Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

A Study on Energy Artificial Neuron Model and Its Applications in Self-growing and Self-organizing Neural Network

Hao Liu*, **. Xiaojuan Ban*, *. Zhuoran Xu*, **. Masahito Kurihara****

**Department of Computer Science and Technology, University of Science and Technology, Beijing, Beijing, China*
***Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Japan*

Abstract: Neurons have been always considered as the active and main players in the brain's activities all the time. However, According to recent studies in the neuroscience field, the glial cells is playing a more and more important role in the brain's activities and the brain should be regarded as a system consisted of both neurons and glial cells. Furthermore, it has been proved to be related to the growth of neurons. In this paper, a new artificial neuron model called EAN Model (Energy Artificial Neuron Model) which is based on the energy concept from the glial cells is proposed, and a way to demonstrate EAN model in mathematics is suggested. Based on EAN model, a self-growing and self-organizing neural network called ESGSONN (EAN Based Self-growing and Self-organizing Neural Network) is realized, which has following features: rapid growing, density persevering, no or less dead neurons and incremental learning. New features of ESGSONN have been shown by comparable experiments.

Keywords*:* adaptive algorithms, neural networks, self-organizing system, entropy.

1. INTRODUCTION

The active vertebrate brain always represents highly self-organizing. To simulate the mechanism of the biological activities, researchers began their studies from 1970s (D. J. Willshaw, 1976). Self-organizing Feature Map (SOFM) was proposed at 1982 (Kohonen, T, 1982, J. Kangas& T. Kohnen, 1990). The basic SOFM takes an unsupervised learning and it has a solid structure of the network which usually consists of a set of neurons arranged in a 2-D structure. Each neuron is attached to a reference vector that is adjusted during the training procedure. When training process is completed, SOFM divides the input space into several regions which represent corresponding clusters in the input space. The density probability of input data is preserved in those regions.

SOFM has been widely used in many areas, such as pattern recognition, image processing, solving TSP and etc. But some limitations have been noted: First, dead neuron, a neuron which will not learn in any case, may exist if the input data distribute as a complex structure. Second, SOFM can't overcome the so-called stability-plasticity dilemma. Third, SOFM must pre-define the map size, which usually results in difficulties of selecting the appropriate map size.

To solve those drawbacks, another kind of self-organizing neural networks, self-growing and self-organizing networks, have been proposed. This kind of network has a feature of growing, and it means that they can dynamically change their structures according to the input data. Some of those networks are considered below.

*** Corresponding author. *E-mail addresses:* banxj@ustb.edu.cn

This work was Supported by National 863 Program (No. 2009AA062800, No.2009AA04Z163), National Nature Science Foundation of P. R. China (No.60973063) and supported by the Fundamental Research Funds for the Central Universities (FRF-TP-09-016B).

Growing Cell Structures (GCS) (Fritzke B, 1994): GCS is an algorithm motivated by SOFM, but it uses a triangle network which is defined by the connections of nodes instead of the two-dimensional grid. Each triangle cover the area of input space of nonzero distribution, and the whole network try to cover all over those areas. The algorithm begins with a random triangle and inserts new nodes to decrease the local highest error until the terminal conditions are reached. However, because the whole process is performed in the input space, visualizing high-dimensional data into a two-dimension plane with their topology cause a problem.

Growing Neural Gas (GNG) (Fritzke B, 1995): GNG is an algorithm based on neural gas algorithm, which begins with two neurons. During the training procedure, new nodes are added as the way in GCS and two nodes with the highest activity are moved together with a link created between them. The algorithm ends when stopping criterions is reached. Because of preserving the topology in the input space, GNG also has the visualizing problem as GCS.

Cell Splitting Grid (CSG) (Tommy W.S, 2004): CGS is a new approach of self–growing network, which simulate the mechanism of cell splitting in biological organs. Each neuron has an activation level which is decreased when its weight vector is adjusted. A neuron is split up into four neurons when its activation level equals zero. The algorithm starts with one neuron and split existing neurons. CGS results in a regular square structure which can easily reflect the density in the input space.

More algorithms of self-growing and self-organizing networks are considered below: Self-creating and Organizing Neural Networks (SCONN) (Doo-Il Choi et al, 1991); Dynamic Self-organizing Map (DSOM) (D. Alahakoon et al, 2000); Self-organizing Incremental Neural Network (SOINN) (Shen F and Hasegawa O, 2006);

These algorithms try to overcome the limitation of fixed structure of SOFM and some other drawbacks aforementioned. Unfortunately, they also have the limitation of slow growing which is partly due to using Local Error as the condition of growing, or Mean Quantization Error as the convergence condition of the network.

