Reducing topological defects in self-organizing maps using multiple scale neighborhood functions

Kazushi Murakoshi *, Yuichi Sato

Department of Knowledge-based Information Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tenpaku-cho, Toyohashi 441-8580, Japan

Received 24 April 2006; revised 4 July 2006; accepted 9 July 2006

Abstract

In this paper, we propose a method of reducing topological defects in self-organizing maps (SOMs) using multiple scale neighborhood functions. The multiple scale neighborhood functions are inspired by multiple scale channels in the human visual system. To evaluate the proposed method, we applied it to the traveling salesman problem (TSP), and examined two indexes: the tour length of the solution and the number of kinks in the solution. Consequently, the two indexes are lower for the proposed method. These results indicate that our proposed method has the ability to reduce topological defects.

Key words:
Multiple scale neighborhood functions; Self-organizing maps; Topological defects; Traveling salesman problem

1 Introduction

The use of self-organizing maps (SOMs) (Kohonen, 2001) is a prominent unsupervised learning method, performing clustering and visualization of high-dimensional input data. Therefore, SOMs are widely used in various applications, such as speech recognition (Kohonen, 1988), motor control (Martinetz et al., 1990), and financial analysis (Deboeck and Kohonen, 1998). The success

* Corresponding author. phone: +81-532-44-6899; fax: +81-532-44-6873.
Email address: `mura@tutkie.tut.ac.jp` (Kazushi Murakoshi).

Article published in Biosystems, 90(1): 101-104, 2007. doi:10.1016/j.biosystems.2006.07.004

of these applications depends on how correctly the feature maps are topologically ordered. SOMs, however, have the problem of topological defects such as kinks in one-dimensional maps and twists in two-dimensional maps in limited learning time (Kohonen, 2001; Van Hulle, 2000).

Topological order is ensured by a neighborhood function, which defines how local activity determines the learning rate in its neighborhood. Kohonen (2001) interprets the learning process of a SOM as a summary of synaptic plasticity in the brain. The similarities between SOMs and brain maps suggest that the same principle might underlie the emergence of feature maps (Kohonen, 2001).

The human visual system has multiple scale processes to improve the efficiency of recognition (Wilson et al., 1983). The human brain is highly specialized and tuned to solve problems in the surroundings. Inspired by the multiple scale processes in the brain, we suggest applying these processes to the learning process of the SOM.

In this paper, we propose a method of reducing topological defects in SOMs using multiple scale neighborhood functions. To evaluate the proposed method, we applied it to the well-known traveling salesman problem (TSP) (Reinelt, 2005). The TSP is the problem of determining the shortest tour visiting all cities exactly once and returning to the starting point. We selected the TSP because it is easy to evaluate the reduction of topological defects by consideration of the indexes for the TSP.

In Section 2, we briefly introduce the conventional SOM, and propose a method for reducing topological defects in SOMs using multiple scale neighborhood functions. Section 3 shows the results of computer experiments to evaluate the proposed method. Section 4 concludes this paper.

2 Method using multiple scale neighborhood functions

A typical learning process of the SOM is as follows. For each node i, the reference vector m_i is updated as

$$m_i \leftarrow m_i + f \cdot (x - m_i),$$

where x is an input vector and the neighborhood function f has an essential role as a smoothing kernel. The neighborhood function f is frequently defined as the Gaussian function

$$f = K \exp\left(-\frac{d_{ic}^2}{\sigma^2}\right),$$

where K is the gain parameter, σ specifies the neighborhood region of the kernel, and d_{ic} is the distance between node i and the winning node c.

To solve the TSP, Angéniol et al. (1988) use a one-dimensional ring shaped SOM, where input x is the position of the city, the gain parameter K equals $1/\sqrt{2}$, and d_{ic} is the distance measured along the ring between nodes i and c. Moreover, the value of σ is decreased by

$$\sigma \leftarrow \alpha \cdot \sigma,$$

where the value of σ is decreased by the factor α ($0 < \alpha < 1$).

Inspired by the multiple scale processes in the human visual system (Wilson et al., 1983), we propose a method for reducing topological defects in SOMs using multiple scale neighborhood functions. The following multiple scale neighborhood functions are introduced,

$$f = \beta_j K \exp\left(-\frac{d_{ic}^2}{(\gamma_j \sigma)^2}\right),$$

where β_j and γ_j ($j = 1, \cdots, N$) are the gain and width factors of the j-th neighborhood function, respectively. On the basis of the psychophysical experiment conducted by Wilson et al. (1983), we set $N = 6$ and the other values as shown in Table 1. The gains for the narrowest and broadest neighborhood functions are relatively low, whereas the gains for the central neighborhood functions are relatively high. In our method, a neighborhood function is randomly chosen.

