ELSEVIER

Artificial Intelligence 116 (2000) 265–286

Artificial Intelligence

www.elsevier.com/locate/artint

Conformality in the self-organization network ☆

Cheng-Yuan Liou *, Wen-Pin Tai

Department of Computer Science and Information Engineering, National Taiwan University, Taipei, 10636, Taiwan, ROC

Received 17 May 1999

Abstract

The conformality of the self-organizing network is studied in this work. We use multi-dimensional deformation analyses to interpret the self-organizing mapping. It can be shown that this mapping is quasi-conformal with a convergent deformation bound. Based on analyses, a deformation measure and a non-conformality measure are derived to indicate the evolution status of the network. These measures can serve as new criteria to evolve the network. We test these measures with simulations on surface mapping problems. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Conformality; Self-organization network; Natural coordinate; Data manifold; Deformation measure; Simplices; Learning algorithm; Pattern recognition

1. Introduction

The development of topological mapping of sensory input using the self-organization network has been analyzed in different ways. Much effort [5,8,11,14,17,21] has been expended in attempts to derive exquisite mapping dynamics based on certain energy functions. Bishop et al. [3] introduce a generative topographic mapping with an alternative foundation. Various mappings have been obtained with varying degrees of success. These energy functions are pre-defined disorder costs or measures. Many of them are realized as locally weighted mean mismappings. The issue of mapping deformation has been rarely discussed.

It has been observed that a self-organizing network prefers approximately conformal mapping [19] during its synapse adaptation (see Figs. 1(a)–(f)). This means that the

☆ This work has been supported by National Science Council under contract number NSC-82-0408-E-002-255 and NSC-84-2213-E-002-012.

* Corresponding author. Email: cyliou@csie.ntu.edu.tw.

0004-3702/00/$ – see front matter
PII: S0004-3702(99)00093-4

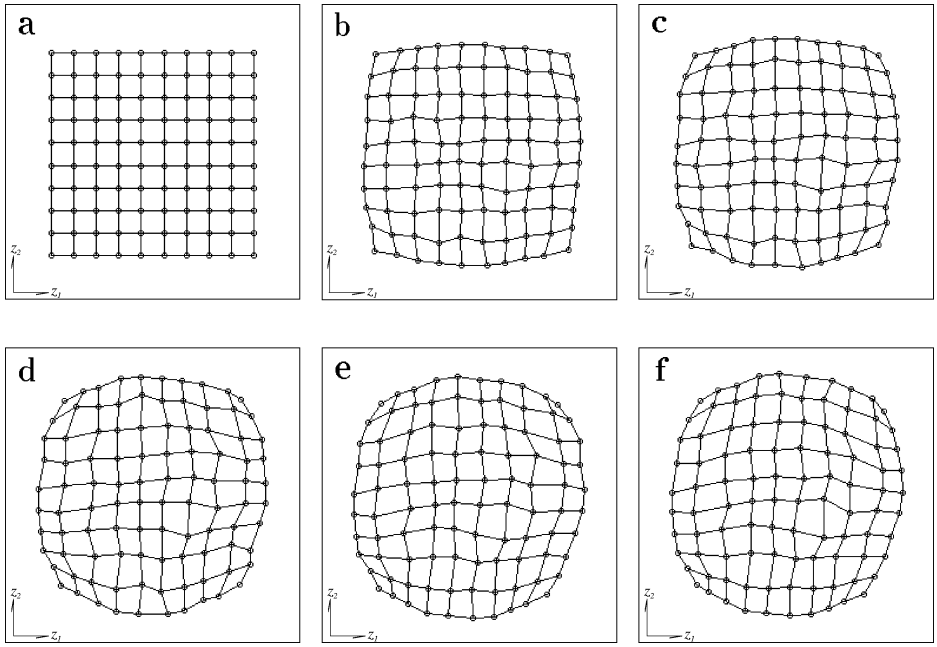


Fig. 1. Observations of the approximately conformal property during consecutive synapse mapping. We used the formal self-organization algorithm to obtain the results in (a)–(f). The inputs are uniformly distributed within a circle as shown in part (f). The neurons are on the grid nodes of a square as shown in part (a). The initial assignments for the synapse vectors are the coordinates of the neurons' positions shown in part (a). After training over 1000 random input samples, we plotted the results.

transformation of all corresponding quadrilaterals in each consecutive mapping remains as conformal as possible. We will show that a self-organizing mapping is quasi-conformal. This conformality (see, for example, [1]), which has received less attention, will be studied in this work. The issues which affect network performance, such as the learning rate, mapping status, and convergence results, can be properly related to the conformality.

We apply multi-dimensional deformation analyses to the network. A deformation measure and a non-conformality measure will be derived for the mapping. These measures can be used to indicate the evolution status of the network. We will derive such measures for the local deformation of the mapping. The crumples of the measure values show the mapping distortion locations. More important, one basic property of the self-organization, the locally isotropic development can be well characterized by these measures.

Based on the conformal mapping assumption, Tanaka [19] investigated the synaptic vector density. In his work, this assumption was used in the mapping from the network field to the same dimensional input space. So far there has been no theoretical evidence which justifies this assumption. Instead, we will study the approximately conformal property in consecutive synapse mapping using both theoretical analyses and simulations.

In our studies, the network configuration has been regarded as a collection of disjoint simplices which are deformed by adaptation of the synapses. It will be shown in a

later section that the deformation from current synapses to their adapted synapses is convergently bounded for each input stimulus; i.e., specifically, the mapping is quasi-conformal. When the bound converges toward its limit, which is equal to one, the synapse vectors adapt to match the input distribution. This bound reach its limit whenever perfect matching to inputs is obtained.

Both the deformation measure and non-conformality measure are bounded and are minimized when the input distribution matches the network. With these measures, appropriate network configurations can be designed according to how close the limit is approached. New dynamics can be derived for these measures to evolve the network. The cumulative measures for consecutive self-organizing mappings can be applied to pattern matching problems, such as surface mapping problems. They are suitable for measuring local mismatches between similar images. Simulations and comparisons will be presented to show the performance of the measures.

In the next section, the formal self-organization model will be reviewed in order to define the notations, including concise formulas, network configurations, and important parameters. Investigation into the mapping using multi-dimensional deformation analyses is included in Section 3. The derivation of the deformation measure and the non-conformality measure are included in this section. Following the mathematical analyses, several properties of the mapping will be studied in Section 4. In Section 5, the averaged non-conformality measure for all neurons will be used to show the evolution of a whole network. We include two other averaged measures, the mean square error (MSE) and the degree of topographicity (TPG) [4], for comparison. Applications of the proposed method will be presented in Section 6. Simulations show that the mismatches in surface mapping problems can be suitably located using these measures. Finally, we will provide discussion and suggest extension of this work in Section 7.

2. The network

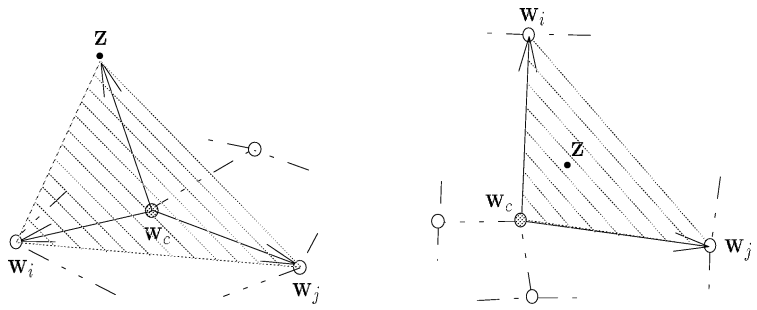
Let the set of input data in the training process be Z, and let $Z \subseteq \mathbb{R}^p$. The self-organization model employs a set of neurons, generally arranged in a network with dimension q, to process the samples from Z. Let the synapse vector set be $W \subseteq \mathbb{R}^p$. By iterative adaptation, the network can be evolved to transform the high-dimensional input space into a low-dimensional network space. In each evolutionary step, an input $\boldsymbol{z} \in Z$ is selected randomly. The model determines the winning or best-matching neuron, evolution criterion, such that

$$\|\boldsymbol{w}_c - \boldsymbol{z}\| = \min_i \|\boldsymbol{w}_i - \boldsymbol{z}\|, \quad \boldsymbol{w}_i \in W. \tag{1}$$

Let the vector $\boldsymbol{w}_c$ denote the synapse vector of the winning neuron for the corresponding input $\boldsymbol{z}$. Then, the synapse vectors of all the neurons in the network are updated with different values, according to the following equation:

$$\Delta \boldsymbol{w}_i = \alpha H\big(D(c,i)\big)(\boldsymbol{z} - \boldsymbol{w}_i), \tag{2}$$

where the parameter $\alpha \in [0,1)$ is the adaptation rate and H is the neighborhood function which decreases monotonously with the distance metric D in the network coordinate.



(a) (b)

Fig. 2. The simplicial representation of the self-organizing mapping, where $\boldsymbol{w}_i$, $\boldsymbol{w}_j$, $\boldsymbol{w}_c$, and $\boldsymbol{z}$ denote the synapse vectors and the input in the simplex, respectively. (a) The simplex is formed by the vectors $\boldsymbol{w}_i - \boldsymbol{w}_c$, $\boldsymbol{w}_j - \boldsymbol{w}_c$, and $\boldsymbol{z} - \boldsymbol{w}_c$ for $p = 3$, $q = 2$, and $n = 3$. (b) The simplex is formed by the vectors $\boldsymbol{w}_i - \boldsymbol{w}_c$ and $\boldsymbol{w}_j - \boldsymbol{w}_c$ for $p = 2$, $q = 2$, and $n = 2$.