In this paper, a new artificial neuron model called EAN (Energy Artificial Neuron) model is proposed based on the energy concept from the glial cells according to the recent discoveries in the neuroscience field. We demonstrate EAN model in mathematics. Additionally, we implement a new self-growing and self-organizing neural network based on EAN model called ESGSONN (EAN Based Self-growing and Self-organizing Neural Network). ESGSONN considers the energy in EAN, the entropy productions in the network and the measure of similarity as its conditions of growing as well as competitions. Its main features are described as below: rapid growing, density persevering, no or less dead neurons and incremental learning. New features of ESGSONN have been shown by comparable experiments with other self-organizing network.

2. THE ENERGY ARTIFICIAL NEURON MODEL

2.1 Recent discoveries in the neuroscience field on glial cells

The cells that construct central nervous system (CNS) can be separated into two kinds: neurons and glial cells. In the past biologic researches, neurons were regarded as the active and main objects in the neural activities. And glial cells were viewed as passive elements which just provide structural and metabolic support to the neurons. However, recent studies show that glial cells may play a very important role in the information processing and it has been proved that glial cells are related to the growth of neurons: Glial cells may increase the amount of synapse and enhance the synaptic efficacy violently (Pfrieger FW& Barres, 1997). BA. Ulian has demonstrated that the glial cells are necessary to sustain the synapses in his experiment (Ullian EM et al, 2001). And astroglias (a kind of glial cell) can apparently induce the neurogenesis (Song H. et al, 2002). Furthermore, glial cells around synapses are essential for the growing of synapses (Slezak M, Pfrieger FW, and et al, 2006). One of the major functions of glial cell is its energy support to neurons. Glial cell can enhance transfer efficiency of glucose through oxidative metabolism and release lactic acid, which creates energy in the form of ATP (Joachim W. Deitmer, 2000, 2001) that can modulate neurons around the glial cell. This modulation can present activation or inhibition (Eric A. Newman, 2003, Joachim W. Deitmer, 2006) and provides a foundation to other complex functions.

Those biologic researches demonstrate that glial cells are responsible to complex and important functions in nervous system, such as growing of synapse, constructing nervous system, learning and memories. So, in a brief conclusion, the brain should be regarded as a system consisted of both neurons and glial cells (Paola Bezzi and Andrea Volterra. 2001). And these features of biological glial cell should be also useful to set up an artificial network for the neurons.

2.2 The Energy Artificial Neuron Model (EAN Model)

In this part, we introduce a new artificial neuron model called EAN (Energy Artificial Neuron) model, which brings the concept of energy to the traditional M-P artificial neuron model. The main contribution of EAN model is providing the threshold of network growing. However, to avoid increasing the complexity of the network system, only energy feature is extracted and implemented.

Definition 1 An EAN model can be defined by a 7-tuple as below:

$$\mathbf{EAN} = \langle \boldsymbol{X}, \boldsymbol{O}, \boldsymbol{M}, \boldsymbol{I}, \boldsymbol{F}, \boldsymbol{E_{AN}}, \boldsymbol{L_{AN}} \rangle$$

where

$\boldsymbol{X}$ is the input vector, such as $\mathrm{X} = (x_1, x, \dots, x_n)^T$.

$\boldsymbol{O}$ is the output vector, such as $O = \{\mathrm{o}(t)\}$.

$\boldsymbol{M}$ is a unit of local memory which has two components:

$$\boldsymbol{M} = \boldsymbol{M_w} \cup \boldsymbol{M_L}$$

where $M_w = \{u\}$ is the STM (Short Term Memory) and $M_L = \{w = \{w_i\}_{i=1}^{n}, b\}$ is the LTM (Long Term Memory).

$\boldsymbol{I}$ is the integrated mapping.

$\boldsymbol{F}$ is the activation mapping which could be: $o = f(\sum_{i=1}^{n} X_i W_i - \theta)$.

$\boldsymbol{E_{AN}}$ is the total energies of the current neuron.

$\boldsymbol{L_{AN}}$ represents the generalized Hebb learning rule.

The EAN model is an improved model of M-P model, and the structure of EAN model is illustrated as Fig. 1.

