \dots, c_n^1)$, that is, unit i is desensitized when $c_i^1 = 1$, the network state x is projected onto a subspace composed of the valid units and changes into $S(C^1)$. Then x moves from $S(C^1)$ to $T^1(C^1)$ along the trajectory attractor formed in the subspace. By removing the contextual modification or releasing desensitization thereafter, x changes into T^1 when T^1 is an attractor. In the same way, when the network is modified by C^2 and C^3 , x reaches T^2 and T^3 , respectively, via state transitions along the trajectory attractors in respective subspaces.

IV. THE MODEL

We constructed the model for the DMS/PACS task by applying the selective desensitization method to the DPA model, and thus it has much in common with the DPA model. We here describe the model briefly, concentrating on the different parts (for a detailed explanation for the common parts, refer to [4]).

The model is composed of two interconnected neural networks: association network N_1 in which memories are stored and trainer network N_2 which generates the learning signal required for memory formation (Fig. 4). The output pattern $x = (x_1, \dots, x_n)$ of N_1 is sent to N_2 and also fed recurrently into N_1 , and the output pattern $r = (r_1, \dots, r_n)$ of N_2 is fed back to N_1 as the learning signal. The external input pattern $s = (s_1, \dots, s_m)$ is fed into N_2 ; although N_2 should receive s via N_1 , the direct input path to N_1 is omitted for simplicity. In addition, the color signal $c = (c_1, \dots, c_n)$ is given to N_1 to desensitize some units.

The internal structures of networks N_1 and N_2 are shown in Fig. 5 and Fig. 6, respectively, and their dynamics are

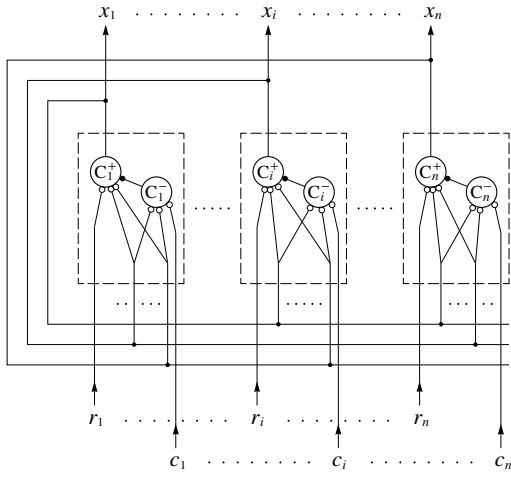


Fig. 5. Structure of network N_1 . A pair of excitatory cell C_i^+ and inhibitory cell C_i^- composes a unit, and all units are mutually connected.

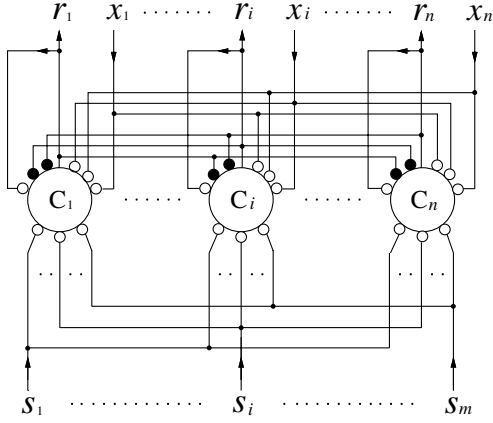


Fig. 6. Structure of network N_2 . For this network, the present model is the same as the DPA model [4].

described mathematically by

$$y_i = f \left(\sum_{j=1}^n w_{ij}^- x_j + \zeta c_i - \theta \right), \quad (2)$$

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^n w_{ij}^+ x_j - w_i^* y_i + \lambda r_i, \quad (3)$$

$$x_i = f(u_i), \quad (4)$$

$$\tau \frac{dv_i}{dt} = -v_i + \sum_{j=1}^m p_{ij} s_j + \sum_{j=1}^n q_{ij} x_j - \rho \sum_{j \neq i} r_j + \sigma r_i + h, \quad (5)$$

$$r_i = f(v_i), \quad (6)$$

$$f(u) = \frac{1}{1 + e^{-10u}}, \quad (7)$$

where w_{ij}^+ and w_{ij}^- are the synaptic weights from the j -th unit to excitatory and inhibitory cells C_i^+ and C_i^- of the i -th