For the input $\boldsymbol{z}$ and the synapse vector $\boldsymbol{w}_i$, we will adopt an approximated neighborhood function $h(d(\boldsymbol{z}, \boldsymbol{w}_i))$, which decreases monotonously with the distance metric d in the input space, and in which $h(d(\boldsymbol{z}, \boldsymbol{w}_c)) = 1$. This is found in some similar work by Erwin et al. [6]. This kind function can avoid the discontinuity from the derivation of the energy functions.

Further, we need to consider the topological representation of the synapse vectors. The topology formed by the self-organization can be regarded as a collection of disjoint n-dimensional simplices [10], where $n = \min\{(q + 1), p\}$. In most cases, p is much larger than $q + 1$. We assume that $n = q + 1$ in order to focus on the network configuration and to simplify the expressions. Each simplex is generated by the $(q + 1)$ vectors, obtained by subtracting the synapse vector of the winning neuron from the input vector and from the synapse vectors of any q neighboring neurons. In cases where p is smaller than $q + 1$, any p vectors obtained in a similar way can compose a simplex. Examples of synapse adaptation in simplicial representation are illustrated in Fig. 2.

When we change the coordinate to the n-dimensional simplicial system with the synapse vector of the winning neuron as the origin, the updating equations are reformulated as follows. Let $\boldsymbol{z}$, $\boldsymbol{w}_c$, and $\boldsymbol{w}$ be the input, the synapse vector of the winning neuron, and one of the q synapse vectors of the neighboring neurons in the simplex, respectively. By updating rule, we have

$$\boldsymbol{w}_c^{\text{new}} = \boldsymbol{w}_c + \alpha(\boldsymbol{z} - \boldsymbol{w}_c) \quad \text{and} \quad \boldsymbol{w}^{\text{new}} = \boldsymbol{w} + \alpha h\big(d(\boldsymbol{z}, \boldsymbol{w})\big)(\boldsymbol{z} - \boldsymbol{w}). \tag{3}$$

Let $\boldsymbol{v} = \boldsymbol{w} - \boldsymbol{w}_c$, $\boldsymbol{x} = \boldsymbol{z} - \boldsymbol{w}_c$, and let h denote $h(d(\boldsymbol{x}, \boldsymbol{v})) = h(d(\boldsymbol{z}, \boldsymbol{w}))$. We will call $\boldsymbol{v}$ the relative synapse vector and $\boldsymbol{x}$ the relative input. The update of $\boldsymbol{v}$, $\Delta\boldsymbol{v}$, in the simplex is

$$\begin{aligned}\Delta\boldsymbol{v} &= \boldsymbol{v}^{\text{new}} - \boldsymbol{v} \\ &= \big(\boldsymbol{w}^{\text{new}} - \boldsymbol{w}_c^{\text{new}}\big) - (\boldsymbol{w} - \boldsymbol{w}_c) \\ &= \alpha(h - 1)\boldsymbol{x} - \alpha h\boldsymbol{v}.\end{aligned} \tag{4}$$

Furthermore, for computational efficiency, we can formulate Eq. (4) in the low-dimensional simplicial coordinate. If the simplex region does not degenerate, it is easy to find a set of coordinate bases for the relative input vector and the q relative synapse vectors using, for example, the Gram–Schmidt method. But if it does degenerate, additional unit vectors which are orthogonal to the simplex space can be included. With these new bases, any vector in the simplex will have a simpler form.

This simplicial representation is in accordance with the description of the energy system. Erwin et al. [6] have shown that there does not exist a global energy function in general. A set of energy functions for neurons, formulated using the average distribution errors, has been used to obtain the dynamics of the network based on the method of stochastic approximation. The energy function for the neuron i is

$$\mathcal{E}_i \equiv \varepsilon \int \widehat{H}(H, W)\tfrac{1}{2}(\boldsymbol{z} - \boldsymbol{w}_i)^2 P(\boldsymbol{z})\,\mathrm{d}^p\boldsymbol{z} + X_i, \tag{5}$$

where $\widehat{H}$ is the generalized neighborhood function [6], $P(\boldsymbol{z})$ is the probability density of input $\boldsymbol{z}$, and X_i is a potential term for the neuron i, which corresponds to the network configurations during the adaptation process. When the training inputs are presented in a stochastic way, the sample energy function $\mathcal{E} = \widehat{H}(H, W)(\boldsymbol{z} - \boldsymbol{w}_i)^2$ for $\mathcal{E}_i$ in (5) will be considered for the purpose of approximation [18].

3. The deformation measure and the non-conformality measure

To formulate the new measures, we use the above simplicial expression in the $n = \min\{(q+1), p\}$ dimensional space. For each relative input vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_i, \ldots, x_n)$ and each relative synapse vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_i, \ldots, v_n)$ within the simplex, the mapping function $\boldsymbol{f}$, where $\boldsymbol{v}^{\text{new}} = \boldsymbol{f}(\boldsymbol{v})$ as in Eq. (4), is defined as

$$\boldsymbol{f}(\boldsymbol{v}) = \alpha(h-1)\boldsymbol{x} + (1 - \alpha h)\boldsymbol{v}. \tag{6}$$

Note that the mapping $\boldsymbol{f}$ is defined at any vector $\boldsymbol{v}$ with finite $\boldsymbol{f}(\boldsymbol{v})$. Furthermore, the vector $\boldsymbol{f}(\boldsymbol{x})$, which is equal to $(1-\alpha)\boldsymbol{x}$, can be regarded as the mapping of $\boldsymbol{x}$ in the simplicial coordinate.

Let us derive the mapping deformation for $\boldsymbol{f}$. For a given simplex, when the relative synapse vectors are updated by Eq. (6), the function $\boldsymbol{f}$ can be regarded as a mapping from the simplex region to the corresponding one which is formed by the updated synapse vectors. By examining the conformality condition of this mapping function, we can derive a deformation measure and a non-conformality measure from the Jacobian matrix of such mapping. These measures have accurate bounds.

3.1. The Jacobian matrix

Let the function $\boldsymbol{f}(\boldsymbol{v})$ be $(f_1, f_2, \ldots, f_i, \ldots, f_n)$, and let each component f_i be a function of $\boldsymbol{v} = (v_1, \ldots, v_j, \ldots, v_n)$. For any $\boldsymbol{v}$ at which $\boldsymbol{f}$ is defined and the Jacobian $J(\boldsymbol{v}, \boldsymbol{f})$, the determinant of the Jacobian matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$, is greater than zero, the deformation of the mapping $\boldsymbol{f}$ can be defined. To further analyze the deformation, we need to derive the explicit form of the Jacobian matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$ for the mapping.

Given a simplex region in the self-organizing network, the mapping of relative synapse vectors is defined as in (6). To evaluate the matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$, we focus on each component f_i of $\boldsymbol{f}$, i.e.,

$$f_i(\boldsymbol{v}) = \alpha(h-1)x_i + (1-\alpha h)v_i, \quad i = 1, \ldots, n, \tag{7}$$

where $h = h(d(\boldsymbol{x}, \boldsymbol{v}))$. We will use the Euclidean distance metric for d in the simplicial coordinate, i.e., $d = \sum_{k=1}^{n} (x_k - v_k)^2$. Hence, we have the partial derivatives of $\boldsymbol{f}$, for $1 \leqslant i, j \leqslant n$, $j \neq i$:

$$\begin{aligned} \frac{\partial f_i}{\partial v_i} &= -2\alpha h'(x_i - v_i)^2 + (1-\alpha h), \\ \frac{\partial f_i}{\partial v_j} &= -2\alpha h'(x_i - v_i)(x_j - v_j), \end{aligned} \tag{8}$$

where h' denotes $\mathrm{d}h/\mathrm{d}d$. The value of $\mathrm{d}h/\mathrm{d}d$ is always non-positive for a monotonous decreasing function h, i.e., $h' \leqslant 0$. Note the matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$ can be written in a compact vector form. The eigenspace structure of this matrix will be obvious from this form. We will use the expanded form in all the following derivations as we did in earlier efforts.

Because $\partial f_i/\partial v_j = \partial f_j/\partial v_i$ for $1 \leqslant i, j \leqslant n$, the matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$ is symmetric. Every eigenvalue of $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$ is real. The results given in Appendix A show that the matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$ has two eigenvalues:

$$\lambda_1 = 1 - \alpha h \quad \text{and} \quad \lambda_2 = -2\alpha h' \sum_{i=1}^{n} (x_i - v_i)^2 + (1-\alpha h), \tag{9}$$

with multiplicities $n-1$ and 1, respectively.

3.2. The deformation measure

The deformation of the mapping $\boldsymbol{f}(\boldsymbol{v})$ is defined by the *deformation measure*

$$Q \equiv \sqrt{(e_{\max}/e_{\min})}, \tag{10}$$

where $e_{\max} > 0$ and $e_{\min} > 0$ are the maximal and minimal eigenvalues of the Jacobian matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$, respectively [15]. The quantity Q provides a good evaluation for measuring the deformation of the mapping function $\boldsymbol{f}$.