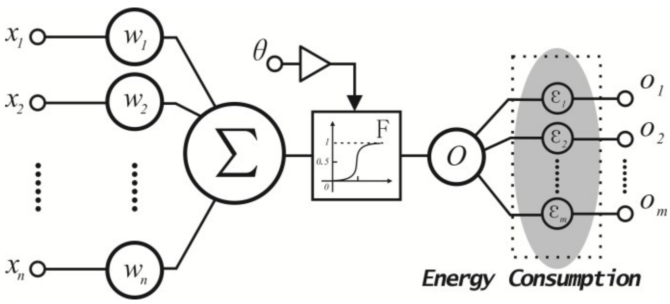


Fig. 1. The structure of the EAN model.

Definition 2 The total energy of each EAN_i is defined as $E_{AN}^{(i)}$:

$$\boldsymbol{E_{AN}^{(i)}} = \boldsymbol{E_{AN}^{consumption(i)}} \cup \boldsymbol{E_{AN}^{surplus(i)}}$$

There are two parts to compose an E_{AN} . One is $E_{AN}^{consumption(i)}$, which shows the energy consumption when two EANs establish a connection between each other. $E_{AN}^{consumption(i)} = \{e_{ij} \mid j = 1,2,\dots,m\}$ where e_{ij} is the energy consumption when EAN_i connects to EAN_j at the moment of t+1. And the other is $E_{AN}^{surplus}$ which is the surplus energies in an EAN.

Definition 3 In the Definition 2, we has demonstrated the energy consumptions as $E_{AN}^{consumption(i)}$, so each e_{ij} in the

$E_{AN}^{consumption(i)}$ can be calculated by $e_{ij} = \varepsilon_{ij} \cdot \bar{e}$, where $\bar{e}$ is the average expectation for the energy consumption when the connection is established between EAN_i and EAN_j, and ε_{ij} is a factor for this energy consumption. As a result, if each EAN_i wants to create a connection to EAN_j, all the factors of EANs in the network can be demonstrated as a matrix, such as W_e:

$$\boldsymbol{W_e} = \begin{bmatrix} \boldsymbol{\varepsilon_{11}} & \boldsymbol{\varepsilon_{12}} & \cdots & \boldsymbol{\varepsilon_{1j}} \\ \boldsymbol{\varepsilon_{21}} & \boldsymbol{\varepsilon_{22}} & \cdots & \boldsymbol{\varepsilon_{2j}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\varepsilon_{i1}} & \boldsymbol{\varepsilon_{i2}} & \cdots & \boldsymbol{\varepsilon_{ij}} \end{bmatrix}$$

Then, the total energies' consumption of EAN_i can be calculated by this formula below:

$$\boldsymbol{E_{AN}^{consumption(i)}} = \sum_{j=1}^{m} \bar{\boldsymbol{e}} \cdot \boldsymbol{\varepsilon_{ij}}$$

And the $E_{AN}^{surplus(i)}$ can be calculated by:

$$\boldsymbol{E_{AN}^{surplus(i)}} = \boldsymbol{E_{AN}^{(i)}} - \boldsymbol{E_{AN}^{consumption(i)}}$$

According to the Definition1 to Definition 3, the EAN model has two features below:

Feature 1 If $\forall\varepsilon_{ij} \to \varepsilon_{ij} \cdot \bar{e} \le E_{AN}^{surplus}$, in another word, if $\exists\varepsilon_{ij} \to \varepsilon_{ij} \cdot \bar{e} > E_{AN}^{surplus}$, the connection cannot be created.

Feature 2 The energy mechanism in EAN model could provide a threshold in the network growing. However, if the connection has been established, the connection-strength between two EANs is determined by their connection weights, not the energy consumption.