3 Computer experiments

To evaluate the proposed method, we applied both the proposed method and the method devised by Angéniol et al. (1988) (AVT method) to the TSP. The sets of data used for the computational experiments are TSP instances available on TSPLIB (Reinelt, 2005). All the selected problems are Euclidean instances of the TSP for simplicity. The parameters are set as follows: the decreasing factor specifying the kernel width α is 0.98, and the initial value specifying the kernel width σ is set at half the number of cities to cover all nodes.

To evaluate the reduction of topological defects, we examine two indexes: the tour length of the solution and the number of kinks in the solution. Table 2 shows the tour lengths of solutions obtained for the problem instances. The data is the average tour length of 100 trials for each instance and method. The values in parentheses are the percentage deviations from each optimal tour length. The average deviation obtained using our proposed method is 0.59 percent better than that obtained using the AVT method. Table 3 shows the numbers of kinks in solutions obtained for the problem instances. The data is the total number of kinks from 100 trials for each instance and method. The values in parentheses are the percentage deviations of the total number of kinks obtained using our proposed method from those obtained using the AVT method. The average number of kinks obtained using our proposed method is 42.65 percent less than that obtained using the AVT method.

As the number of kinks markedly decreases, we can show more information about the kinks. Figure 1 illustrates examples of solutions for the instance bier127. The solution shown in Fig. 1 (a) is obtained using the AVT method. This solution has two kinks: near the center and at the lower right. The solution shown in Fig. 1 (b) is obtained using our proposed method. This solution has no kinks. Figure 2 shows the number of kinks in each solution for the instance bier127. A noteworthy point is that the number of solutions without kinks increases significantly using the proposed method. For the other instances, similar results are observed.

From an other point of view, we evaluate the performance of the proposed method in terms of execution time. Table 4 shows the execution times for the obtained solutions. The data is the total execution time for 100 trials for each instance and method. The values in parentheses are the percentage deviations of the execution times obtained using our proposed method from those obtained using the AVT method. The average execution time obtained using our proposed method is 9.83 percent less than that obtained using the AVT method.

4 Conclusion

We propose a method for reducing topological defects in SOMs using multiple scale neighborhood functions. To evaluate the proposed method, we applied it to the TSP, and examined two indexes: the tour length of the solution and the number of kinks in the solution. Consequently, the two indexes are lower for the proposed method. These results indicate that our proposed method has the ability to reduce topological defects. Additionally, the average execution time obtained using our proposed method is less than that obtained using the conventional method.

Our proposed method is inspired by the multiple scale processes in the brain. The results obtained using our proposed method imply that the multiple scale processes in the brain are effective in the following two points. The first is in the reduction of topological defects. The second is in the rapid convergence of learning processes.

SOMs are widely used as a powerful tool. Therefore, in future studies, we will examine whether the proposed method is more generally effective.

References

Angéniol, B., de la Caroix Vaubois, G., and le Texier, J.-Y., 1988. Self-organizing feature maps and the travelling salesman problem. Neural Netw., 1, 289-293.

Deboeck, G., and Kohonen, T. 1998. Visual explorations in finance : with self-organizing maps. Berlin: Springer-Verlag.

Kohonen, T., 1988. The 'neural' phonetic typewriter. Computer, 21(3), 11-22.

Kohonen, T. 2001. Self-organizing maps (3rd edition). Berlin; New York: Springer-Verlag.

Martinetz, T., Ritter, H., and Schulten, K., 1990. Three-dimensional neural net for learning visuomotor coordination of a robot arm. IEEE Trans. Neural Netw., 1(1), 131-136.

Reinelt, G., 2005. Traveling salesman problem. (Retrieved in 2005 from http://www.iwr.uni-heidelberg.de/groups/como pt/software/TSPLIB95/)

Van Hulle, M. M. 2000. Faithful representations and topographic maps: From distortion- to information-based self-organization. New York: Wiley-Interscience.

Wilson, H. R., McFarlane, D. K., and Phillips, G. C., 1983. Spatial frequency tuning of orientation selective units estimated by oblique masking. Vision Res., 23(9), 873–882.

Table 1
Values of β_j and γ_j.

j	1	2	3	4	5	6
β_j	2^{-2}	2^{-1}	2^{0}	2^{-1}	2^{-2}	2^{-3}
γ_j	2^{-2}	2^{-1}	2^{0}	2^{1}	2^{2}	2^{3}

Table 2
Optimal tour length and comparison of solutions.