To evaluate the deformation measure Q, we examine the Jacobian matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$ of the network. There are two distinct eigenvalues, λ_1 and λ_2 in (9), of the matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$. For $\alpha \in [0, 1)$ and $h' \leqslant 0$,

$$\lambda_2 - \lambda_1 = -2\alpha h' \sum_{i=1}^{n} (x_i - v_i)^2 \geqslant 0. \tag{11}$$

Thus, the deformation measure for the network will be

$$Q = \left(\frac{\lambda_2}{\lambda_1}\right)^{1/2} = \left(\frac{-2\alpha h'}{(1-\alpha h)} \sum_{i=1}^{n} (x_i - v_i)^2 + 1\right)^{1/2}. \tag{12}$$

The value of Q is no less than 1, i.e., $Q \geqslant 1$. If $Q = 1$ in (12), the measure Q will indicate that there is no deformation in the mapping $\boldsymbol{f}$.

3.3. The non-conformality measure

We now study the elegant conformal and quasi-conformal mappings. This kind mapping is defined as a topological mapping which satisfies certain geometric restrictions. Reshetnyak [15] proposed conditions for both conformal mapping and quasi-conformal mapping. First, let us consider the conformal mapping. For any $\boldsymbol{v}$ at which $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$ is defined and $J(\boldsymbol{v}, \boldsymbol{f}) > 0$, the inequality

$$n^{n/2} J(\boldsymbol{v}, \boldsymbol{f}) \leqslant \left(\sum_{i=1}^{n} \left\| \nabla f_i(\boldsymbol{v}) \right\|^2 \right)^{n/2} \tag{13}$$

holds. The function $\boldsymbol{f}$ is conformal if and only if the sign of equality holds.

The geometric interpretation of the inequality (13) will give us the sense of this criterion. On the left-hand side, the term $J(\boldsymbol{v}, \boldsymbol{f})$ is the volume of the hyper-parallelepiped determined by the vectors $\nabla f_i(\boldsymbol{v}),\ i = 1, \ldots, n$. On the other side, the term

$$\left(\sum_{i=1}^{n} \left\| \nabla f_i(\boldsymbol{v}) \right\|^2 \right)^{1/2}$$

is the length of the diagonal in the hypercube formed by the n orthogonal vectors with length $\|\nabla f_i(\boldsymbol{v})\|,\ i = 1, \ldots, n$, and $(1/n^{n/2})(\sum_{i=1}^{n} \|\nabla f_i(\boldsymbol{v})\|^2)^{n/2}$ is the maximum volume of the hypercube inside a hypersphere which has a diameter with length $(\sum_{i=1}^{n} \|\nabla f_i(\boldsymbol{v})\|^2)^{1/2}$. When the mapping $\boldsymbol{f}$ is conformal, the vector set $\{\nabla f_i(\boldsymbol{v}) \mid i = 1, \ldots, n\}$ will be an orthonormal set. The sign of equality in (13) will hold.

Further, let us focus on the less rigid condition, quasi-conformality. Quasi-conformal mapping is a natural generalization of conformal mapping. If the inequality

$$\left(\sum_{i=1}^{n} \left\| \nabla f_i(\boldsymbol{v}) \right\|^2 \right)^{n/2} \leqslant n^{n/2} k J(\boldsymbol{v}, \boldsymbol{f}) \tag{14}$$

holds for a finite k, then the function $\boldsymbol{f}$ is quasi-conformal with the deformation bound k [15].

We then study the measure of bound K $(k \geqslant K)$ for mapping deformation from the inequality (14) and then have

$$K \equiv \frac{\left(\sum_{i=1}^{n} \|\nabla f_i(\boldsymbol{v})\|^2 \right)^{n/2}}{n^{n/2} J(\boldsymbol{v}, \boldsymbol{f})}. \tag{15}$$

Note that $K = 1$ implies that the sign of equality holds in the inequalities (13) and (14); i.e., the mapping $\boldsymbol{f}$ is conformal.

For any function $\boldsymbol{f}$ where $J(\boldsymbol{v}, \boldsymbol{f}) > 0$, the measure K is always greater than 1. This can be derived by the Cauchy's and the Hadamard's inequalities successively [16]. Thus, the value of K can be considered to be a measure of deformation, i.e., a measure of non-conformality. The larger the value is, the heavier the deformation is in the mapping, and vice versa.

By (8), the terms in (15) can be calculated as follows:

$$\sum_{i=1}^{n} \left\| \nabla f_i(\boldsymbol{v}) \right\|^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial v_j} \right)^2$$

$$= \left((-2\alpha h') \sum_{i=1}^{n} (x_i - v_i)^2 + (1 - \alpha h) \right)^2 + (n-1)(1-\alpha h)^2, \quad (16)$$

and

$$J(\boldsymbol{v}, \boldsymbol{f}) = (-2\alpha h')(1-\alpha h)^{n-1} \sum_{i=1}^{n} (x_i - v_i)^2 + (1-\alpha h)^n. \quad (17)$$

The derivation of (16) and (17) is given in Appendix B.

Finally, we have the *non-conformality measure* K from (15):

$$K \equiv \frac{\left[(-2\alpha h' \|\boldsymbol{x} - \boldsymbol{v}\|^2 + 1 - \alpha h)^2 + (n-1)(1-\alpha h)^2 \right]^{n/2}}{n^{n/2}(-2\alpha h' \|\boldsymbol{x} - \boldsymbol{v}\|^2 + 1 - \alpha h)(1-\alpha h)^{n-1}}. \quad (18)$$

We also introduce here a *deformation potential* E as $(\sum_{i=1}^{n} \|\nabla f_i(\boldsymbol{v})\|^2)^{n/2} - n^{n/2} J(\boldsymbol{v}, \boldsymbol{f})$, i.e.,

$$E \equiv \left[\left(-2\alpha h' \|\boldsymbol{x} - \boldsymbol{v}\|^2 + 1 - \alpha h \right)^2 + (n-1)(1-\alpha h)^2 \right]^{n/2}$$
$$- n^{n/2} \left(-2\alpha h' \|\boldsymbol{x} - \boldsymbol{v}\|^2 + 1 - \alpha h \right)(1-\alpha h)^{n-1}. \quad (19)$$

The non-conformality measure K in (18) will overflow when its denominator (18) is equal or close to zero. This overflow cannot be expected at all in general, unless there are infinitely many reference vectors. The potential E can be well applied to measure the non-conformality of the mapping. In many cases, E is simple to manipulate in further dynamic analyses, and the results obtained using E are similar to those obtained using K.

Using (13), it is easy to check that the value of the numerator will not be smaller than that of the denominator, which must be positive, in (18); i.e., $K \geqslant 1$ (or $E \geqslant 0$) under the condition where $0 \leqslant \alpha, h < 1$ and $h' \leqslant 0$.

4. Properties of conformality

In this section, we will discuss the measures obtained during convergence of the mapping and show some important properties on the conformality of the network. All measures Q, K, and E show similar behaviors from our experiences. We will focus on the non-conformality measure K in our discussion.

Since the relative expressions in the simplicial coordinate are adopted for both inputs and synapse vectors, we may wish to find similar expressions for K in the input coordinate. It is easy to show that

$$\|\boldsymbol{x} - \boldsymbol{v}\|_n = \left\| (\boldsymbol{z} - \boldsymbol{w}_c) - (\boldsymbol{w} - \boldsymbol{w}_c) \right\|_p = \|\boldsymbol{z} - \boldsymbol{w}\|_p. \quad (20)$$

This relationship provides an effective way to measure K in the input coordinate. The value of $K(\boldsymbol{z}, \boldsymbol{w})$ can be obtained by substituting $\boldsymbol{z}$, $\boldsymbol{w}$ and H for $\boldsymbol{x}$, $\boldsymbol{z}$ and h respectively, in (18).

4.1. Convergence of the measures

With (20), we can study the measure $K(\boldsymbol{z}, \boldsymbol{w})$. When the self-organization is in its final state of equilibrium, it will provide a good approximation of the input, and the average distribution errors will be minimized. If $\|\boldsymbol{z} - \boldsymbol{w}\|$ approaches 0 in the final state, we have

$$K \approx \frac{[n(1-\alpha h)^2]^{n/2}}{n^{n/2}(1-\alpha h)^n} = 1, \quad \text{or} \quad E \approx 0 \tag{21}$$

by (18), (19), and (20). In most applications, the minimum is not reached (guaranteed) because of the deformation in the mapping. Our experiences show that this measure tends to reduce its value to a certain equilibrium during self-organization evolutions. Whenever such an equilibrium is reached, we say it is the convergence of the evolution.

From the properties of the measure $K(\boldsymbol{z}, \boldsymbol{w})$, the mapping can be said to be quasi-conformal with a bound which has a minimum value of one. When a perfect approximation of the input distribution is obtained by means of synapse adaptation, the mapping will be conformal, i.e., $K = 1$, and there will be no topological deformation in the mapping.

The training parameters in the self-organization process may affect the non-conformality measure. When the learning rate α is close to zero, the measure will approach one. This condition means that the adaptation of the synapse vectors is small so that the mapping is nearly conformal. Furthermore, when $h' \approx 0$, the non-conformality measure $K \approx [n(1-\alpha h)^2]^{n/2}/[(n^{n/2}(1-\alpha h)^n] = 1$ for any fixed learning rate α. This is because the extent of adaptation remains constant for the whole simplex; i.e., scaling of the simplex takes place. The conformality will be maintained by means of this scaling, and this is shown in Fig. 3.

