2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

Kohonen Feature Map Probabilistic Associative Memory based on Weights Distribution and Area Neuron Increase and Decrease

Takahiro HADA† and Yuko OSANA†

†School of Computer Science, Tokyo University of Technology
1404-1 Katakura, Hachioji, Tokyo, 192-0982, Japan Email: osana@cs.teu.ac.jp

Abstract—In this paper, we propose a Kohonen Feature Map Probabilistic Associative Memory based on Weights Distribution and Area Neuron Increase and Decrease. This model is based on the conventional Kohonen Feature Map Probabilistic Associative Memory based on Weights Distribution. In the proposed model, the winner neuron is selected from the neurons in the Map Layer whose connection weights are similar to the input pattern, and the associations based on weights distribution are realized. Moreover, the weight distribution in the Map Layer can be modified by the increase and decrease of neurons in each area. This model has enough robustness for noisy input and damaged neurons. We carried out a series of computer experiments and confirmed the effectiveness of the proposed model.

1. Introduction

Recently, neural networks are drawing much attention as a method to realize flexible information processing. Neural networks consider neuron groups of the brain in the creature, and imitate these neurons technologically. Neural networks have some features, especially one of the important features is that the networks can learn to acquire the ability of information processing.

In the field of neural network, many models have been proposed. In these models, the learning process and the recall process are divided, and therefore they need all information to learn in advance. However, in the real world, it is very difficult to get all information to learn in advance, so we need the model whose learning process and recall process are not divided. As such model, Grossberg and Carpenter proposed the ART (Adaptive Resonance Theory)[1]. However, the ART is based on the local representation, and therefore it is not robust for damaged neurons. While in the field of associative memories, some models have been proposed[2]-[4]. Since these models are based on the distributed representation, they have the robustness for damaged neurons. However, their storage capacities are small because their learning algorithm is based on the Hebbian learning.

On the other hand, the Kohonen Feature Map (KFM) associative memory[5] has been proposed. Although the KFM associative memory is based on the local representation as similar as the ART[1], it can learn new patterns successively[6], and its storage capacity is larger than that of models in refs.[2]-[4]. It can deal with auto and hetero associations and the associations for plural sequential patterns including common terms[7]. Moreover, the KFM associative memory with area representation[8] has been proposed. In the model, the area representation[9] was introduced to the KFM associative memory, and it has robustness for damaged neurons. However, it can not deal with one-to-many associations, and associations of analog patterns. As the model which can deal with analog patterns, the Kohonen Feature Map Associative Memory with Refractoriness based on Area Representation (KFMAM-R-AR)[10] has been proposed. In the model, one-to-many associations are realized by refractoriness of neurons. Moreover, the Kohonen Feature Map Probabilistic Associative Memory based on Weights Distribution (KFMPAM-WD)[11] has been proposed. It is based on the conventional KFM associative memory with area representation[8] and can realize probabilistic association for the training set including one-to-many relations.

In this paper, we propose a Kohonen Feature Map Probabilistic Associative Memory based on Weights Distribution and Area Neuron Increase and Decrease. This model is based on the conventional Kohonen Feature Map Probabilistic Associative Memory based on Weights Distribution[11]. In the proposed model, the winner neuron is selected from the neurons in the Map Layer whose connection weights are similar to the input pattern, and the associations based on weights distribution are realized. Moreover, the weights distribution in the Map Layer can be modified by the increase and decrease of neurons in each area. This model has enough robustness for noisy input and damaged neurons.

2. KFM Probabilistic Associative Memory based on Weights Distribution and Area Neuron Increase and Decrease

Here, we explain the proposed Kohonen Feature Map Probabilistic Associative Memory based on Weights Distribution and Area Neuron Increase and Decrease.

2.1. Structure

Figure 1 shows the structure of the proposed model. As shown in Fig.1, the proposed model has two layers; (1)

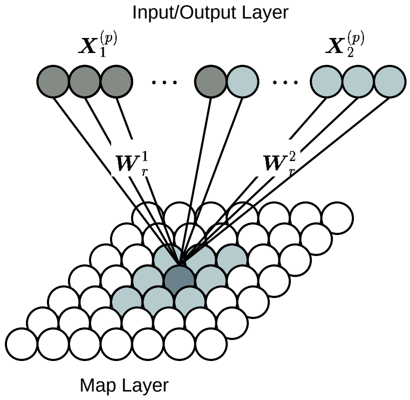


Figure 1: Structure of Proposed Model.