Instance	Optimum	AVT method	Proposed method
bier127	118282	124341.43 (5.12)	122783.32 (3.81)
kroD100	21294	22044.51 (3.52)	21946.00 (3.06)
lin318	42029	44668.68 (6.28)	44369.17 (5.57)
p654	34643	36701.64 (5.94)	36502.92 (5.37)
pcb442	54302	55980.50 (10.25)	55609.48 (9.51)
pcb1173	56892	62880.76 (10.53)	62928.32 (10.61)
rd100	7910	8346.54 (5.52)	8189.02 (3.53)
rd400	15281	16203.79 (6.04)	16215.56 (6.12)
tsp225	3919	4177.75 (6.60)	4150.01 (5.89)
pr439	107217	114063.81 (6.39)	114129.34 (6.45)
vm1084	239297	258431.39 (8.00)	257787.21 (7.73)
Average	—	— (6.74)	— (6.15)

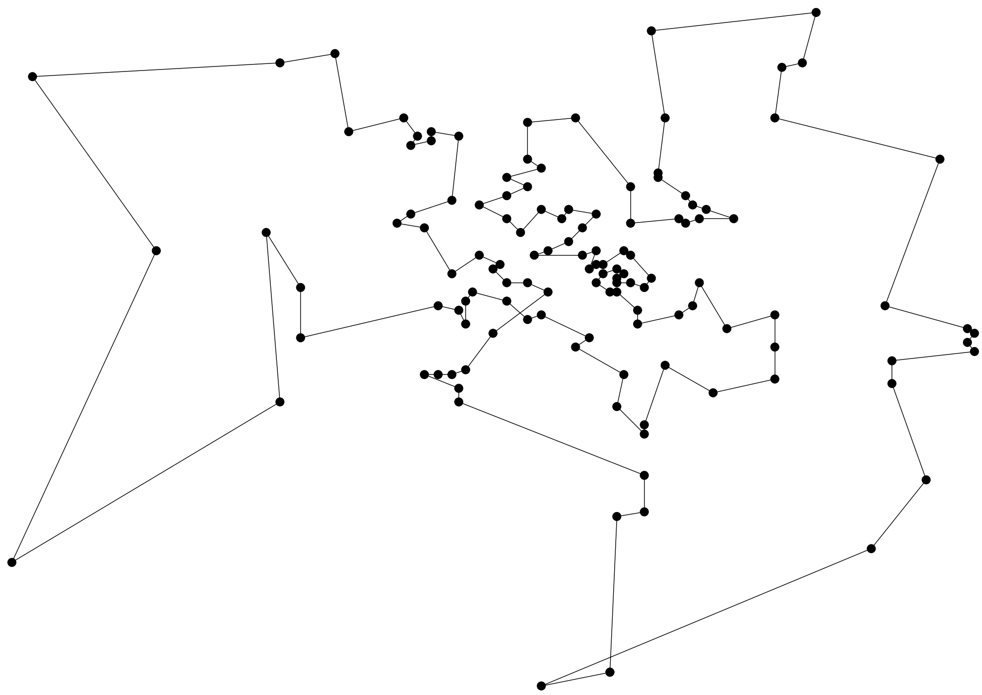
Table 3
Comparison of number of kinks.

Instance	AVT method	Proposed method
bier127	111	31 (−72.07)
kroD100	83	44 (−46.99)
lin318	451	289 (−59.20)
p654	434	248 (−42.86)
pcb442	435	221 (−49.20)
pcb1173	615	332 (−46.02)
rd100	42	24 (−42.86)
rd400	252	201 (−20.24)
tsp225	207	88 (−57.49)
pr439	472	319 (−32.42)
vm1084	1162	893 (−23.15)
Average	—	— (−42.65)

Table 4
Comparison of execution times (seconds).

Instance	AVT method	Proposed method
bier127	411.46	354.17 (−13.92)
kroD100	235.80	194.14 (−17.67)
lin318	3309.19	3207.29 (−3.08)
p654	15952.71	15364.27 (−3.69)
pcb442	6770.78	6366.33 (−5.27)
pcb1173	62928.32	56708.08 (−9.88)
rd100	233.46	176.91 (−24.22)
rd400	5465.58	5162.94 (−5.54)
tsp225	1477.40	1208.23 (−18.22)
pr439	6687.42	6364.00 (−4.84)
vm1084	48570.69	48041.31 (−1.09)
Average	—	— (−9.83)

(a)



(b)

