

## Topological maps by STDP for input space of discrete and continuous features

Yutaka Sakai

Human informatics course, Faculty of Engineering, Tamagawa University,  
Tamagawa-gakuen 6-1-1, Machida-shi, Tokyo 194-8610, Japan

**Abstract**—Position-and-scale-free representations of shapes are acquired by neurons in the inferior temporal (IT) cortex. So each neuron receives information from the whole visual field. Familiar shapes are extremely limited from all the possible shapes on the whole visual field. So they must be clustered in the shape space to have mixed structure of continuity and discreteness. We demonstrate that multiple representation can be acquired in a spike-based model for topological maps based on the spike-timing-dependent synaptic plasticity (STDP), subjected to a set of inputs on two rings, which is a simple example of mixed structure. In this representation, the position on each ring is represented by a center of active neurons and the difference of rings is represented by a detail pattern of active neurons. Neurons in the same region exhibit high activities for an input on the other ring. The result is consistent with the fact observed in IT cortex that neighboring neurons exhibit different preferences while the region of active neurons is continuously shifted for continuous changes of object.

### 1. Introduction

It is known that spike frequency of a neuron in the primary visual cortex (V1) depends on the orientation of edge in a specific small region of visual field. In other words, a V1 neuron has its preferred position and orientation. The feature space composed of position and orientation is covered by population of V1 neurons. These neurons are arranged on a cortex so that neurons preferring more close features might be more close on a cortex. Such an arrangement of neurons is called as a “topological map”. Topological maps are generally observed in sensory neural systems as well as V1. These maps are considered to be self-organized by changing of connection strength, and various models have been proposed [1, 2, 3].

On a higher stage of extracting visual features, a position-free representation is acquired. For example, in the inferior temporal (IT) cortex, a neuron raises its spike frequency selectively to a certain specific shape and independently from the visual position and size. It shows that an IT neuron receives information from the whole visual field. Sources of inputs to each neuron have quite high dimension at the retina. The number of IT neurons is too few to cover the whole possible shapes. So IT neurons are considered to cover only shapes that the animal frequently sees, which are extremely limited from the whole possi-

ble shapes. It is unlikely that all of the familiar shapes are continuously linked. On the other hand, there is a set of shapes that are linked continuously, for example a set of shapes obtained by rotation of a familiar object. Therefore, shapes of familiar objects are considered to be on separate clusters in the space of shapes. We can say that IT neurons receive inputs that have mixture of continuous and discrete structure.

In the case of such clustered inputs, the topological map models acquire patchwork-like structures dividing a cortex discretely. Continuous components in each cluster are represented by position in each patch. In fact, continuous changes of selectivity are observed in neighboring IT neurons[4, 5], and patch-like neurons selective to a shape are observed in IT cortices [5]. These facts are consistent with the models for topological maps. However, an IT neuron generally exhibits high activity even for entirely different shapes from its preferred shape. The activity pattern for a given set of shapes differs among neighboring neurons[6]. This fact is contradictory to the topological map models. It is found that IT cortex acquires a multiple representation different from simple topological maps.

In order to overcome this conflict, Wada et al. introduced random recurrent connections between local neurons[7], while classical models for topological maps have symmetric and homogeneous connection patterns. They applied it to a set of inputs distributed on two parallel rings in 3-dimensional space. As a result, a multiple representation is acquired where the position on each ring is represented by the center position of active neurons localized in a region and the difference of rings is represented by the detail activity pattern of neurons in the region. In this representation, neighboring neurons exhibit different activity patterns for the set of inputs, but they are likely to prefer similar inputs. It is consistent with the physiological results [4, 5, 6].

In these models for topological maps [1, 2, 3, 7], each neuron is assumed as an abstract unit that outputs value of spike frequency. The rule for updating connections also depends on spike frequency. On the other hand, a biological neuron outputs spikes. It is necessary to assume a unit as an ensemble of identical and independent neurons in order to reduce the realistic model to a frequency-based model. Every unit should be a neuronal ensemble. For example, in the simulation of Wada et al.[7], we should assume that many neuronal ensembles on a cortex receive inputs commonly from three neuronal ensembles. There is not necessarily such a discrete grouping of neurons in a cortex. The