Input/Output(I/O) Layer and (2) Map Layer, and the I/O Layer is divided into some parts.

2.2. Learning Process

In the learning algorithm of the proposed model, the patterns are learned as follows:

(1) In the network with the Map Layer composed of $x_{max} \times y_{max}$ neurons, the connection weights are initialized randomly. Here, x_{max} is the initial number of neurons of a horizontal direction, and y_{max} is the initial number of neurons of a vertical direction. In the initial state, $x_{max} \times y_{max}$ neurons are arranged at the coordinates $(0, 0)$, $(1, 0)$, $\cdots$, $(x_{max} - 1, 0)$, $(1, 0)$, $\cdots$, $(x_{max} - 1, y_{max} - 1)$.

(2) The Euclidean distance between the learning vector $\boldsymbol{X}^{(p)}$ and the connection weights vector $\boldsymbol{W}_i$, $d(\boldsymbol{X}^{(p)}, \boldsymbol{W}_i)$ is calculated.

$$d(\boldsymbol{X}^{(p)}, \boldsymbol{W}_i) = \sqrt{\sum_{k=1}^{M} (X_k^{(p)} - W_{ik})^2} \qquad (1)$$

where M is the number of neurons in the Input/Output Layer. If $d(\boldsymbol{X}^{(p)}, \boldsymbol{W}_i) > \theta^t$ is satisfied for all neurons, the input pattern $\boldsymbol{X}^{(p)}$ is regarded as an unknown pattern. If the input pattern is regarded as a known pattern, go to (6).

(3) The neuron which is the center of the learning area c is determined as follows:

$$c = \underset{\substack{i\,:\, D_{iz} + D_{zi} \leq d_{iz} \leq d_{iz} + 1 \\ (\text{for } \forall z \in F)}}{\operatorname{argmin}} d(\boldsymbol{X}^{(p)}, \boldsymbol{W}_i). \qquad (2)$$

In this equation, the neuron whose Euclidean distance between its connection weights and the learning vector is minimum in the neurons which can be take areas without overlaps to the areas corresponding to the patterns which are already trained. In Eq.(2), F is the set of the weight-fixed neurons, d_{iz} is the distance between the neuron i and the weight-fixed neuron z. And D_{ij} is the radius of the ellipse area whose center is the neuron i for the direction to the neuron i for the direction to the neuron j, and is given by

$$D_{ij} = \begin{cases} \sqrt{\dfrac{{a_i}^2 {b_i}^2}{{b_i}^2 + {m_{ij}}^2 {a_i}^2} ({m_{ij}}^2 + 1)}, & \\ & (d_{ij}^x \neq 0 \text{ and } d_{ij}^y \neq 0) \\ a_i, & (d_{ij}^y = 0) \\ b_i, & (d_{ij}^x = 0) \end{cases} \qquad (3)$$

where a_i and b_i are the long and short radius of the ellipse, and m_{ij} is the slope of the line through the neurons i and j, and is given by

$$m_{ij} = \frac{d_{ij}^y}{d_{ij}^x}, \qquad (d_{ij}^x \neq 0). \qquad (4)$$

(4) If $d(\boldsymbol{X}^{(p)}, \boldsymbol{W}_c) > \theta^t$ is satisfied, the connection weights of the neurons in the ellipse whose center is the neuron c are updated as follows:

$$\boldsymbol{W}_i(t+1) = \begin{cases} \boldsymbol{X}^{(p)}, & (\theta_1^{learn} \leq H(\overline{d_{ci}})) \\ \boldsymbol{W}_i(t) + H(\overline{d_{ci}})(\boldsymbol{X}^{(p)} - \boldsymbol{W}_i(t)), & \\ & (\theta_2^{learn} \leq H(\overline{d_{ci}}) < \theta_1^{learn} \\ & \text{and } H(\overline{d_{i^*i}}) < \theta_1^{learn}) \\ \boldsymbol{W}_i(t), & (\text{otherwise}) \end{cases} \qquad (5)$$