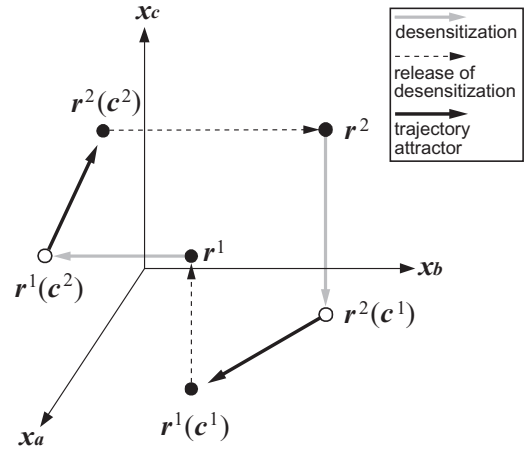


Fig. 7. Attractors formed in the state space of N_1 . Patterns r^1 and r^2 are point attractors, and paths from $r^1(c^2)$ to $r^2(c^2)$ and from $r^2(c^1)$ to $r^1(c^1)$ are trajectory attractors. N_1 can then undergo state transitions from r^1 to r^2 and vice versa by receiving color signal c^2 and c^1 , respectively.

unit, respectively, w_i^* represents the efficiency of the inhibitory synapse from C_i^- to C_i^+ , and p_{ij} and q_{ij} are random weights of input synapses to the i -th cell C_i of N_2 . u_i and v_i are the potentials of C_i^+ and C_i^- , respectively, λ denotes the input intensity of r , h is an offset, and θ , ζ , τ , ρ and σ are positive constants.

Learning of this model is performed by modifying the synaptic weights of valid N_1 units according to

$$\tau' \frac{dw_{ij}^+}{dt} = -w_{ij}^+ + \alpha r_i x_j, \quad (8)$$

$$\tau' \frac{dw_{ij}^-}{dt} = -w_{ij}^- - \beta_1 r_i x_j + \beta_2 x_i x_j + \gamma, \quad (9)$$

while each cell of the model is running according to Eqs. (2–7) [5]. Here, τ' is a time constant ($\tau' \gg \tau$), α , β_1 , and β_2 are learning coefficients ($\beta_1 < \beta_2$), and γ is a positive constant representing the lateral inhibition between units. Coefficient α may be a positive constant, but because the learning performance is better when α is a decreasing function of x_i , we adopt

$$\alpha = \begin{cases} \alpha'(\kappa - x_i) & (x_i < \kappa) \\ 0 & (x_i \geq \kappa), \end{cases} \quad (10)$$

where $\kappa \equiv \beta_1/\beta_2$ and α' is a positive constant.

By repeating this synaptic modification several times, a pattern ‘close’ (in the sense of the angle between pattern vectors) to learning signal r becomes a point attractor in the state space when r stands still. If r varies continuously, however, r leads the move of the network state x , producing a gentle flow in the direction of motion of r ; that is, a trajectory attractor is formed along the track of r [6].

The purpose of learning is to form attractors in the state space of N_1 so that N_1 can make bidirectional state transitions between two patterns, as schematically shown in Fig. 7. In this figure, r^1 and r^2 are point attractors coding paired stimulus