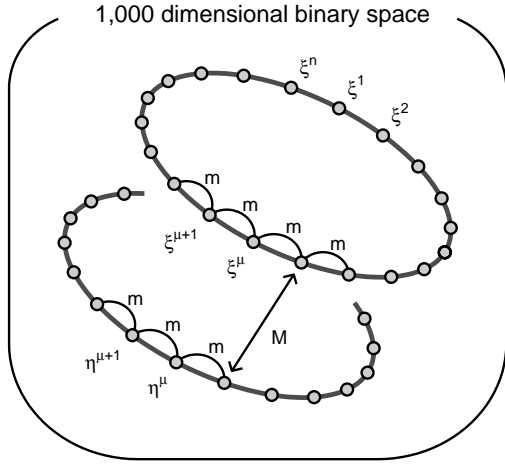


Figure 1: Schematic representation of the structure of input pattern set.

present study attempts to examine reproducibility for multiple topological maps on a spike-based model that is free from the restriction of discrete groupings.

A spike-based model for topological maps has been proposed by Song and Abbott[8], based on the physiological phenomena known as the spike-timing-dependent synaptic plasticity(STDP)[9]. In their simulation, a topological map for inputs on one ring in high dimensional input space is organized on one-dimensional cortex. The present study examines what maps are organized in the equivalent model for a set of inputs on two rings like Wada et al.[7].

## 2. Model

In the present study, we use a model equivalent to the model in [8]. There are 1,000 neurons on the input layer and 200 neurons arranged circularly on the output layer. The output layer corresponds to a cortex. Each neuron on the output layer receives excitatory connections from randomly selected input neurons at rate 0.2, excitatory connections from 80 neighboring neurons on the output layer, and inhibitory connections from all of the output neurons. In addition, each neuron receives random inhibitory inputs (Poisson process) of size  $w_{\text{inh}}$  at rate 400Hz.

Each input neuron emits random spikes (Poisson process) at the current rate depending on its state: “up-state” with high spike rate  $\lambda^{\text{up}}$  or “down-state” with low spike rate  $\lambda^{\text{down}}$ . Up-down pattern is given as an input. Let up-state and down-state correspond to the values 1 and 0 respectively, and an input pattern is described by 1,000-dimensional binary vector  $\xi$ . We prepare two sets of  $n$  random binary patterns  $\{\xi^\mu | \mu = 1, 2, \dots, n\}$  and  $\{\eta^\mu | \mu = 1, 2, \dots, n\}$  at constant bit rate,  $\sum_j \xi_j^\mu = \sum_j \eta_j^\mu = R$ , which are arranged on respective rings at constant intervals measured by the over-

lap or Hamming distance,

$$\begin{aligned} \xi^1 \cdot \xi^2 &= \dots = \xi^\mu \cdot \xi^{\mu+1} = \dots = \xi^n \cdot \xi^1 = Rm, \\ \eta^1 \cdot \eta^2 &= \dots = \eta^\mu \cdot \eta^{\mu+1} = \dots = \eta^n \cdot \eta^1 = Rm, \\ \xi^1 \cdot \eta^1 &= \dots = \xi^\mu \cdot \eta^\mu = \dots = \xi^n \cdot \eta^n = RM, \end{aligned} \quad (1)$$

where  $m$  and  $M$  represent the intra-ring and inter-ring overlaps. This set of inputs  $\xi^\mu, \eta^\mu$  is topologically equivalent to the parallel two rings introduced by Wada et al.[7]. When the intra-ring overlap  $m$  is sufficiently larger than the inter-ring overlap  $M$ , the position on a ring described by  $\mu$  is considered to be a relatively continuous feature to the discreteness of the difference of rings.

Each output neuron obeys the leaky integrate-and-fire model. We use the current-based synaptic inputs instead of conductance-based inputs. So we introduce a constant synaptic delay of 5msec to include effectively the open-and-closing time scale of conductance. Let  $s_j(t)$  denote series of delta functions shifted 5msec from the spike timings of the  $j$ -th neuron,

$$s_j(t) = \sum_k \delta(t - t_k^j - 5), \quad (2)$$

where  $\{\dots, t_k^j, t_{k+1}^j, \dots\}$  are spike timings of the  $j$ -th neuron. The index  $j$  is assigned serially through every input and output neuron. The membrane potential of the  $i$ -th output neuron obeys the following equation,

$$\begin{aligned} \frac{dv_i}{dt} &= -\frac{v_i}{\tau} + \sum_{j \in \text{exc.}} w_{ij} s_j(t) - w_{\text{inh}} \sum_{j \in \text{inh.}} s_j(t) \\ v_i > v_{\text{th}} &\Rightarrow \text{spike} \ \& \ v_i = 0, \end{aligned} \quad (3)$$