where θ_1^{learn} and θ_2^{learn} are the thresholds, and i^* is the nearest weight-fixed neuron from the neuron i. $H(\overline{d_{ci}})$ and $H(\overline{d_{i^*i}})$ are the semi-fixed functions and are given by

$$H(\overline{d_{ci}}) = \frac{1}{1 + \exp\left(\dfrac{\overline{d_{ci}} - D}{\varepsilon}\right)} \qquad (6)$$

where ε is the steepness parameter of the function $H(\overline{d_{ci}})$, and D $(1 < D)$ is the size of the neighborhood area. $\overline{d_{ci}}$ is the normalized distance between the center neuron of the area c and the neuron i, and is given by

$$\overline{d_{ci}} = \frac{d_{ci}}{D_{ci}}. \qquad (7)$$

If there is no weight-fixed neuron,

$$H(\overline{d_{ci}}) = 0 \qquad (8)$$

is used.

(5) The connection weights of the neuron c $\boldsymbol{W}_c$ are fixed.

(6) (2)~(5) are iterated when a new pattern set is given.

2.3. Neuron Increase and Decrease

In the proposed model, the weights distribution in the Map Layer can be modified by the increase and decrease of neurons in each area.

2.3.1. Neuron Increase

In the proposed model, the neuron is added at the position corresponding to the neurons that exist in the initial Map Layer.

When the N_z neurons whose connection weights are same as that of the center neuron are in the area z, the neuron is added as follows. Here, N_z^{ini} is the number of the neurons when the area z is generated, N_z^{min} is the minimum number of the neurons in the area z, and N_z^{max} is the maximum number of the neurons in the area z.

(1) $N_z^{ini} \leq N_z < N_z^{max}$

If $N_z^{ini} \leq N_z < N_z^{max}$ is satisfied, the new neuron is added in the area z.

The reference neuron j^* corresponding to the adding neuron j is given by

$$j^* = \left((N_z - N_z^{ini}) \bmod (N_z^{ini} - 1)\right) + 1 \tag{9}$$

where mod shows remainder operation, and j^* shows the sequential serial number for the neurons except for the center neuron which exist when the area is generated (See Fig.2). The coordinates of the new neuron is given by

$$x_j = x_{j^*} - \mathrm{sgn}(x_{j^*}) \cdot \frac{S_j^{add}}{C_z} \tag{10}$$

$$y_j = y_{j^*} - \mathrm{sgn}(y_{j^*}) \cdot \frac{S_j^{add}}{C_z} \tag{11}$$

where x_{j^*}, y_{j^*} are the coordinates of the reference neuron j^*, and sgn$(\cdot)$ is the sign function. C_z is the coefficient for the distance between neurons in the area z, and is given by

$$C_z = \left\lceil \frac{N_z^{max} - N_z^{ini}}{N_z^{ini} - 1} \right\rceil + 1. \tag{12}$$

S_j^{add} is

$$S_j^{add} = \left\lfloor \frac{N_z - N_z^{ini}}{N_z^{ini} - 1} \right\rfloor. \tag{13}$$

where $\lceil\cdot\rceil$ is the ceiling function and $\lfloor\cdot\rfloor$ is the floor function.

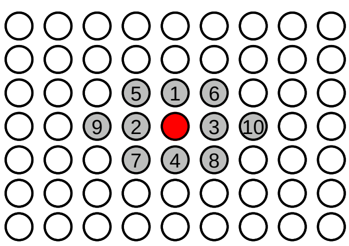


Figure 2: Sequential Serial Number for Original Neurons in Area.

(2) $N_z^{min} < N_z \leq N_z^{ini}$

If the number of the neurons in the area N_z is smaller than N_z^{ini}, the connection weights of the neuron in the area (the neuron which satisfy $d_{ci} \leq D_{ci}$ and whose connection weights are same as that of the center neuron) are initialized randomly. The connection weights are updated in the neuron whose sequential serial number is $N_z - 1$.

