New Method for Measuring the Topology Preservation of Self-organizing Feature Maps

MURASHIMA Sadayuki, KASHIMA Masayuki, and FUCHIDA Takayasu
Department of Information and Computer Science Faculty of Engineering
Kagoshima University
Koorimoto 1-21-40 Kagoshima Japan
mura@ics.kagoshima-u.ac.jp

Abstract

The topology preservation of self-organizing feature maps is an important property which is used in many applications. Various qualitative and quantitative approaches are known for measuring the degree of topology preservation. However, well received measure for determining the topology preservation are not presented yet. In this paper, we present new method for measuring the degree of topology preservation based on the masked Delaunay triangulation. The topology preservation is completed when the masked Delaunay triangulation coincides with the output network shape. We demonstrate the usefulness of this measure for various examples of data manifolds. This method is also applied to measure the degree of topology preservation of the topology representing network by Martinez and Schulten. Throughout various simulations this measure turned out to bring us reasonable degrees of topology preservation.

1 Introduction

The self-organization of a topology preserving map is had an interest in many persons as a central theme of the field of neural network and various algorithm [1, 2, 5] have been proposed. As for a topology preserving map various fields from the feature extraction to image processing are studied.

However a mathematical definition of topology preservation was unclear for long. Although many proposal regarding measures of topology preservation have been done by many researchers [3, 4, 5] even henceforth these definitions are based on the evaluations of the positions of neurons in the lattice and, on the other hand, on the evaluations of their weight vectors only.

To this algorithm that decides the topological structure with the form that reacts to the input called Topology Representing Network was proposed in 1994 by Martinez and Schujten [4]. They were attempting a rigid definition of the topology preservation in Topology Representing Network by using the masked Voro-
expressed adaptively in \( n \) reference vectors which are taken in the same space. The nearest reference vector \( w_i \ast (v) \) to input data \( v \) in Euclid distance is called the winner vector. This winner vector and plural vectors which are in the topological neighborhood of this winner vector change so that they approach to \( v \). The amount of modification of reference vector \( w_i \) is

\[
\Delta w_i = \varepsilon h_{i*}(v - w_i), \quad \text{for all } i \in A. \quad (1)
\]

The vicinity function \( h_{i*} \) is the topological neighborhood which decreases with time.

### 2.2 Topology Representing Network by Martinetz and Schulten

The procedure for constructing the connections between the unit \( i, j(i, j = 1, \ldots, N) \) can be formulated as follows:

1. **Assign initial values to the pointers** \( w_i \in R^D_i, i = 1, \ldots, N \) and set all connection strengths \( C_{ij} \) to zero.
2. **Select an input pattern** \( v \in M \) with equal probability for each \( v \).
3. **For each unit** \( i \) determine the number \( k_i \) of units \( j \) with

\[
\| v - w_j \| < \| v - w_i \|
\]

by, for example, determining the sequence \( (i_0, i_1, \ldots, i_{N-1}) \) with

\[
\| v - w_{i_0} \| < \| v - w_{i_1} \| < \cdots < \| v - w_{i_{N-1}} \|
\]

4. **Perform an adaptation step of the pointers** \( w_i \) according to the neural gas algorithm by setting

\[
w_i^{new} = w_i^{old} + \varepsilon e^{-k_i/\lambda} (v - w_i^{old}), \quad i = 1, \ldots, N;
\]

5. **If** \( C_{i_0i_1} = 0 \) \( \text{set} \ C_{i_0i_1} = 0 \) and \( t_{i_0i_1} = 0 \), that is, connect \( i_0 \) and \( i_1 \). If \( C_{i_0i_1} > 0 \) \( \text{set} \ t_{i_0i_1} = 0 \), that is refresh connection \( i_0 - i_1 \).
6. **Increase the age of all connection of** \( i_0 \) by setting \( t_{i_0j} = t_{i_0j} + 1 \) for all \( j \) with \( C_{i_0j} > 0 \);
7. **Remove those connections of unit** \( i_0 \) the age of which exceeds \( T \) by setting \( C_{i_0j} = 0 \) for all \( j \) with \( C_{i_0j} > 0 \) and \( t_{i_0j} > T \); continue with (2).