where the summation  $\sum_{j \in \text{exc.}}$  is for all the excitatory connections received from the input and output neurons, and the summation  $\sum_{j \in \text{inh.}}$  is for all the inhibitory inputs received from all the other output neurons and the source of random inhibitory inputs. All the excitatory connections  $w_{ij}$  is updated according to the STDP rule[8], while the recurrent connections on the output layer are fixed in the frequency-based models [1, 2, 3, 7]. Let the time scales of timing window of STDP rule be equal to the membrane time scale  $\tau$ , and the connections  $w_{ij}$  obey the following equations,

$$\frac{da_i}{dt} = -\frac{a_i}{\tau} + s_i(t), \quad (4)$$

$$\frac{dw_{ij}}{dt} = A^+ a_j \Theta[1 - w_{ij}] s_i(t) - A^- a_i \Theta[w_{ij}] s_j(t), \quad (5)$$

where the function  $\Theta[\cdot]$  denotes the unit step function (Heaviside function), introduced in order to limit the strength of connection in the region  $w \in [0, 1]$ .

The model parameters are fixed so that this model should be equivalent to that of Song and Abbott[8],  $\tau = 20\text{ms}$ ,  $w_{\text{inh}} = 0.2$ ,  $v_{\text{th}} = 20$ ,  $A^+ = 0.005$ ,  $A^- = 1.05A^+$ ,  $\lambda^{\text{up}} = 36.7\text{Hz}$ ,  $\lambda^{\text{down}} = 3.33\text{Hz}$ , and  $R = 200$ . The number of input patterns and the intra-ring overlap are also fixed throughout the following simulations,  $n = 1,000$ ,  $m = 0.995$ .