3. ESGSONN: A EAN BASED SELF-GROWING AND SELF-ORGANIZING NEURAL NETWORK

3.1 Basic Concepts of ESGSONN

It is different from the other SONNs (Self-organizing Neuron Network, SONN). The ESGSONN use the EAN model as its neurons and it started from TNU (Treble Neuron Unit, TNU). Here are some basic concepts of ESGSONN:

Definition 4 NU (Neuron Unit) is the fundamental unit in the procedure of network growing. (See Fig. 2 (a)) A NU can be defined as 3-tuple as below:

$$\boldsymbol{EAN^{unit}} = \langle \boldsymbol{G_{AN}^{unit}}, \boldsymbol{W_G}, \Delta \boldsymbol{H_G^{unit}}(\boldsymbol{t+1}) \rangle,$$

where:

- $\boldsymbol{G_{AN}^{unit}}$ is a graph that is composed of six EANs placed as hexagon structure in 2-D. A G_{AN}^{unit} can be detailed as: $G_{AN}^{unit} = \langle V_c, E_c \rangle$, where $V_c = \{EAN_i | i = 1,2,\ldots,6\}$ is a set of EANs and $E_c = \{e_{ij} | i,j = 1,2,\ldots,6\}$ is a set of borders which start from EAN_i and end at EAN_j.
- $\boldsymbol{W_G}$ is a matrix of weights which are attached to the corresponding neurons.
- $\Delta \boldsymbol{H_G^{unit}}(\boldsymbol{t+1})$ is the entropy production of EAN^{unit} at the moment of t+1, however, $\Delta H_G^{unit}(t+1)$ is a property of EAN^{unit} at the moment of t.

A NU has two features below:

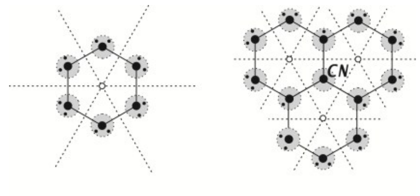
- The amount of energy in each EAN of NU is four, and the expectation of energy-consumption is one. Every energy-consumption for the connection establishing is the same. In another word, $\forall EAN_i \in EAN^{unit}$ must meets:

$$\boldsymbol{E_{AN}} = \boldsymbol{E_{AN}^{consumption(i)}} \cup \boldsymbol{E_{AN}^{surplus(i)}} = \boldsymbol{4}, \boldsymbol{W_e} = \begin{bmatrix} \boldsymbol{1} & \cdots & \boldsymbol{1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{1} & \cdots & \boldsymbol{1} \end{bmatrix}$$

- The incremental entropy production in EAN^{unit} can be calculate by Shannon Entropy:

$$\Delta \boldsymbol{H_G^{unit}}(\boldsymbol{t+1}) = -\sum_{\boldsymbol{R \in S(R)}} \boldsymbol{P(R_{t+1})\, log_2\, P(R_{t+1})},$$

where the $P(R_{t+1})$ is determined by the $E_{AN}^{surplus}$ at the moment of t.



(a) (b)

Fig. 2. Each black point in this Fig. presents one 'energy' in EAN model. (a) A NU (Neuron Unit) topological graph. (b) A TNU (Treble Neuron Unit) topological graph; and a CN (Central Neuron) locates in the center topological graph.

Definition 5 TNU (Treble Neuron Unit) has thirteen nodes in total including four public nodes. The TNU defined as below.

$$\boldsymbol{EAN^{unit3}} = \langle \boldsymbol{G_{AN}^{unit3}}, \boldsymbol{W_G}, \Delta \boldsymbol{H_G^{unit3}}(\boldsymbol{t+1}) \rangle$$

The TNU's structure is shown in Fig. 2 (b).

Definition 6 CN (Central Neuron) is a neuron which locates at the central of TNU's topological graph. Usually, CN is a neuron that has max count of learning-circles; however, because of the limitation of the energy, a CN cannot be a public node or connect to other neurons. A CN is shown in Fig. 2(b).

Definition 7 GP (Growing Point) is the place where network could add new nodes. It can be described as 2-tuple as below:

$$\mathbf{GP} = \langle \boldsymbol{V_G}, \boldsymbol{E_G} \rangle,$$

where:

- $V_G = \{EAN_i, EAN_j, \ldots, EAN_n\}$ is a set of EANs in GP.
- $E_G = \{e_{ij} | i,j = 1,2,\ldots,n\}$ is a set of connection between EAN_i and EAN_j which exist in GP.

Moreover, a GP must meet two conditions below:

1. The topological graph must be a proper sub-graph of NU; the nodes existing in GP and the new nodes must be placed in a hexagon-topology strictly. In another word, each EAN in GP must meets:

$$\text{GP} \subset \boldsymbol{G_t} \ and \ GP_{(\boldsymbol{t+1})} \subsetneq \boldsymbol{G_{AN}^{unit}}$$

where G_t is the current network topology. And $GP_{(t+1)}$ is the new topology after grown at the moment of t+1.