Figs. 4(a) and (b) show values of the measure K as a function of $\|\boldsymbol{z} - \boldsymbol{w}\|$ with different network parameters α and β. The potential E has a behavior similar to that of K. In all cases, the neighborhood function h is defined as $h(d) = \exp(-d^2/\beta^2)$ in normalized distance scales, i.e., $d \in [0, 1]$. We can see that the measure K approaches 1 when $\|\boldsymbol{z} - \boldsymbol{w}\| \approx 0$. When $\|\boldsymbol{z} - \boldsymbol{w}\| \approx 1$, the measure K is also close to 1 for small

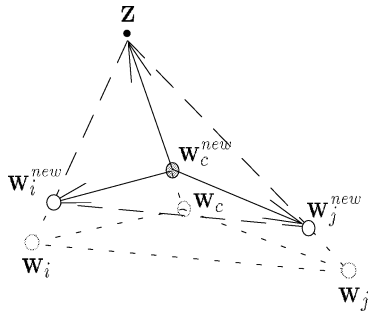


Fig. 3. Conformal mapping of the simplex when $\alpha \approx 0$ or $h' \approx 0$ in the cases of $p = 3$, $q = 2$, and $n = 3$.

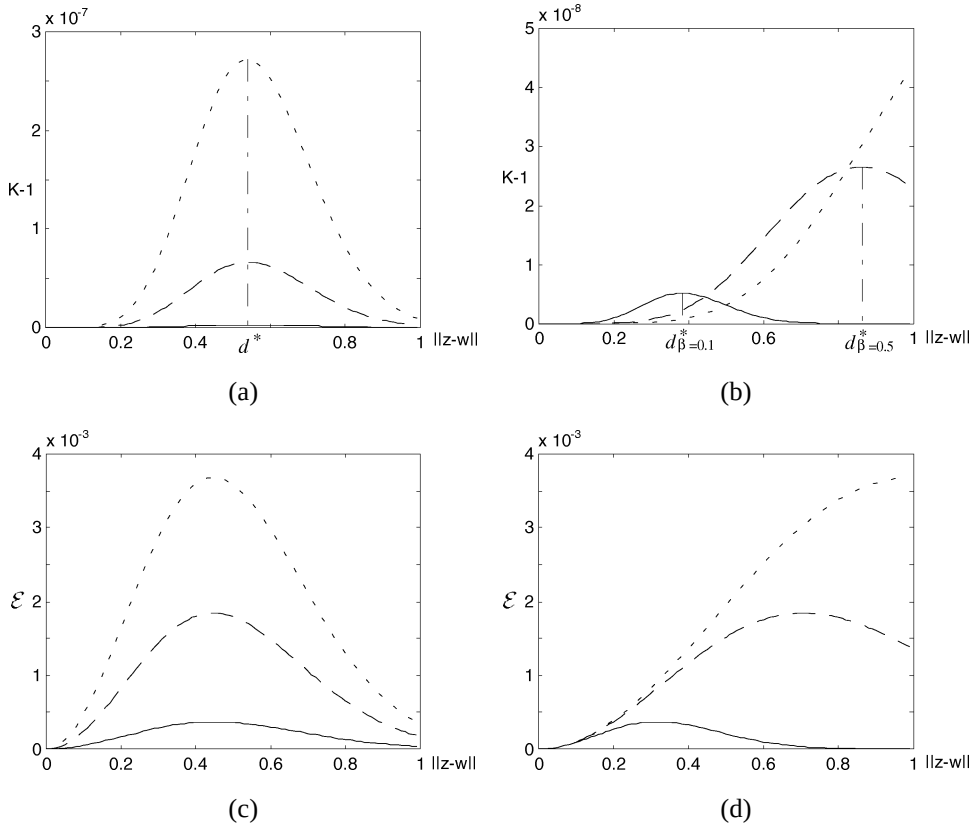


Fig. 4. Plot of measures K (E) and $\mathcal{E}$ for different $\|z-w\|$ in the examples with the following training parameters: (a) $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.1$ correspond to the solid, dashed, and dotted lines, respectively, ($\beta = 0.02$) for K (E). (b) $\beta = 0.01$, $\beta = 0.05$, and $\beta = 0.1$ correspond to the solid, dashed, and dotted lines, respectively, ($\alpha = 0.02$) for K (E). (c) $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.1$ correspond to the solid, dashed, and dotted lines, respectively, ($\beta = 0.02$) for $\mathcal{E}$. (d) $\beta = 0.01$, $\beta = 0.05$, and $\beta = 0.1$ correspond to the solid, dashed, and dotted lines, respectively, ($\alpha = 0.02$) for $\mathcal{E}$.

h' or α. This means that a conformal mapping can be constructed either by updating the synapse vector $\boldsymbol{w}$ close to the input $\boldsymbol{z}$ or by updating the synapse vector $\boldsymbol{w}$ to the opposite of the input $\boldsymbol{z}$. From the figures that the non-conformality measure depends on the distance $\|\boldsymbol{z}-\boldsymbol{w}\|$, when $\|\boldsymbol{z}-\boldsymbol{w}\| \approx 0$ or $\|\boldsymbol{z}-\boldsymbol{w}\| \approx 1$, the mapping is nearly conformal. When the value $\|\boldsymbol{z}-\boldsymbol{w}\|$ is smaller than d^*, where the value of K is maximum, further reduction of the distance ($\|\boldsymbol{z}-\boldsymbol{w}\|$) will cause the non-conformality measure to decrease. On the other hand, when the value $\|\boldsymbol{z}-\boldsymbol{w}\|$ is greater than d^*, further reduction of the distance will cause the non-conformality measure to increase. The formal self-organization dynamics tend to reduce the distance only. For comparison, Figs. 4(c) and (d) show the values of the sample energy function $\mathcal{E}$ in (5) with different network parameters α and β. Obviously, the $\mathcal{E}$ is not a suitable energy function for the formal dynamics when $\|\boldsymbol{z}-\boldsymbol{w}\|$ is greater than d^*. We will discuss more on these figures in the last section.

4.2. Conformality of two dimensions

Let us consider a simple case for the measure K, where the value of n is equal to 2. Using the simplicial coordinate, the mapping function $\boldsymbol{f}$ comprises (f_1, f_2), and each component f_i, $i = 1, 2$, is a function of (v_1, v_2), where (v_1, v_2) is a point in the two-dimensional simplex. By the definition of $K_{n=2}$ derived above, we have

$$K_{n=2} = \frac{\left(\left(\frac{\partial f_1}{\partial v_1}\right)^2 + \left(\frac{\partial f_1}{\partial v_2}\right)^2 + \left(\frac{\partial f_2}{\partial v_1}\right)^2 + \left(\frac{\partial f_2}{\partial v_2}\right)^2\right)}{2\left(\frac{\partial f_1}{\partial v_1}\frac{\partial f_2}{\partial v_2} - \frac{\partial f_1}{\partial v_2}\frac{\partial f_2}{\partial v_1}\right)}$$

$$= \frac{(-2\alpha h'\|\boldsymbol{x} - \boldsymbol{v}\|^2 + 1 - \alpha h)^2 + (1 - \alpha h)^2}{2(-2\alpha h'\|\boldsymbol{x} - \boldsymbol{v}\|^2 + 1 - \alpha h)(1 - \alpha h)}. \tag{22}$$

We also obtain a simplified formula for the deformation potential:

$$E_{n=2} = \left(-2\alpha h'\|\boldsymbol{x} - \boldsymbol{v}\|^2\right)^2 \geqslant 0. \tag{23}$$

Eq. (23) shows that the mapping deformation can be measured by means of the matching error between the relative input vector $\boldsymbol{x}$ and the relative $\boldsymbol{v}$ (or $\boldsymbol{w}$ and z). This potential is in a form that is similar to the approximated relaxation version of the sample energy function $\mathcal{E}$ in (5). This may partially explain the conjecture regarding the conformal mapping assumption for the self-organizing network. This formula provides insight into the distribution error, topology mapping, and conformality.

5. Comparison with other measures

As unsupervised learning, mapping quality may be evaluated in different ways. Usually the map is utilized to transform the topological structure from a high-dimensional input space to a low-dimensional visible reference space, specifically a 2D surface. When the dimensions of the two corresponding spaces are different, the relationships among the inputs may not be perfectly preserved in the map. Thus, mapping deformation always exists in the self-organized representation. Several measures have been devised [2,4] which are related to the distribution error and the topology preservation. As we have discussed in the previous two sections, the non-conformality measure can be used to express both the distribution error and the topology preservation for the self-organizing process.