2.4. Recall Process

In the recall process of the proposed model, when the pattern $\boldsymbol{X}$ is given to the I/O Layer, the output of the neuron i in the Map Layer, x_i^{map} is calculated by

$$x_i^{map} = \begin{cases} 1, & (i = r) \\ 0, & (\text{otherwise}) \end{cases} \tag{14}$$

where r is selected randomly from the neurons which satisfy

$$\frac{1}{N^{in}} \sum_{k \in C} g(X_k - W_{ik}) < \theta^{map} \tag{15}$$

where N^{in} is the number of neurons which receive the input in the I/O Layer. $g(\cdot)$ is given by

$$g(h) = \begin{cases} 1, & (|h| < \theta^d) \\ 0, & (\text{otherwise}) \end{cases} \tag{16}$$

where θ^d is the threshold, and θ^{map} is the threshold of the neuron in the Map Layer.

In the proposed model, one of the neurons whose connection weights are similar to the input pattern are selected randomly as the winner neuron. So, the probabilistic association can be realized based on the weights distribution. For example, if the training patterns including the common term such as $\{\boldsymbol{X}, \boldsymbol{Y}_1\}$, $\{\boldsymbol{X}, \boldsymbol{Y}_2\}$ are memorized, and the number of the neurons whose connection weights are similar to the pattern pair $\{\boldsymbol{X}, \boldsymbol{Y}_1\}$ is larger than the number of the neurons whose connection weights are similar to the pattern pair $\{\boldsymbol{X}, \boldsymbol{Y}_2\}$, then the probability that the pattern pair $\{\boldsymbol{X}, \boldsymbol{Y}_1\}$ is recalled is higher than the probability that the pattern pair $\{\boldsymbol{X}, \boldsymbol{Y}_2\}$ is recalled.

When the binary pattern $\boldsymbol{X}$ is given to the I/O Layer, the output of the neuron k in the I/O Layer x_k^{io} is given by

$$x_k^{io} = \begin{cases} 1, & (W_{rk} \geq \theta_b^{in}) \\ 0, & (\text{otherwise}) \end{cases} \tag{17}$$

where θ_b^{in} is the threshold of the neurons in the I/O Layer.

When the analog pattern $\boldsymbol{X}$ is given to the I/O Layer, the output of the neuron k in the I/O Layer x_k^{io} is given by

$$x_k^{io} = W_{rk}. \tag{18}$$

3. Computer Experiment Results

Here, we show the computer experiment results to demonstrate the effectiveness of the proposed model.

Table 1: Relation between Area Size and Recall Times.

Input	Output	# of Neurons	Recall Times
Pattern 1	Pattern A	35 (3.2)	270 (3.3)
	Pattern B	11 (1.0)	82 (1.0)
	Pattern C	19 (1.7)	148 (1.8)
Pattern 2	Pattern D	23 (1.4)	208 (1.5)
	Pattern E	17 (1.0)	152 (1.1)
	Pattern F	17 (1.0)	140 (1.0)

Table 2: Relation between Area Size and Recall Times (After Resizing of Areas).

Input	Output	# of Neurons	Recall Times
Pattern 1	Pattern A	105 (9.5)	294 (9.8)
	Pattern B	55 (5.0)	176 (5.9)
	Pattern C	11 (1.0)	30 (1.0)
Pattern 2	Pattern D	35 (5.0)	176 (5.2)
	Pattern E	70 (10.0)	290 (8.5)
	Pattern F	7 (1.0)	34 (1.0)

3.1. Relation between Area Size and Recall Times

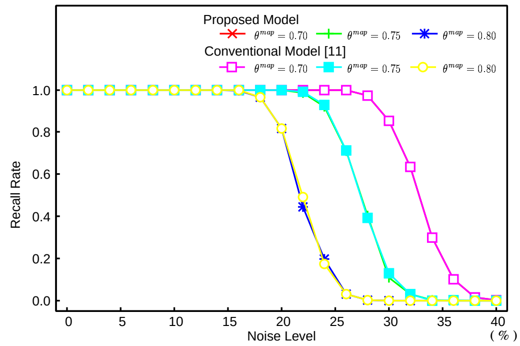
Here, we examined the relation between the area size and the recall times of the proposed model. In this experiment, six pattern pairs including 1-to-3 relations were memorized in the proposed model. Table 1 shows the relation between the area size and the recall times. And Table 2 shows the relation between the area size and the recall times after the resizing of areas. As shown in these tables, the proposed model can realize probabilistic association based on the weights distribution, and can modify the weights distribution by the increase and decrease of neurons in each area.