### 2.3 Topology preserving map

The widely accepted definition of topology preservation does not exist yet. We consider that the thought to the topology preservation given by Martinetz and Schulten is adequate in direction. On the basis of the lattice \( A \) that consists of \( n \) neurons a self-organizing feature map of Kohonen composes a topology preserving map \( M_A. \) To neuron \( i \) \((i \in A)\) reference vector \( w_i \in M \) is assigned. A map \( M_A = (\Psi_{M \rightarrow A}, \Psi_{A \rightarrow M}) \) is topology preserving if both the inverse mapping \( \Psi_{A \rightarrow M} \) from \( A \) to \( M \) and the mapping \( \Psi_{M \rightarrow A} \) from \( M \) to \( A \) are neighborhood preserving. These 2 maps are

\[
M_A = \begin{cases}
\Psi_{M \rightarrow A}; & v \in M \rightarrow \Psi^*(v) \in A \\
\Psi_{A \rightarrow M}; & i \in A \rightarrow w_i \in M.
\end{cases} \quad (2)
\]

Generally when the map \( \Psi_{M \rightarrow A} \) from \( M \) to \( A \) and similarly the inverse map \( \Psi_{A \rightarrow M} \) from \( A \) to \( M \) are both a vicinity preservation map, \( M_A \) is called a topology preservation. Accordingly these two vicinity preservation need to be calculated, to decide whether or not the self-organizing feature map is topology preserving. How much the topological relation between input and output is preserved rely to a selection of lattice structure \( A \) although it depends on the definition of a topology preservation measure. The lattice structure, i.e., 1 dimensional, 2 dimensional or the other type of lattice must be chosen as the supreme result is obtained according to the feature of input data space \( M \).

However in many applications, because the feature of \( M \) is not known, what kind of lattice structure should be chosen is not obvious in advance. It is said generally that it must decide which lattice shows the highest topology preservation rate observing results obtained by using the various lattice structures. However the measure that decides an extent of topology preservation as mentioned before have not been announced yet. But it is considered that the topology preservation is materialized automatically with the stage that self-organization finished and in many cases the rectangular lattice is used. Because so far it is considered that the Kohonen network with rectangular lattice is able to attain a topology preservation. A method like the Topology Representing Network in which the lattice structure of network is decided adaptively in self-organizing manner is more desirable than the method in which some lattice structure is fixed in advance like Kohonen network for a tool of feature extraction map.
3 New Method for Measuring the Topology Preservation

3.1 Vicinity relation of input data manifold

It is assumed that input data are restricted to the limited k dimensional space M. When Kohonen network that has n neurons is formed in self organizing manner, reference vector w_i is obtained. In the Kohonen network the neuron that has a reference vector most close to input data answers. The area of influence (receptive field) of each reference vector is conceivable the area where each neuron answers. In the field of data processing the figure of influence area is known as Voronoi figure. Each reference vector makes kernel point. We assumed that the set of reference vectors is given as follows:

\[ W = \{ w_i | i = 1, 2, \ldots, n \}. \]  

(3)

According to the set W of reference vectors, the whole space of k dimension are divided depending upon which reference vector is the closest. This is called Voronoi division. By using Euclid distance d(v, w_i) between point v and reference vector w_i of k dimensional space, Voronoi area V_i of the i-th reference vector is defined as follows:

\[ V_i = \{ v \in \mathbb{R}^k | d(v, w_i) \leq d(v, w_j) \}, \quad \text{for} \ j \neq i. \]  

(4)

The Delaunay figure is also very important in the field of data processing. The element of adjacent matrix C_ij is defined as follows:

\[ C_{ij} = \begin{cases} 1, & \text{for } V_i \cap V_j \neq \emptyset \\ 0, & \text{for } V_i \cap V_j = \emptyset \end{cases} \]  