For comparison, three measures, the mean square error (MSE), the degree of topographicity (TPG) [4], and the averaged non-conformality measure K (or the deformation potential E), for adapted neurons will be used to illustrate the mapping quality of the self-organization. The MSE, which is a simplified and discrete function of (5), is expressed as the mean square quantization error between the input vector and the corresponding clustering center, i.e.,

$$\text{MSE} \equiv \sum_{\boldsymbol{w}} \int_{z \in V_{\boldsymbol{w}}} d(z, \boldsymbol{w}) P(z)\, \mathrm{d}^p z, \tag{24}$$

where $d(\boldsymbol{z}, \boldsymbol{w})$ denotes the Euclidean distance metric between the input $\boldsymbol{z}$ and synapse vector $\boldsymbol{w}$, and $V_{\boldsymbol{w}}$ is the Voronoi partition corresponding to $\boldsymbol{w}$ [9]. The TPG, related to the topology relationships, is the mean of the average distances between two neighboring neurons, i.e.,

$$\mathrm{TPG} \equiv \sum_{\boldsymbol{w}} \sum_{\widetilde{\boldsymbol{w}} \in N_{\boldsymbol{w}}} d(\boldsymbol{w}, \widetilde{\boldsymbol{w}}), \tag{25}$$

where $N_{\boldsymbol{w}}$ denotes the neighboring synapse vector set of $\boldsymbol{w}$. The MSE and the TPG have been used in many studies. They do not have any direct relationship with the conformality. The TPG is a function of the network states only. The MSE is a function of the inputs and network states.

For each neuron, there may exist several neighboring simplices, as shown in Fig. 5. Although there are other possible simplices composed of adapted synapse vectors, depending on the effective range of the neighborhood function h, the main neighboring simplices shown in Fig. 5 indicate the major information contribution to the mapping deformation and will be used in the discussion of our simulations. For each neighboring simplex, we can calculate a K value. All the K values that belong to a winning neuron are averaged to obtain a mean value of K for this winning neuron. The averaged measure is equal to this mean value of K for the winning neuron in an evolutionary step. Note in many applications the input z is a random variable and K is a function of this random variable. K is also a random variable. We take average of them for smoothness.

Fig. 6 shows the results obtained using the three different measures to reveal the self-organizing process. The neurons in the network were arrayed in the regular square grid in a 2D plane. The results for the 1024 input samples from the uniform distribution of a unit square are plotted in Fig. 6(a). Fig. 6(b) shows another example where 1024 input samples from Gaussian distribution were used. Each line was obtained by averaging over 100 training simulations. From the results shown in Fig. 6, the averaged non-conformality measure K suitably displays the evolution of the self-organizing network, with results comparable to those for the MSE and the TPG. The results for the deformation potential E are almost coincident with the K curve, so we omit the curve for E.

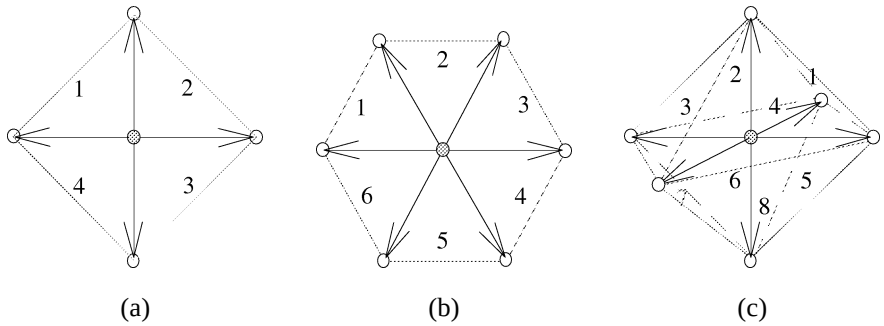
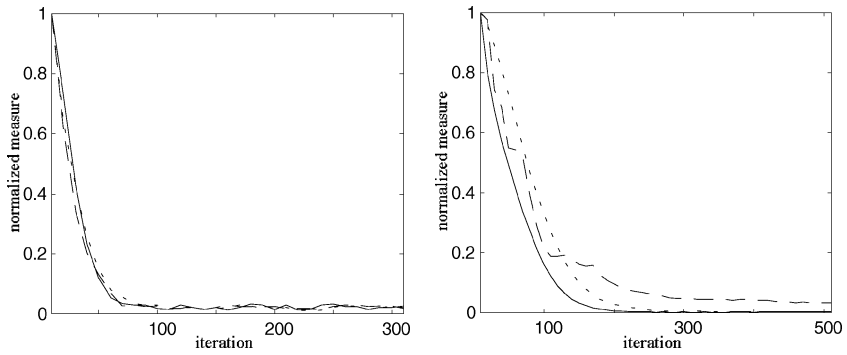


Fig. 5. The neighboring simplices of a winning neuron which is in the center in each case. The simplices are indicated by different numbers. (a) Four simplices in the case of $q = 2$. (b) Six simplices in the case of $q = 2$. (c) Eight simplices in the case of $q = 3$.



(a) (b)

Fig. 6. Different measures to indicate the self-organization learning process using the MSE, the TPG, and the averaged non-conformality measure K, plotted in the dashed, dotted, and solid lines, respectively. (a) For the 1024 uniformly distributed input samples. (b) For the 1024 Gaussian distributed input samples.

It should be noted that the training parameters α and h affect the bound of the mapping deformation. When the adaptation rate α decays too quickly or the neighborhood size is reduced too quickly, the distribution errors will remain significant [13]. Monitoring these measures provides a good way to determine the reduction of the parameters, such that preservation of both the input distribution and topology can be obtained smoothly in the evolution.

6. Simulations of surface mapping

In practical applications, it is necessary to derive a metric to monitor the accumulative activities of the consecutive quasi-conformal mappings. Because the composite of a K_1-quasiconformal and K_2-quasiconformal mapping is (K_1K_2)-quasiconformal (see, for example, [1]), we define a metric of *total non-conformality* as the product of consecutive measures, i.e.,

$$M_i = \prod_{t=1}^{T} K(\boldsymbol{z}_t, \boldsymbol{w}_i), \tag{26}$$

for the synapse vector $\boldsymbol{w}_i$ of the neuron i. It is a direct measure of how much each neuron has experienced nonconformality deformation in the entire course of the adaption process. By means of this metric (26), the self-organizing network can be applied to pattern matching problems.

We apply these measures to surface mapping problems. The cases considered for presentation are the mappings from regularly organized network configurations to the configurations of given patterns. The sampled inputs of a pattern are processed by the self-organizing network, which is initialized with a regular configuration, for example, a square grid pattern. Along with the training, we use both the averaged measure of non-conformality in (18) for the evolution of the process and the composite metric (26) for the accumulated activity of each neuron. The proper number of training evolutionary steps can be

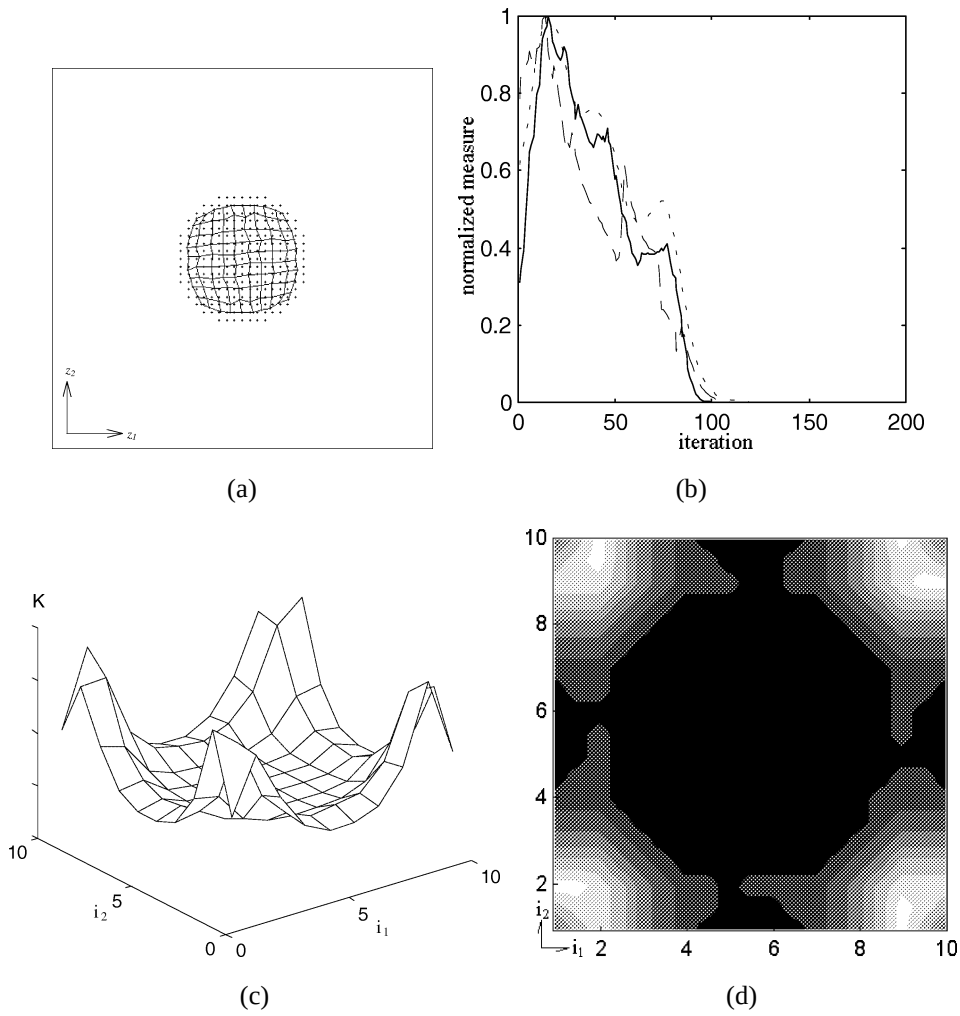


Fig. 7. Simulation results of mapping from a 10×10 square grid to a circular pattern. (a) The final trained configuration. (b) The measures of the MSE, the TPG, and K, plotted as dashed, dotted, and solid lines, respectively. (c) Composite non-conformality metric for the adaptation of each neuron in a 3D mesh surface plot and (d) in a contour plot.

determined by examining the convergence of the averaged non-conformality measure. Using formula (26), we can estimate the matching activity between the two patterns. Heavy activity may indicate a mismatch of the locations of the two patterns. In our simulations, the adaptation rate was set to be a small constant, and the neighborhood was within a fixed range.