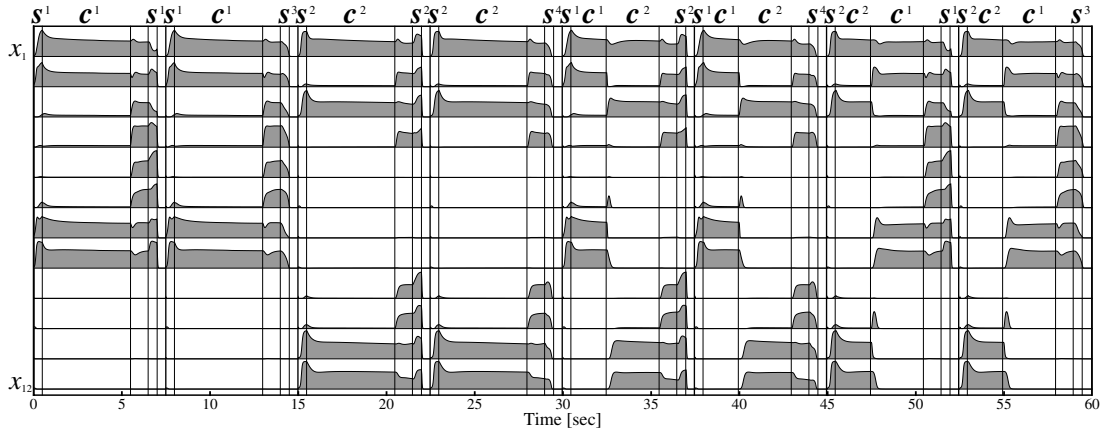


Fig. 8. Behavior of the model after learning. Response of individual units in N_1 to various cue and test patterns are vertically arranged. Thin and thick vertical lines represent the borders between periods and between trials, respectively. The first four trials are for the DMS task and the last four trials are for the PACS task.

patterns s^1 and s^2 , respectively, and line arrows from $r^1(c^2)$ to $r^2(c^2)$ and from $r^2(c^1)$ to $r^1(c^1)$ represent trajectory attractors, where $r^\mu(c^\nu)$ denotes r^μ modified by color signal c^ν .

After learning, the model performs the DMS/PACS task as follows. First, when a cue pattern, for example, s^1 , is given with color signal c^1 in the cue period, the network state x changes into $r^1(c^1)$. This state is maintained after the cue period, since it is a point attractor in the subspace corresponding to the modification by c^1 and c^1 is continuously fed during d1.

In the DMS trial in which no color switch occurs, d3 starts after d1, when the color signal is removed or c is set to $O \equiv (0, \dots, 0)$; it should be noted that in the actual task, the monkey no longer needs the color signal after the onset of d3. Consequently, x moves to the point attractor r^1 and stays there. If we then input s^1 again as a test pattern, it is thought that many units show a stronger response than when the test pattern is s^μ ($\mu \neq 1$).

In contrast, in the PACS task, the color signal is switched from c^1 to c^2 at the onset of d2, and x changes immediately into a state near $r^1(c^2)$; then x moves successively toward $r^2(c^2)$ along the trajectory attractor. When the color signal is removed in d3, x is attracted by r^2 , and thus strong responses to the test input of s^2 are expected.

V. COMPUTER SIMULATION

We carried out computer simulation using the model with a size of $n = 1000$. First, we randomly generated 12 pairs of patterns that were 1000-dimensional ($m = 1000$) binary vectors. We then input these pairs of patterns sequentially, applying the above learning procedure. Training was repeated 20 times for each pair. The parameters were

$$\begin{aligned} \tau &= 100\text{ms}, \tau' = 50000\tau, \theta = 2.1, w_i^* = 10, \\ \zeta &= 0.12, \lambda = 0.3, \rho = 0.0105, \sigma = 0.6 \\ \alpha' &= 50, \beta_1 = 25, \beta_2 = 50, \gamma = 0.05. \end{aligned}$$

Synaptic weights p_{ij} and q_{ij} were set to random numbers with mean of 0.005 and 0.001 and variance of 0.05 and 0.01, respectively.

After training, we tested the model in performing each task, varying the combination of cue and test patterns. The response of the model in 8 trials is shown in Fig. 8, where the time course of the outputs of 12 units is plotted. These units were selected from the units encoding s^1 or s^2 or both such that various kinds of activity are displayed. The first half and second half of the trials are for the DMS and PACS tasks, respectively; the odd-numbered trials are matching trials in which the test and target patterns are identical and the others are nonmatching trials.

We can see that units showing high activity in d3 further increase their output in matching trials, whereas they are depressed in nonmatching trials. We confirmed that we can distinguish between the target and nontarget patterns according to the output distribution in the test period. This indicates that the model can perform the DMS/PACS task well.

To compare the results with physiological data, we plotted the graphs shown in Fig. 9, where outputs of 20 units are plotted in the same manner as in Fig. 1. These units were randomly selected from among the units encoding s^1 or s^2 or both. We see that the graphs in Fig. 9 are very similar to those in Fig. 1 if we compensate the difference in the activity level during the warning period.

VI. DISCUSSION AND CONCLUSION

The most notable point in the above results is that in d3 of the DMS task, some units rapidly increase their output. This is because these units are desensitized during d1 and released from desensitization in d3.