(5)

We can not think with infinitely wide k dimensional space because input data are restricted in the limited space in Kohonen network actually. Accordingly the Voronoi figure of the top must be changed as follows:

\[ V_i^{(M)} = \{ v \in M | d(v, w_i) \leq d(v, w_j) \}, \quad \text{for} \ j \neq i. \]  

(6)

The Voronoi figure considered in the limited space M not in infinite space \( \mathbb{R}^k \) is called the masked Voronoi figure. Each masked Voronoi area is a closed set and makes complete division of input data manifold M as follows:

\[ M = \bigcup_{i=1}^{n} V_i^{(M)} \in \mathbb{R}^k. \]  

(7)

By connecting the kernel points whose Voronoi area are adjoining in the masked Voronoi figure, we can get a graph. This is called the masked Delaunay figure D_{M}. Connection relation between nodes i and j are given below:

\[ C_{ij}^{(M)} = \begin{cases} 1, & \text{for } V_i^{(M)} \cap V_j^{(M)} \neq \emptyset \\ 0, & \text{for } V_i^{(M)} \cap V_j^{(M)} = \emptyset \end{cases} \]  

(8)

This gives us adjacent relation of the masked Voronoi area (receptive field) making each reference vector a kernel point. If \( C_{ij}^{(M)} = 0 \) node i and j are not adjoining and if \( C_{ij}^{(M)} = 1 \) nodes i and j are adjoining. The masked Delaunay figure for the case of k = 2 is called the masked Delaunay triangulation.

3.2 Vicinity relation between nodes of output network

The element \( C_{ij}^{(A)} \) of the connection matrix of network A on the side of output is

\[ C_{ij}^{(A)} = \begin{cases} 1, & \text{for } i \text{ and } j \text{ are connected} \\ 0, & \text{for } i \text{ and } j \text{ are unconnected} \end{cases}. \]  

(9)

The digital distance \( d(A)(i, j) \) between nodes i and j of network A on the output side means the smallest number of connections necessary to pass in order to travel from node i to node j. This is same even in the case that the digital distance \( d(D_{M})(i, j) \) between nodes on the masked Delaunay figure is calculated.

3.3 Topology preservation by an agreement of topological structure of input and output

The primary information or feature of input data is given by values of reference vector which each neuron has. These values are determined in the process of self-organizing feature map. The secondary information about input data manifold is given by the vicinity relation between reference vectors. The masked Delaunay triangulation \( C_{ij}^{(D_{M})} \) expresses the compound of the primary and secondary information about input data. We propose a measure that we think a natural definition of topology preservation without unreasonableness. In this paper we adopt the idea that the topology preservation is completed when the masked Delaunay triangulation coincides with the shape of output network.

3.4 Topology preservation rate

According to the definition above, local topology preservation rates of a neuron i are given by

\[ \psi^+(i) = \frac{\# \{ j | d(A)(i, j) = 1; d(D_{M})(i, j) = 1 \}}{\# \{ j | d(D_{M})(i, j) = 1 \}}, \]  

(10)

and

\[ \psi^-(i) = \frac{\# \{ j | d(A)(i, j) = 1 \}}{\# \{ j | d(A)(i, j) = 1 \}}. \]  

(11)

Here \( \# \{ j | d(A)(i, j) = 1 \} \) is the number of j that fills d(A)(i, j) = 1. Namely \( \psi^+(i) \) is defined as the proportion of the node that fills d(A)(i, j) = 1 in the output network to the node that fills d(D_{M})(i, j) = 1 in a masked
Delaunay figure when node $i$ is fixed. From the definition the following holds:

$$0 \leq \psi^+(i), \psi^-(i) \leq 1.0.$$  \hspace{1cm} (12)

If $\psi^+(i)$ and $\psi^-(i)$ become 1.0 with all neurons we consider that the form of masked Delaunay figure of input data manifold and the form of output network topologically agree in other words the topology preservation is completed. When we want to check the topology preservation rate globally, average topology preservation rate $\Psi_{\text{mean}}$ is used.

$$\Psi^+_{\text{mean}} = \frac{\sum_{i=1}^{n} \psi^+(i)}{n}$$  \hspace{1cm} (13)

$$\Psi^-_{\text{mean}} = \frac{\sum_{i=1}^{n} \psi^-(i)}{n}$$  \hspace{1cm} (14)

The topology preservation rate proposed here is applicable to the topology representing network by Martinetz and Schulten straightforwardly.