Let us consider general cases, mappings from a two-dimensional square grid pattern to a different two-dimensional pattern. The initial configuration of the network was a 10×10 square grid pattern, regularly arranged in the input space. The training parameters were

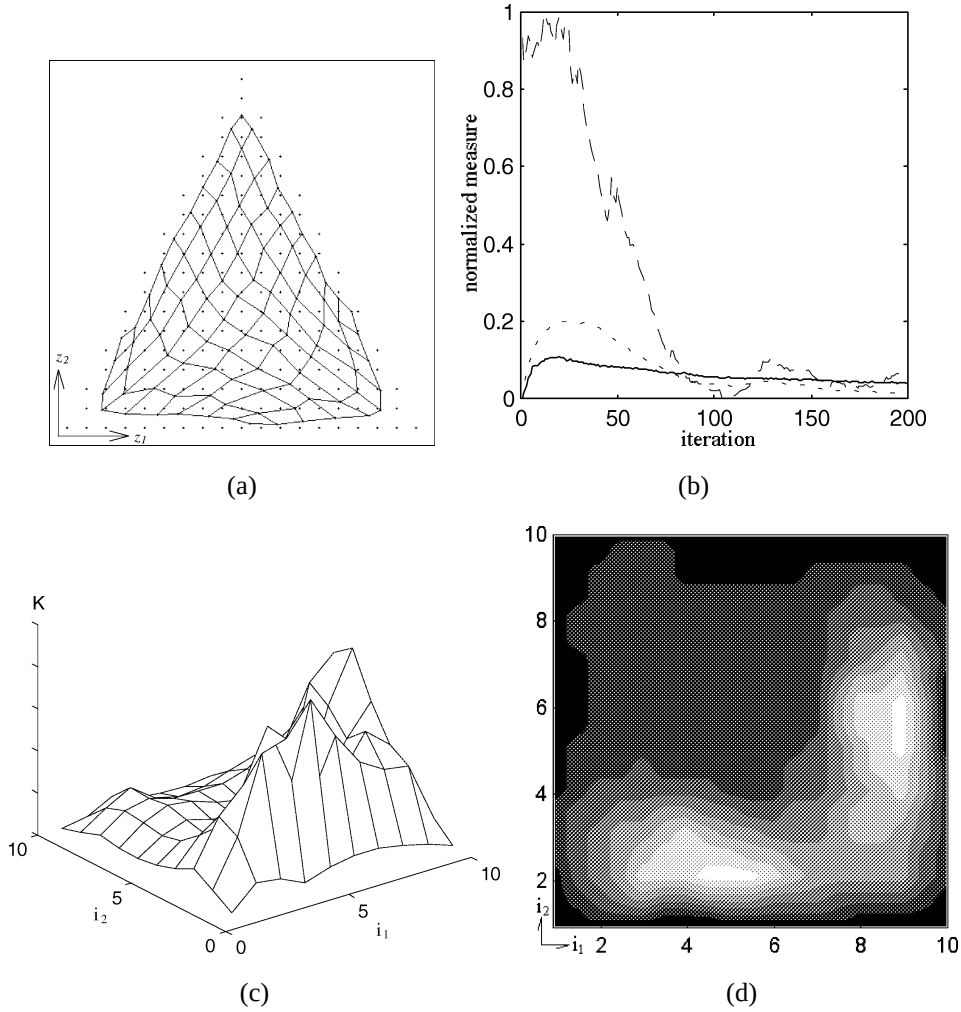


Fig. 8. Simulation results of mapping from a 10×10 square grid to a triangular pattern [9]. (a) The final trained configuration. (b) The measures of the MSE, the TPG, and K, plotted as dashed, dotted, and solid lines, respectively. (c) Composite non-conformality metric for the adaptation of each neuron in a 3D mesh surface plot and (d) in a contour plot.

the same in all the simulations. We set $\alpha = 0.02$ and $h(d) = \exp(-d^2/S^2)$, where d is the distance metric in the network and S is a quarter of the side length of the grid. The simulation results are shown in Figs. 7–9. In the figures, part (a) shows the plot of the trained network configuration, which approximates the input in the dotted area. Part (b) shows the averaged non-conformality measure, plotted as a solid line, as a function of the evolution steps. For comparison, part (b) includes the measures of the MSE and the TPG, plotted as dashed and dotted lines, respectively. Part (c) and part (d) show the composite

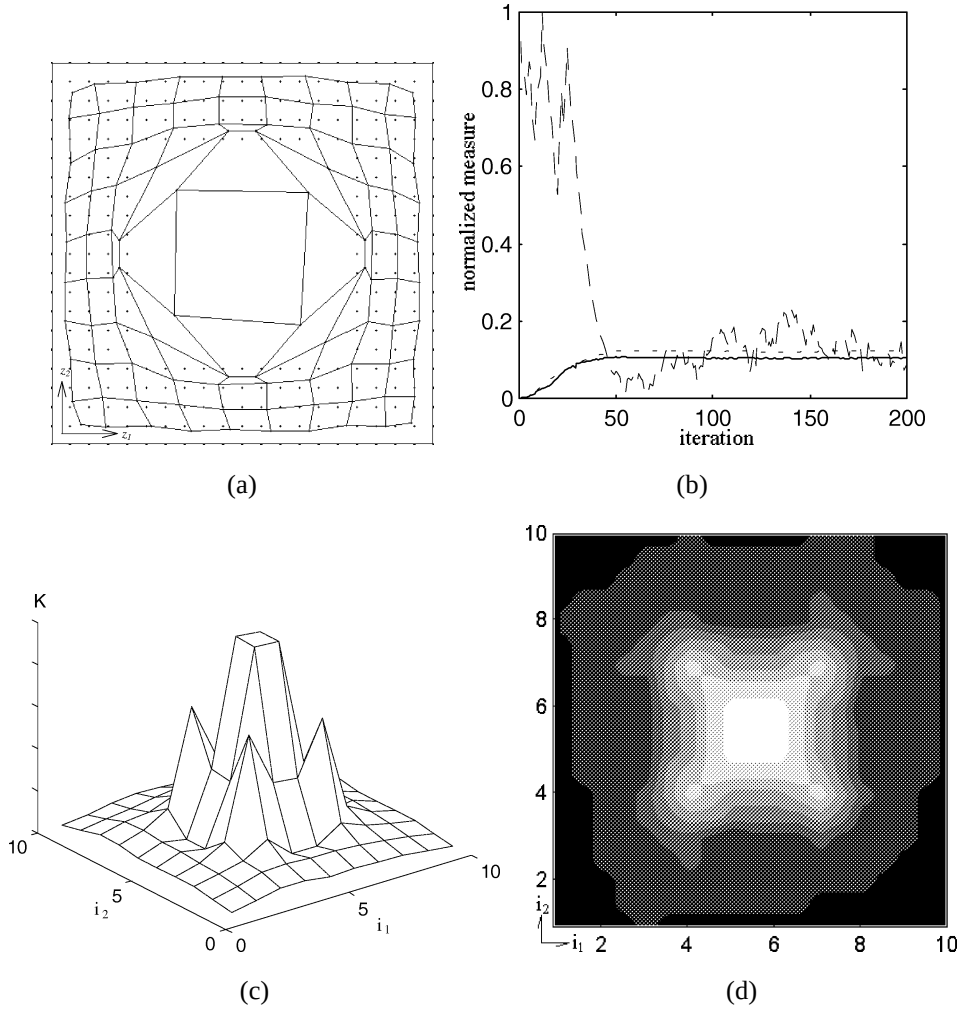


Fig. 9. Simulation results of mapping from a 10×10 square grid to a hollow square pattern. (a) The final trained configuration. (b) The measures of the MSE, the TPG, and K, plotted as dashed, dotted, and solid lines, respectively. (c) Composite non-conformality metric for the adaptation of each neuron in a 3D mesh surface plot and (d) in a contour plot.

metric (26) for each neuron. The metric is displayed with a 3D mesh surface and contour, respectively.