Such activity enhancement was actually observed in the IT neurons, as shown in Fig. 1a, in which activity increase of three to four times is seen for 4 of 15 neurons. This phenomenon is unprecedented and has not been explained so far, since in the DMS task, the monkey is thought to

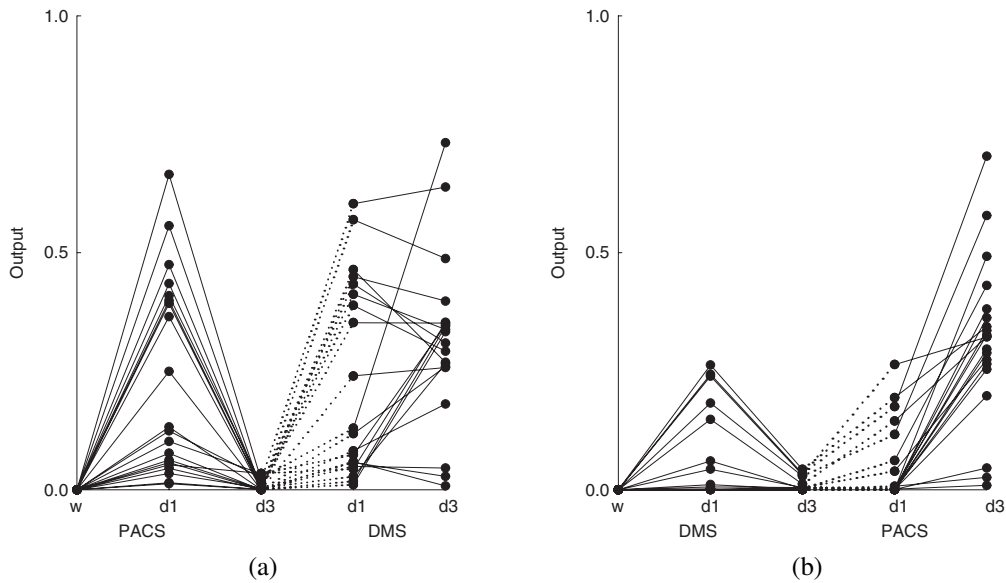


Fig. 9. Responses of each unit in both tasks when the cue is (a) the pattern encoded by the unit and (b) its paired associate.

remember the same picture throughout the delay period. The agreement between the simulation results and the empirical data, therefore, not only provides a reasonable explanation for this phenomenon, but also suggests strongly that selective desensitization actually occurs in IT.

In conclusion, we have constructed a neural network model for context-dependent association and demonstrated that the model can reproduce the distinctive activities, including an inexplicable one, of the monkey IT neurons. This supports our hypothesis that the selective desensitization method is used for context-dependent information processing in IT, and presumably in other areas of the brain as well.

ACKNOWLEDGMENT

This work was supported by Grant-in-Aid for Scientific Research (B) from JSPS (#15300068) and Grant-in-Aid for Scientific Research on Priority Areas “Informatics” (Area #006) from MEXT (#16016207).

REFERENCES

- [1] Y. Naya, K. Sakai, and Y. Miyashita, “Activity of primate inferotemporal neurons related to a sought target in pair-association task,” *Proc. Natl. Acad. Sci. USA*, vol. 93, pp. 2664–2669, 1996.
- [2] M. Morita, K. Murata, and S. Morokami, “Context-dependent sequential recall by a trajectory attractor network with selective desensitization,” *Proc. ICNNAI2003*, pp. 235–238, 2003.
- [3] M. Morita, K. Matsuzawa, and S. Morokami, “A model of context-dependent association using selective desensitization of nonmonotonic neural elements,” *Systems and Computers in Japan*, vol. 36, no. 7, pp. 73–83, 2005.
- [4] M. Morita and A. Suemitsu, “Computational modeling of pair-association memory in inferior temporal cortex,” *Cognitive Brain Research*, vol. 13, pp. 169–178, 2002.
- [5] M. Morita, “Computational study on the neural mechanism of sequential pattern memory,” *Cognitive Brain Research*, vol. 5, pp. 137–146, 1996.
- [6] M. Morita, “Memory and learning of sequential patterns by nonmonotone neural networks,” *Neural Networks*, vol. 9, pp. 1477–1489, 1996.