4 Computer Simulation on the Self-Organizing Feature map

4.1 Calculation of a masked Delaunay triangulation

Carrying out simulation of several Kohonen networks the local topology preservation rate and also an average topology preservation rate were measured. For the calculation of the receptive field, the Voronoi division method proposed by Watanabe and Murashima [8] was used. Checking adjacent relation of each masked Voronoi area in the top the masked Delaunay triangulation was obtained.

4.2 Kohonen network with 1 dimensional lattice and 2 dimensional input

Input data manifold $M$ is assumed to be limited to the square area of length 10 and height 1 and the Kohonen network is assumed to have $n$ neurons connected in 1 dimensional chain. So far it is considered that the dimensional conflict like this does not yields a topology preservation. However in this paper we can show the case in which a topology preservation is obtained in spite of the dimensional conflict. The learning results to input data of uniform distribution are shown in Fig.1. In the cases of $n = 100$ and $n = 50$ the average topology preserving rates $\Psi^+_{\text{mean}} = 0.43$ and 0.74, respectively, but in the case of $n = 10$, $\Psi^+_{\text{mean}}$ becomes 1.0. For each case $\Psi^-_{\text{mean}} = 1.0$. This means that a topology preservation is completed for $n = 10$ but not for $n = 50$ and $n = 100$. This fact also teaches us that there is the maximum neuron numbers of Kohonen Network with 1 dimensional lattice which can attain the topology preservation in spite of the dimensional conflict [6].

4.3 Kohonen network with rectangular and hexagonal lattice

The variations of local topology preservation rate in the learning process of Kohonen network is shown for the case of the Kohonen network with rectangular lattice in Fig.2 and with the hexagonal lattice in Fig.3, respectively. In the early stage of self-organizing process ($t = 100$, $t = 1000$) the local topology preservation rates $\psi^+$ and $\psi^-$ are very low but as the process goes on they rise gradually. Finally they both reach to 1.00 for the case of hexagonal lattice. But for the case of rectangular lattice $\psi^-$ does not reach to 1.00 although $\psi^+$ goes to 1.00. Here the final stage of self-organizing process means that the number of iterations $t$ reaches to 45000.

The learning results of Kohonen network with rectangular lattice and hexagonal lattice for the two dimensional input manifold are compared in Fig.4 and Fig.5. (a) is masked Voronoi partition, (b) is masked Delaunay triangulation and (c) is network.

The average topology preserving rate $\Psi^+_{\text{mean}} = 0.70$, $\Psi^-_{\text{mean}} = 1.00$ and $\Psi^+_{\text{mean}} = \Psi^-_{\text{mean}} = 1.00$, respectively. The input data manifold is assumed to be limited to the square area with uniform distribution. For this type of input, the well known learning result of Kohonen network with rectangular lattice is considered to complete a topology preservation at the end of self organization of learning process. However the topology preserving rate proposed here teach us that it means only the end of self organization but not the completion of a topology preservation.