2. Each EAN in GP must have enough energy. So according to the feature1 of EAN model, $\forall EAN_i \in GP$ must meet:

$$\sum_{i=1}^{n} \bar{e} \cdot \varepsilon_{ij} < E_{AN}^{surplus}$$

A GP has two features below:

- New nodes must be created in a GP. Existing nodes in GP will become public nodes when network is growing. In addition, a new NU will be created by these public nodes and the new nodes after grown.
- The energies in EAN will reduce when the network is growing because new connections have been established. The energy decrement ΔE_{AN}^{unit} can be calculated by:

$$\Delta E_{AN}^{unit} = \sum_{i=1}^{n} \Delta E_{AN}^{consumption(i)} = \sum_{i,j=1}^{n} \bar{e} \cdot \varepsilon_{ij}$$

where $\Delta E_{AN}^{consumption(i)}$ is each energy decrement, and according to Definition 4, the ΔE_{AN}^{unit} can be simply calculated by the amount of nodes in GP:

$$\Delta E_{AN}^{unit} = \sum_{i=1}^{n} \Delta E_{AN}^{consumption(i)} = n$$

where n is the amount of nodes in GP.

Definition 8 OGP (Optimal Growing Point) is an optimal growing point at the moment of t. It is possible that there are many GPs when the network is growing. A GP which wants to become the OGP must win in the competition. The competition rule is described as below:

- The GP which will become OGP must insure that the production entropy of current network must be the minimum one among all the GPs at the moment of t+1.
- The weight of the neuron which located in the central of GP's topological graph must have the max similarity with the input data, which is similar with SOFM competition rule.

So, in another word, if any GP meet this condition below will become an OGP:

$$\Delta H_{G'_{optimal}}^{unit}(t+1) \le \Delta H_{G'_i}^{unit}(t+1)$$
$$and \ \|X - W_{C_{optimal}}\| \le \|X - W_{C_i}\|.$$

where G'_i is any GP in the moment of t, and W_{C_i} is the weight of any GP's central node in the moment of t.

Definition 9 CDN (Cover Domain Neuron) is a neuron whose weight is the best matching unit. The CDN represents the activity of a certain neuron to a corresponding input data $\boldsymbol{X_i}$. We can simply calculate the Euclidian distance between them as the similarity. If such distance is smaller than a threshold **ξ**, we can call this neuron CDN.

$$\|W - X_i\| \le \xi$$

The CDN usually reflects a hyper-sphere in the high-dimension space. The weight of the neuron is the corresponding centre of this hyper-sphere and the parameter **ξ** is the corresponding radius of this hyper-sphere. Thus, this hyper-sphere represents the space occupation of this neuron in the high-dimension space when the mapping procedure is terminated.

3.2 *The Architecture of ESGSONN*

ESGSONN has three layers in its network architecture: the Input Layer, the Growing & Competition Layer and the Combine & Output Layer.

- Input Layer is composed of several input nodes, which is similar with SOFM.
- Growing & Competition Layer is made up by EANs. It will take an unsupervised learning and EANs participate in the competition here. Also, EANs will be generated or deleted in this layer.
- Combine & Output Layer contains output neurons. Each output neurons will be generated after obtaining a decision domain.

The architecture of ESGSONN is shown in Fig. 3.

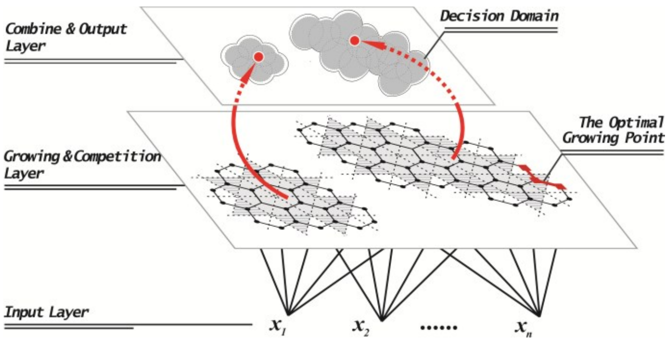


Fig. 3. The architecture of ESGSONN.

3.3 Training Algorithm

There are mainly eight steps in ESGSONN algorithm:

Step 1 Input a sample data X_i from R^n randomly. After executing step1, other sample data can be input in a certain order or randomly.