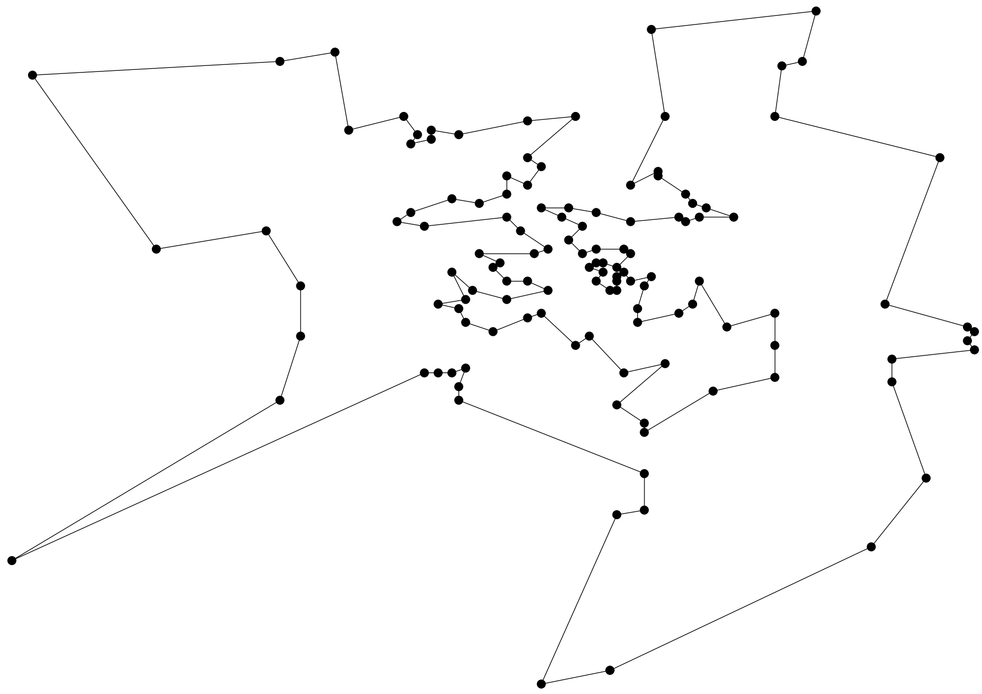


Fig. 1. Solutions for instance bier127. (a) Solution obtained using AVT method. This solution has two kinks. (b) Solution obtained using proposed method. This solution has no kinks.

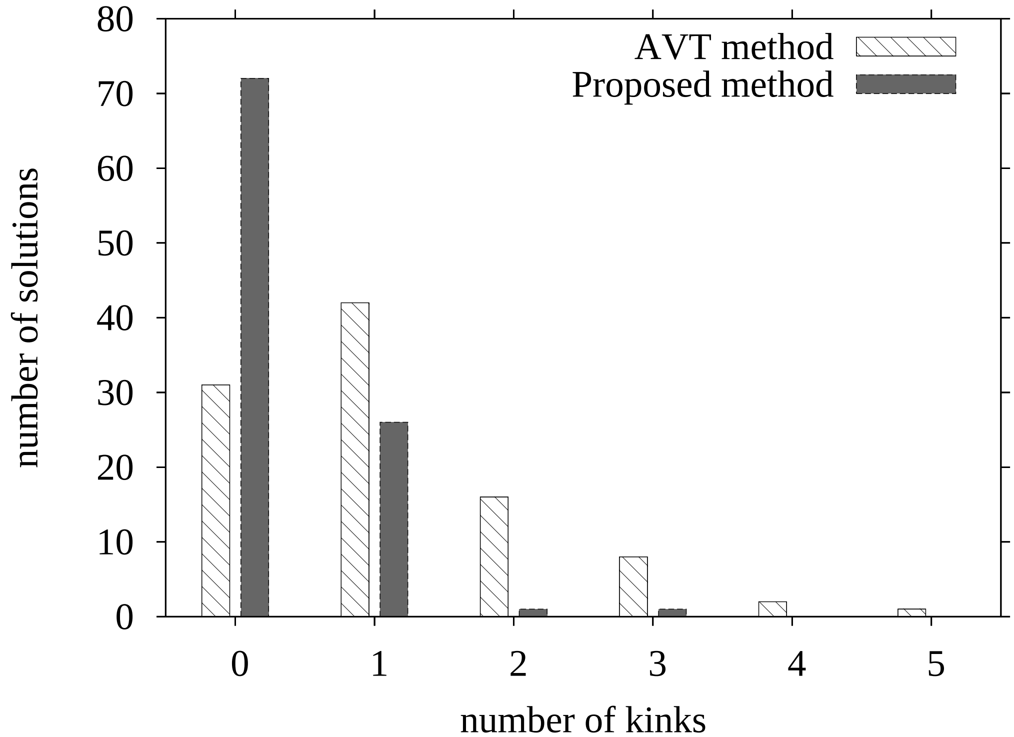


Fig. 2. Number of kinks in each solution for instance bier127.