Mapping from a two-dimensional square grid pattern to a three-dimensional surface pattern was also considered, and the simulation results are shown in Fig. 10. The training inputs were sampled from a 3D surface with facial features. The configuration of the network was a 20×20 square grid in a 2D plane. Each neuron was initialized with its plane grid coordinate. In the figure, part (a) shows the projection of the final network

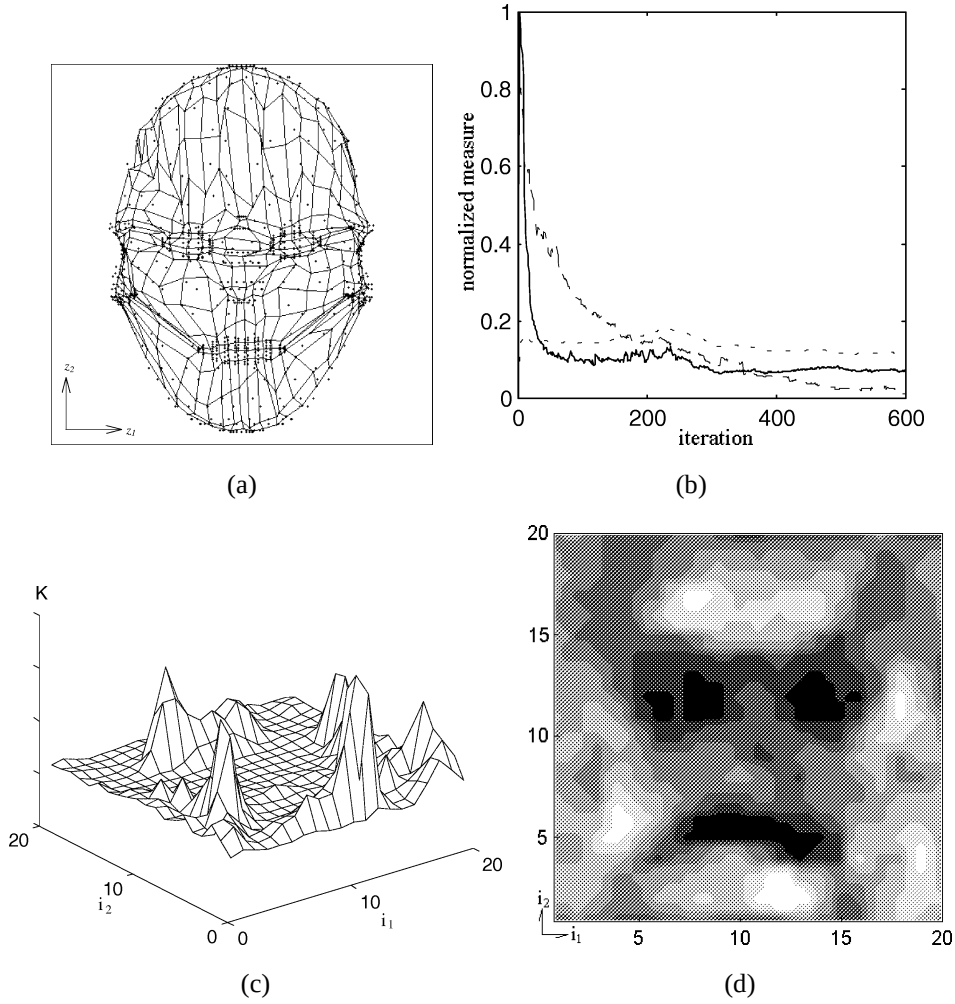


Fig. 10. Simulation results of mapping from a 20×20 square grid to a 3D surface. (a) Projection of the final trained configuration. (b) The measures of the MSE, the TPG, and K, plotted as dashed, dotted, and solid lines, respectively. (c) Composite non-conformality metric for the adaptation of each neuron in a 3D mesh surface plot and (d) in a contour plot.

configuration. Part (b) show the measures obtained during the training process. Parts (c) and (d) show the values of the composite metric.

From these results, it is clear that the convergence of non-conformality is consistent with that of the distribution error and the topology. This fact meets the conformal mapping assumption. In part (b) for each simulation, all the proposed three measures Q, K and E will give similar results as the solid line. The heavy activity in the mismatched parts of the two patterns is properly indicated by the large M_i values as shown in parts (c) and (d). The results shown in (c) and (d) are in accordance with the deformation of the final network

configuration shown in (a) for each case. This composite metric provides a good means of locating the mismatches between two image patterns. This technique has been used for medical images and airbone images in the projects.

7. Discussion

In this work, we have provided measures by using simplices to investigate both the deformation and the conformality of the self-organizing mapping. The deformation measure Q, the non-conformality measure K, and the deformation potential E can be estimated accurately in each evolutionary step. These measures give us additional information indicating the development of the network during its evolution. The total non-conformality can indicate the defective (or mismatched) parts between two images, which is extremely useful in pattern matching applications.

It has been shown that the mapping of synapse vectors is quasi-conformal with a bound, and that the bound has a limit which is equal to one. The formulas for the non-conformality measure and deformation measure can serve as an objective cost function to obtain new dynamics to evolve a network. We now discuss these new dynamics in depth.

The energy functions $\mathcal{E}$ and K will be minimized when $\|z - w\|$ is either close to 0 or close to 1, where $\|z - w\|$ is properly normalized within 0 and 1. This means that an input pattern may have a similar representation or a totally opposite representation on the map. Both representations may be mixed on one network. The exact dynamics derived from such energy functions will give a different mapping. The opposite representation is not considered in the formal self-organization model [9]. This is because the formal model starts from large α and β and reduces them gradually to keep only a single minimum state at $\|z - w\| = 0$ in the network. As seen in Fig. 4, all $\|z - w\|$ will be included on one side of the $\mathcal{E}$ curve with a very large β value. The dynamics of the energy function $\mathcal{E}$ will drive the network state down the curve toward the position with a small $\|z - w\|$ value. In this case, all the synapse vectors w will be forced to match the input z as closely as possible. β must be large enough to enable all the values of $\|z - w\|$ to be on the left side of the curve mode, $\|z - w\| < d^*$. The formal dynamics, (1) and (2), will not work for $\|z - w\| > d^*$. This is the reason for the increase of the MSE when β decreases too quickly [13]. This is also the reason for increasing the similarity (decreasing the $\|z - w\|$) improving the conformality. This explains the conformal mapping assumption.

The intuitively devised energy $\mathcal{E}$ does not match the formal dynamics, (1) and (2), exactly, which will drive the synapse vector w toward the only position, $\|z - w\| \approx 0$. We may slightly modify the dynamics of the formal self-organization model to obtain exact dynamics by selecting the winning neuron with the smallest value $\mathcal{E}$ (or K) for a given input z and updating the synapse vector following the slope of $\mathcal{E}$ (or K) which is $\partial\mathcal{E}/\partial\|z - w\|$.

The measures have been used as criterions to derive learning algorithms for the multilayer perceptron to obtain natural coordinates for the input set. [7] discussed one conformal mapping application. This conformality measure is also necessary in constructing a form mapping among organisms [20]. Both the interactions of neighboring cells in an organism [22] and the conformal form transformation are accomplished using

such a network where neurons are located in the surface grid points representing a 3D organism form, such as a Cyprinus carpio fish. The input data is sampled from another fish surface, such as an Acanthopagrus schlegeli fish. In this case $p = q = 3$. The synapse vector contains the coordinates of the surface point. Using K as an evolutionary criterion, we can obtain the form transformation between these two species. The corresponding parts and natural coordinates can be located according to this transformation [12]. The deformation parts between the fishes can be quantified.

Appendix A. The eigenvalues of $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$

Let $\boldsymbol{A}$ denote the Jacobian matrix $\mathrm{d}\boldsymbol{f}_{\boldsymbol{v}}$ with the components $a_{ij} = \partial f_i/\partial v_j$, which are defined in (8), for $1 \leqslant i, j \leqslant n$. Obviously, the matrix $\boldsymbol{A}$ is symmetric, i.e., $\boldsymbol{A}^{\mathrm{t}} = \boldsymbol{A}$. The characteristic polynomial of $\boldsymbol{A}$, $c(\lambda) = \det(\boldsymbol{A} - \lambda\mathbf{I})$, splits. Every eigenvalue of $\boldsymbol{A}$ is real.

First, we will show that λ_1 and λ_2 in (9) are two eigenvalues of the matrix $\boldsymbol{A}$. For λ_1,

$$
\begin{aligned}
c(\lambda_1) = c(1-\alpha h) &= (-2\alpha h')^n \\
&\quad \times \det \begin{pmatrix} (x_1-v_1)^2 & (x_1-v_1)(x_2-v_2) & \dots & (x_1-v_1)(x_n-v_n) \\ (x_2-v_2)(x_1-v_1) & (x_2-v_2)^2 & \dots & (x_2-v_2)(x_n-v_n) \\ \vdots & \vdots & \ddots & \vdots \\ (x_n-v_n)(x_1-v_1) & (x_n-v_n)(x_2-v_2) & \dots & (x_n-v_n)^2 \end{pmatrix} \\
&= (-2\alpha h')^n \prod_{k=1}^{n} (x_k - v_k) \cdot \det \begin{pmatrix} (x_1-v_1) & (x_2-v_2) & \dots & (x_n-v_n) \\ (x_1-v_1) & (x_2-v_2) & \dots & (x_n-v_n) \\ \vdots & \vdots & \ddots & \vdots \\ (x_1-v_1) & (x_2-v_2) & \dots & (x_n-v_n) \end{pmatrix} \\
&= 0. \qquad \text{(A.1)}
\end{aligned}
$$