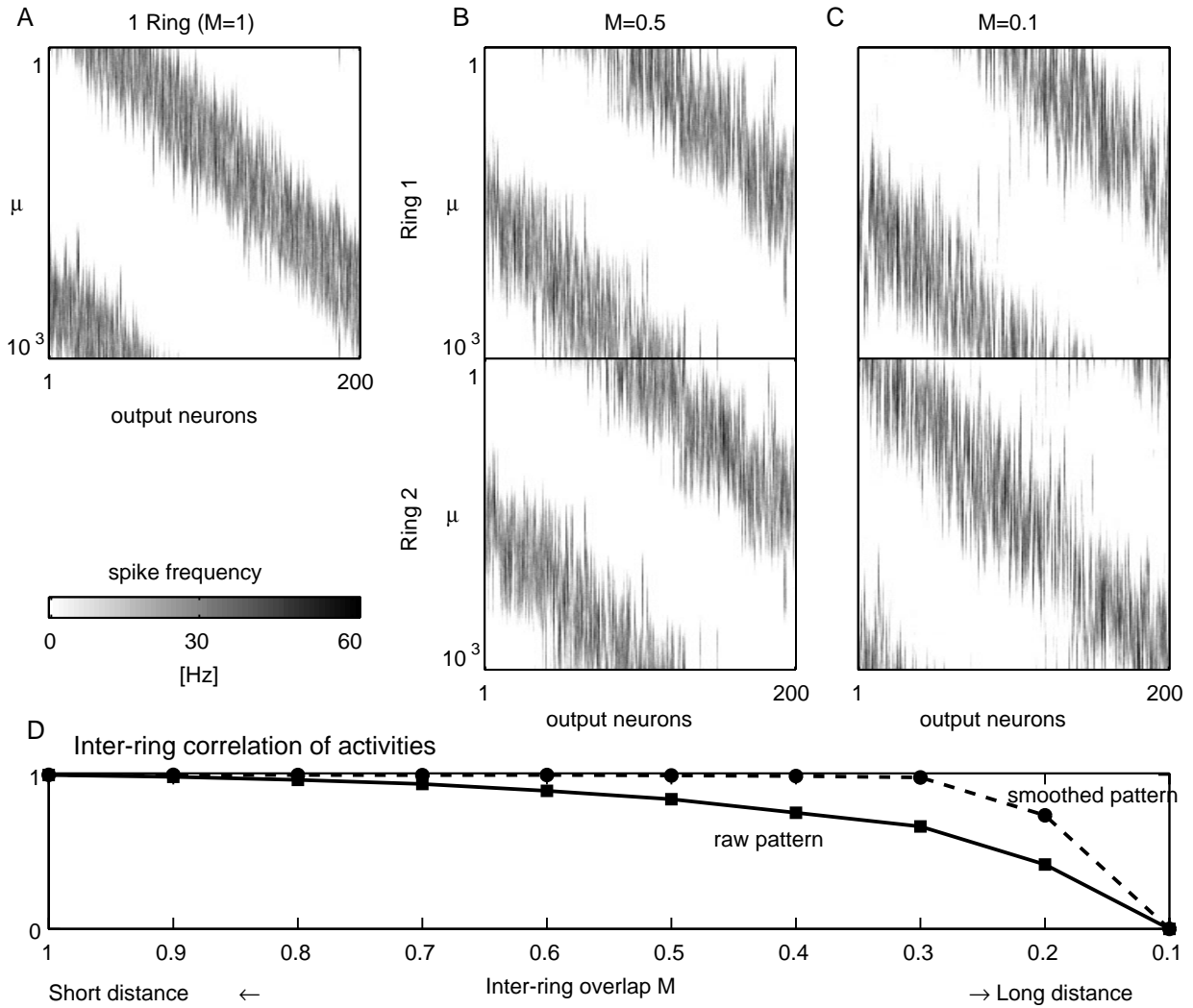


Figure 2: A-C: Spike frequency patterns  $\alpha^\mu$  and  $\beta^\mu$  of output neurons for input patterns  $\xi^\mu$  (Ring1) and  $\eta^\mu$  (Ring2) respectively after sufficient self-organization (10,000sec) for the set of input patterns  $\{\xi^\mu, \eta^\mu\}$ . D: The mean direction cosines of the activity pattern vectors  $\alpha^\mu$  and  $\beta^\mu$  (solid line marked by squares) and of the smoothed activity pattern vectors  $\tilde{\alpha}^\mu$  and  $\tilde{\beta}^\mu$  filtered by the smoothing window with 40 neighbors (dashed line marked by circles).

### 3. Simulation results

A set of input patterns  $\{\xi^\mu, \eta^\mu\}$  are prepared for a certain inter-ring overlap  $M$  according to (1). Current input pattern switches randomly among the prepared set of input patterns at a mean interval of 40msec (Poisson process). Obeying the equations (2–5), the excitatory connections  $\{w_{ij}\}$  are updated depending on the spike timings. The initial values of feed-forward connections from the input layer to the output layer are uniformly distributed in  $[0,1]$  and the initial values of recurrent connections on the output layer are set to zero. After a long enough time about 10,000sec, the distribution of  $\{w_{ij}\}$  approaches a steady distribution. After approaching the steady state, we stop updating the connections (5) and examine activity pattern of the output neurons for each input pattern. The current input pattern is fixed to

a certain pattern  $\mu$  and spike frequencies of output neurons are estimated from spike series for a long enough simulation. In the same way, the frequency patterns or the activity patterns for all the input patterns are obtained,

$$\alpha^\mu = (\alpha_1^\mu, \alpha_2^\mu, \dots, \alpha_{200}^\mu), \quad \beta^\mu = (\beta_1^\mu, \beta_2^\mu, \dots, \beta_{200}^\mu),$$

where  $\alpha^\mu$  and  $\beta^\mu$  correspond to the activity patterns for the input patterns on the two rings,  $\xi^\mu$  and  $\eta^\mu$ , respectively.

Figure 2 shows the frequency patterns in several cases of inter-ring overlap values  $M = 1, 0.5, 0.1$  as grey scale plots. The case that  $M = 1$  corresponds to the case that a ring completely coincides with the other, namely the case of one ring which is equivalent to the simulation of Song and Abbott[8]. In the case that  $M = 1$  (Figure 2A), localized neurons exhibit high activities for each input pattern and the center of active neurons is shifted as the position of

the input pattern on the ring is shifted. It is confirmed that a topological map representing the position on the ring can be self-organized in the present spike-based model as well as the demonstration of Song and Abbott[8].

In the case of the more separate rings,  $M = 0.5$  (Figure 2B), we can see the same type of correspondence between the input position on a ring and the center of active neurons. In addition, neurons in the same region exhibit high activity even for the input pattern on the corresponding position of the other ring. So the center of active neurons represents the input position on the rings regardless of the difference of rings. Is there any difference in the activity patterns between the two rings?