3.2. Robustness for Noisy Input/Damaged Neurons

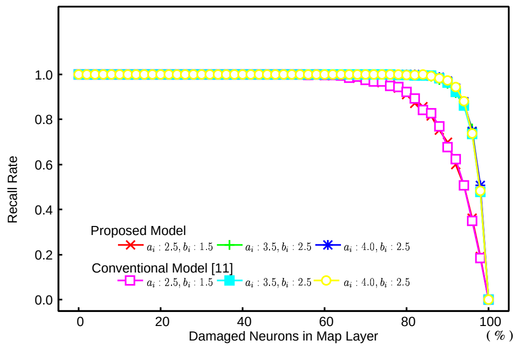
Figure 3 shows the robustness for noisy input/damaged neurons in the proposed model. As shown in these figures, we confirmed that the proposed model has robustness for noisy input/damaged neurons.

4. Conclusions

In this paper, we have proposed the Kohonen Feature Map Probabilistic Associative Memory based on Weights Distribution and Area Neuron Increase and Decrease. We carried out a series of computer experiments and confirmed the effectiveness of the proposed model.



(a) Robustness for Noisy Input



(b) Robustness for Damaged Neurons

Figure 3: Robustness for Noisy Input/Damaged Neurons.

References

[1] G. A. Carpenter, S. Grossberg, Pattern Recognition by Self-organizing Neural Networks, The MIT Press, 1995.

[2] M. Watanabe, K. Aihara, S. Kondo, "Automatic learning in chaotic neural networks," IEICE-A, Vol.J78-A, No.6, pp.686–691, 1995 (in Japanese).

[3] T. Arai and Y. Osana: "Hetero chaotic associative memory for successive learning with give up function — One-to-many associations —," Proceedings of IASTED Artificial Intelligence and Applications, Innsbruck, 2006.

[4] M. Ando,Y. Okuno and Y. Osana: "Hetero chaotic associative memory for successive learning with multi-winners competition," Proceedings of IEEE and INNS International Joint Conference on Neural Networks, Vancouver, 2006.

[5] H. Ichiki, M. Hagiwara and M. Nakagawa, "Kohonen feature maps as a supervised learning machine," Proc. of IEEE International Conference on Neural Networks, pp.1944–1948, 1993.

[6] T. Yamada, M. Hattori, M. Morisawa and H. Ito, "Sequential learning for associative memory using Kohonen feature map," Proc. of IEEE and INNS International Joint Conference on Neural Networks, paper no.555, Washington D.C., 1999.

[7] M. Hattori, H. Arisumi and H. Ito, "SOM Associative Memory for Temporal Sequences," Proc. of IEEE and INNS International Joint Conference on Neural Networks, pp.950-955, Honolulu, 2002.

[8] H. Abe and Y. Osana : "Kohonen feature map associative memory with area representation," Proceedings of IASTED Artificial Intelligence and Applications, Innsbruck, 2006.

[9] N. Ikeda and M. Hagiwara, "A proposal of novel knowledge representation (Area representation) and the implementation by neural network", International Conference on Computational Intelligence and Neuroscience , III, pp,430-433, 1997.

[10] T. Imabayashi and Y. Osana : "Implementation of association of one-to-many associations and the analog pattern in Kohonen feature map associative memory with area representation," Proceedings of IASTED Artificial Intelligence and Applications, Innsbruck, 2008.

[11] M. Koike and Y. Osana : "Kohonen feature map probabilistic associative memory based on weights distribution," Proceedings of IASTED Artificial Intelligence and Applications, Innsbruck, 2010.