For λ_2,

$$
\begin{aligned}
c(\lambda_2) &= c\left(-2\alpha h' \sum_{k=1}^{n} (x_k - v_k)^2 + (1-\alpha h)\right) \\
&= \det \begin{pmatrix} a_{11}-\lambda_2 & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda_2 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda_2 \end{pmatrix} \\
&= \det \begin{pmatrix} a_{11}-\lambda_2 & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda_2 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}, \qquad \text{(A.2)}
\end{aligned}
$$

where

$$b_{nj} = \big(a_{nj} - \delta(n-j)\lambda_2\big) + \sum_{i=1}^{n-1} \frac{(x_i - v_i)}{(x_n - v_n)}\big(a_{ij} - \delta(i-j)\lambda_2\big), \quad j = 1, \ldots, n, \tag{A.3}$$

and δ is the Kronecker delta function. For $1 \leqslant j \leqslant n-1$,

$$\begin{aligned} b_{nj} &= (-2\alpha h')\Bigg((x_n - v_n)(x_j - v_j) \\ &\qquad + (x_j - v_j)\sum_{i=1}^{n-1}\frac{(x_i - v_i)^2}{(x_n - v_n)} - \frac{(x_j - v_j)}{(x_n - v_n)}\sum_{k=1}^{n}(x_k - v_k)^2\Bigg) \\ &= (-2\alpha h')\frac{(x_j - v_j)}{(x_n - v_n)}\Bigg(\sum_{i=1}^{n}(x_i - v_i)^2 - \sum_{k=1}^{n}(x_k - v_k)^2\Bigg) = 0. \end{aligned} \tag{A.4}$$

For $j = n$,

$$\begin{aligned} b_{nn} &= (-2\alpha h')\Bigg(\Bigg((x_n - v_n)^2 - \sum_{k=1}^{n}(x_k - v_k)^2\Bigg) \\ &\qquad + \sum_{i=1}^{n-1}\frac{(x_i - v_i)^2(x_n - v_n)}{(x_n - v_n)}\Bigg) \\ &= (-2\alpha h')\Bigg(\sum_{i=1}^{n}(x_i - v_i)^2 - \sum_{k=1}^{n}(x_k - v_k)^2\Bigg) = 0. \end{aligned} \tag{A.5}$$

Then, we have

$$c(\lambda_2) = \det\begin{pmatrix} a_{11} - \lambda_2 & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda_2 & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix} = 0. \tag{A.6}$$

Thus, λ_1 and λ_2 are two eigenvalues of the matrix $\boldsymbol{A}$ by (A.1) and (A.6).

Secondly, we will determine the multiplicities of the eigenvalues. Suppose that the multiplicities of λ_1 and λ_2 are m_1 and m_2, respectively. Because λ_1 and λ_2 are two of the eigenvalues of $\boldsymbol{A}$, it follows that

$$m_1 + m_2 \leqslant n. \tag{A.7}$$

The matrix $\boldsymbol{A}$ is symmetric; thus,

$$m_1 = \dim(E_{\lambda_1}) \geqslant 1 \quad \text{and} \quad m_2 = \dim(E_{\lambda_2}) \geqslant 1, \tag{A.8}$$

where $\dim(E_\lambda) = \text{nullity}(\boldsymbol{A} - \lambda\mathbf{I}) = n - \text{rank}(\boldsymbol{A} - \lambda\mathbf{I})$. For m_1,

$$\begin{aligned}
&\operatorname{rank}(\boldsymbol{A}-\lambda_1\mathbf{I})\\
&\quad\leqslant \operatorname{rank}\begin{pmatrix}
(x_1-v_1) & (x_2-v_2) & \dots & (x_n-v_n)\\
(x_1-v_1) & (x_2-v_2) & \dots & (x_n-v_n)\\
\vdots & \vdots & \ddots & \vdots\\
(x_1-v_1) & (x_2-v_2) & \dots & (x_n-v_n)
\end{pmatrix}\\
&\quad\leqslant 1. \qquad \text{(A.9)}
\end{aligned}$$

Then,

$$m_1 = \operatorname{nullity}(\boldsymbol{A}-\lambda_1\mathbf{I}) \geqslant n-1. \qquad \text{(A.10)}$$

From (A.7), (A.8), and (A.10), the multiplicities of λ_1 and λ_2 are $m_1 = n-1$ and $m_2 = 1$, respectively.

Appendix B. Derivation of the non-conformality measure K

To evaluate the terms in (15), the results in Appendix A will be utilized. First, we will evaluate the term $\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij})^2$, where $a_{ij} = \partial f_i/\partial v_j$. Let λ be an eigenvalue of the matrix A and $\boldsymbol{u}$ be the corresponding eigenvector. Because $\boldsymbol{A}$ is symmetric, it follows that

$$\boldsymbol{A}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{A}^{\mathrm{t}}\boldsymbol{u} = \lambda\boldsymbol{A}\boldsymbol{u} = \lambda^2\boldsymbol{u}. \qquad \text{(B.1)}$$

Then, λ^2 is an eigenvalue of the matrix $\boldsymbol{A}^{\mathrm{t}}\boldsymbol{A}$:

$$\begin{aligned}
\operatorname{trace}(\boldsymbol{A}^{\mathrm{t}}\boldsymbol{A}) &= \sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij})^2 = \sum_{i=1}^{n}\lambda_i^2 = (n-1)\lambda_1^2+\lambda_2^2\\
&= \left((-2\alpha h')\sum_{i=1}^{n}(x_i-v_i)^2 + (1-\alpha h)\right)^2 + (n-1)(1-\alpha h)^2. \qquad \text{(B.2)}
\end{aligned}$$

Secondly, we can evaluate the Jacobian $J(\boldsymbol{v}, \boldsymbol{f})$:

$$\begin{aligned}
J(\boldsymbol{v},\boldsymbol{f}) &= \det(\boldsymbol{A})\\
&= \prod_{i=1}^{n}\lambda_i = \lambda_2{\lambda_1}^{n-1}\\
&= (-2\alpha h')(1-\alpha h)^{n-1}\sum_{i=1}^{n}(x_i-v_i)^2 + (1-\alpha h)^n. \qquad \text{(B.3)}
\end{aligned}$$

References

[1] L.V. Ahlfors, Lectures on Quasi-conformal Mappings, D. Van Nostrand, Princeton, NJ, 1966.
[2] H.-U. Bauer, K.R. Pawelzik, Quantifying the neighborhood preservation of self-organizing feature maps, IEEE Trans. Neural Networks 3 (1992) 570–578.

[3] C.M. Bishop, M. Svensen, C.K.I. Williams, GTM: The generative topographic mapping, Technical Report NCRG/96/015, Department of Computer Science and Applied Mathematics, Aston University, Birmingham, 1997.
[4] D.I. Choi, S.H. Park, Self-creating and organizing neural networks, IEEE Trans. Neural Networks 5 (1994) 561–575.
[5] E. Erwin, K. Obermayer, K. Schulten, Self-organizing maps: Stationary states, metastability and convergence rate, Biol. Cybernet. 67 (1992) 35–45.
[6] E. Erwin, K. Obermayer, K. Schulten, Self-organizing maps: Ordering, convergence properties and energy functions, Biol. Cybernet. 67 (1992) 47–55.
[7] A. Hyvarinen, P. Pajunen, Nonlinear independent component analysis: Existence and uniqueness results, Neural Networks 12 (1999) 429–439.
[8] T. Kohonen, Self-organizing maps: Optimization approaches, in: Artificial Neural Networks, Proc. ICANN-91, Elsevier, Amsterdam, 1991, pp. 981–990.
[9] T. Kohonen, Self-organizing Maps, Springer, Berlin, 1995.
[10] M. Lemmon, Topologically ordered competitive sampling, Neural Networks 7 (1994) 101–111.
[11] X. Li, J. Gasteiger, J. Zupan, On the topology distortion in self-organizing feature map, Biol. Cybernet. 70 (1993) 189–198.
[12] C.-Y. Liou, W.-P. Tai, Conformal self-organization for continuity on a feature map, Neural Networks 12 (1999) 893–905.
[13] Z.-P. Lo, B. Bavarian, On the rate of convergence in topology preserving neural networks, Biol. Cybernet. 65 (1991) 55–63.
[14] S.P. Luttrell, Self-organization: A derivation from first principles of a class of learning algorithms, in: Proc. 3rd IEEE International Joint Conference on Neural Networks, Washington, DC, 1989, pp. 495–498.
[15] Y.G. Reshetnyak, Estimates of the continuity modulus for various mappings, Sibirskii Matematicheskii Zhurnal 7 (1966) 1106–1114.
[16] Y.G. Reshetnyak, Mappings with bounded deformation as extremal of Dirichlet type integrals, Sibirskii Matematicheskii Zhurnal 9 (1968) 652–666.
[17] H. Ritter, K.J. Schulten, Kohonen's self-organizing maps: Exploring their computational capabilities, in: Proc. IEEE International Conference Neural Networks, Piscataway, NJ, 1988, pp. 109–116.
[18] H. Robbins, S. Monro, A stochastic approximation method, Ann. Math. Statist. 22 (1951) 400–407.
[19] T. Tanaka, On evaluation of reference vector density for self-organizing feature map, IEICE Trans. Inf. and Syst. E77-D (1994) 402–408.
[20] D'A. W. Thompson, On Growth and Form, 2nd Edn., Cambridge University Press, Cambridge, 1952, p. 1083.
[21] V.V. Tolat, An analysis of Kohonen's self-organizing maps using a system of energy functions, Biol. Cybernet. 64 (1990) 155–164.
[22] A.M. Turing, The chemical basis of morphogenesis, Philos. Trans. Roy. Soc. 237 (1952) 37–72.