Figure 2D shows the inter-ring correlation of activity patterns,  $C$  (solid line marked with squares), and of smoothed activity patterns,  $\tilde{C}$  (dashed line marked with circles), which are defined to be the mean direction cosines between the activity vectors  $\alpha^\mu$  and  $\beta^\mu$ , and the smoothed activity vectors  $\tilde{\alpha}^\mu$  and  $\tilde{\beta}^\mu$ ,

$$C = \frac{1}{n} \sum_{\mu=1}^n \frac{\alpha^\mu \cdot \beta^\mu}{|\alpha^\mu| |\beta^\mu|}, \quad \tilde{C} = \frac{1}{n} \sum_{\mu=1}^n \frac{\tilde{\alpha}^\mu \cdot \tilde{\beta}^\mu}{|\tilde{\alpha}^\mu| |\tilde{\beta}^\mu|},$$

$$\tilde{\alpha}_i^\mu = \sum_{k=-20}^{20} \alpha_{i+k}^\mu, \quad \tilde{\beta}_i^\mu = \sum_{k=-20}^{20} \beta_{i+k}^\mu,$$

where the index  $i$  for  $\alpha$  and  $\beta$  satisfies periodic boundary conditions, for example  $\alpha_{-k}^\mu$  means  $\alpha_{200-k}^\mu$ . We can see that the inter-ring correlation of activity patterns decreases as the input rings separate more, while that of smoothed activity patterns stays near 1. So in the middle range of inter-ring overlap,  $M = 0.3 \sim 0.5$ , the position on each ring is represented by the center of active neurons on the output layer, and the difference of rings is represented by the detail activity patterns of the localized active neurons.

In the case of the still more separate rings,  $M = 0.1$  (Figure 2C), topological maps for the position on each ring is still observed, but the correspondence between the two rings is broken. The inter-ring correlation of the smoothed activity patterns also vanishes (dashed line marked by circle in Figure 2D).

#### 4. Discussion

We found that the present spike-based model acquires the dual representation where the position on each ring is represented by the center of active neurons on the output layer, and the difference of rings is represented by the detail activity patterns of the localized active neurons. In addition, in the middle range of inter-ring distance, positions on the two rings represented by a center of active neurons correspond. This dual representation is consistent with the results in Wada et al.[7]. The duality commonly originates in random recurrent connections on the output layer, which are also self-organized by the STDP rule in the present model, while they are prepared in advance in the model of Wada et al.[7].

The present result implies the possibility that multiple topological maps for different features are acquired in a sheet of neural field and the different maps are distinguished with the detail pattern of active neurons. The present study demonstrated that only two topological maps corresponding to two rings are acquired. It is interesting problem how many maps are acquired. The capacity will depend on the number of neurons.

#### Acknowledgments

This work is supported in part by Grants-in-Aid for Scientific Research on Priority Areas, No.17021038, 2005.

#### References

- [1] C. von der Malsburg, "Self-organization of orientation sensitive cells in the striate cortex," *Kybernetik*, vol. 14, 1973. Reprinted in [10].
- [2] A. Takeuchi and S. Amari, "Formation of topographic maps and columnar microstructures in nerve fields," *Biological Cybernetics*, vol. 35, pp. 63–72, 1979.
- [3] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, 1982. Reprinted in [10].
- [4] I. Fujita, K. Tanaka, M. Ito, and K. Cheng, "Columns for visual features of objects in monkey inferotemporal cortex," *Nature*, vol. 360, pp. 343–346, 1992.
- [5] G. Wang, M. Tanifuji, and K. Tanaka, "Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging," *Neuroscience Research*, vol. 32, pp. 33–46, 1998.
- [6] H. Tamura, H. Kaneko, K. Kawasaki, and I. Fujita, "Presumed inhibitory neurons in the macaque inferior temporal cortex: visual response properties and functional interactions with adjacent neurons," *J. Neurophysiology*, vol. 91, pp. 2782–2796, 2004.
- [7] K. Wada, K. Kurata, and M. Okada, "Self-organization of globally continuous and locally distributed information representation," *Neural Networks*, vol. 17, pp. 1039–1049, 2004.
- [8] S. Song and L. F. Abbott, "Cortical development and remapping through spike timing-dependent plasticity," *Neuron*, vol. 32, pp. 339–350, 2001.
- [9] G. Bi and M. Poo, "Activity-induced synaptic modifications in hippocampal culture: dependence on spike timing, synaptic strength and cell type," *J. Neuroscience*, vol. 18, pp. 10464–10472, 1998.
- [10] J. Anderson and E. Rosenfeld, eds., *Neurocomputing: Foundations of Research*. Cambridge: MIT Press, 1988.