4.4 Feature extraction by Topology Representing Network

The learning results of Kohonen networks with rectangular and hexagonal lattice and topology representing network for two dimensional and non-linear input data manifold are shown in Fig.6, Fig.7 and Fig.8, respectively. The input data are assumed to be limited to a square area but is not uniform. There are two big circles in the square and input data do not fall on the two circles. Three figures mean (a) masked Voronoi partition, (b) masked Delaunay triangulation and (c) output network, respectively. In the learning result of both Kohonen networks some reference vectors fall inside the two circles where no input data come. On the contrary in the case of topology representing network there is no reference vector inside the two circles. This means that the ability of vector quantization of Kohonen network
has some problems. The average topology preservation rates of two Kohonen networks with different type of lattice are $\Psi_{\text{mean}}^+ = 0.55$, $\Psi_{\text{mean}} = 0.98$ and $\Psi_{\text{mean}}^- = 0.94$, $\Psi_{\text{mean}}^- = 0.91$, respectively. The average topology preservation rates of TRN are $\Psi_{\text{mean}}^+ = 0.93$ and $\Psi_{\text{mean}}^- = 1.00$.

As a tool of feature extraction we have to check the degree of vector quantization as well as the degree of topological preservation. The average topology preservation rates above teach us that the degree of the topology preservation attained by TRN is higher than that attained by two type of Kohonen networks above. We can say that the degrees of topology preservation obtained by our measure produce reasonable results.

4.5 Feature extraction of characters by attaining the topology preservation

We have reached the stage that we can show the examples of feature extraction of characters by attaining the topology preservation between input and output. Fig. 1 teach us that the topology preservation between input and output by Kohonen network with 1 dimensional lattice is attainable if the two dimensional input is restricted in very narrow area. We applied the Self-organizing feature maps to the input character data "C" and "Y". The extracted data are shown in Fig.9. In these figures (a) are extracted features of characters for the case of Kohonen network with 1 dimensional lattice of 15 neurons and (b) are for the case of Topology Representing Network of 15 neurons,respectively. The perfect topology preservations have been obtained for all cases except for the case of Kohonen network applied to the character "Y". Only in this case $\Psi_{\text{mean}}^+ = 0.88$, $\Psi_{\text{mean}}^- = 0.87$. This means that the perfect topology preservation is not attainable because of the topological conflict in shape between the 1-dimensional lattice and the character "Y". In the other cases, $\Psi_{\text{mean}}^+ = \Psi_{\text{mean}}^- = 1.0$. These experimental results teach us that the Topology Representing Network is more suitable to extract the topological feature of characters by using the topology preservation [9] than Kohonen network with 1 dimensional lattice.

5 Conclusion

A topology preservation measure of a self organizing feature map was proposed. It judges that topology preservation is completed when the lattice structure of output neural network and the masked Delaunay triangulation that means the extracted feature of input data manifold agree. Several topology preserving rates are measured for various self organizing feature map or the topology representing network. Throughout various simulations this measure turned out to be very reasonable as a measure of topology preservation.

References

Figure 2: Local topology preservation rate in Learning process of Kohonen network with rectangular lattice for two dimensional data manifold with uniform distribution.


Figure 3: Local topology Preservation rate in Learning process of Kohonen network with hexagonal lattice for two dimensional data manifold with uniform distribution.

Figure 4: Learning results of Kohonen net with rectangular lattice for two dimensional data manifold with uniform distribution. $\psi^+_\text{mean} = 0.70$, $\psi^-\text{mean} = 1.00$

Figure 5: Learning results of Kohonen network with hexagonal lattice for two dimensional data manifold with uniform distribution. $\psi^+_\text{mean} = \psi^-\text{mean} = 1.00$
Figure 6: Learning results of Kohonen network with rectangular lattice for two dimensional and non-linear input data manifold. $\psi^+_{mean} = 0.55$, $\psi^-_{mean} = 0.98$

Figure 7: Learning results of Kohonen network with hexagonal lattice for two dimensional and non-linear input data manifold. $\psi^+_{mean} = 0.94$, $\psi^-_{mean} = 0.91$

Figure 8: Learning results of Topology Representing Network for two dimensional and non-linear input data manifold. $\psi^+_{mean} = 0.93$, $\psi^-_{mean} = 1.00$

Figure 9: Extracted features of characters by (a) Kohonen network with 1-dimensional lattice and (b) TRN for the characters "C" and "Y", respectively.