Step 2 Initialize the network by generating a TNU structure, EAN_{first}^{unit3}, which has to meet that:

$$\forall EAN_i \in EAN_{first}^{unit3} \rightarrow W_i = X_i$$

Step 3 Obtain an OGP according to the Definition 8.

Step 4 If a neuron becomes a CDN or the count of learning steps has reached to a threshold σ, the network will generate a new NU at the OGP at the moment of t. the weights of new EANs added in new NU can be calculated by:

$$W_{new} = \begin{cases} W_c & EAN_{new} \ is \ diagonal \ to \ CN \\ W_{adjacent} & EAN_{new} \ is \ adjacent \ to \ OGP \end{cases}$$

where the W_c is the weight of CN in the new TNU. And the $W_{adjacent}$ is the weight of neurons in the OGP, which is adjacent to the new ENA at the moment of t+1.

If the neuron in the central topological graph becomes a CN at the moment of t+1, the weights of neurons in OGP will be adjusted by:

$$\boldsymbol{W}_{(t+1)} = \begin{cases} \boldsymbol{W}_{(t)} + \boldsymbol{\eta}_{max} \cdot \Delta\widehat{\boldsymbol{W}}_{it} & \boldsymbol{V}_t = \boldsymbol{V}_c \\ \boldsymbol{W}_{(t)} + \boldsymbol{\eta}_{topo1} \cdot \Delta\widehat{\boldsymbol{W}}_{it} & \boldsymbol{V}_t = \boldsymbol{V}_{topo1} \\ \boldsymbol{W}_{(t)} + \boldsymbol{\eta}_{topo2} \cdot \Delta\widehat{\boldsymbol{W}}_{it} & \boldsymbol{V}_t = \boldsymbol{V}_{topo2} \end{cases}$$

where V_{topo1} and V_{topo2} are two neurons which are adjacent to the CN. In addition, $\boldsymbol{\eta}$ is the learning-efficiency, where η_{topo1} , η_{topo2} and η_{max} are three different learning-efficiency corresponding to V_{topo1}, V_{topo2} and V_c. Make sure that η_{max} is the biggest one. The $\Delta\widehat{W}_{it}$ is a unit vector towards the sample data X_i, which can be calculated as below:

$$\Delta\widehat{\boldsymbol{W}}_{it} = \frac{\boldsymbol{X}_i - \boldsymbol{W}_t}{\|\boldsymbol{X}_i - \boldsymbol{W}_t\|}$$

$$= \left(\frac{X_i^1 - W_t^1}{\sqrt{\sum_{j=1}^{n}\left(X_i^j - W_t^j\right)^2}}, \frac{X_i^2 - W_t^2}{\sqrt{\sum_{j=1}^{n}\left(X_i^j - W_t^j\right)^2}} \dots, \frac{X_i^k - W_t^k}{\sqrt{\sum_{j=1}^{n}\left(X_i^j - W_t^j\right)^2}} \right)^{\boldsymbol{T}}$$

Step 5 If a neuron becomes to a CDN, reserve its weight and generate a new NU, then input next sample until all the samples have been input into the network, or go back to the Step3.

Step 6 After all the samples have been input into the network, delete the neurons and its adjacent neurons if this neuron did not become to a CN and its adjacent neurons were not CDN neither. After the delete operating, we can obtain different clusters.

Step 7 Combine all the adjacent neurons and calculate the decision domains, D_c , which are the volumes of corresponding hyper-spheres in the high-dimensional space.

$$\boldsymbol{D}_c = \bigcup_{i=1}^{n} \boldsymbol{Vol}_{EAN_i}$$

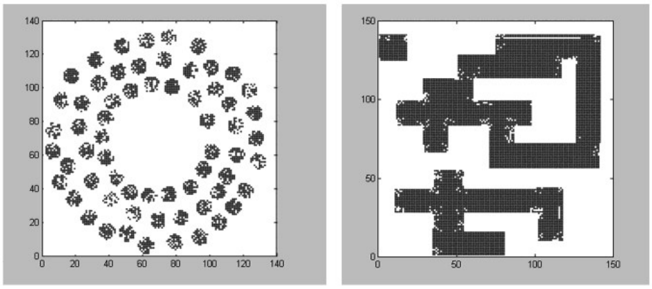
Step 8 Generate the output neurons in the Combine & Output Layer, and make sure that each output neurons has a corresponding decision domain.

4. EXPERIMENTAL RESULTS

4.1 The Stability of Dataset with Complex Structures

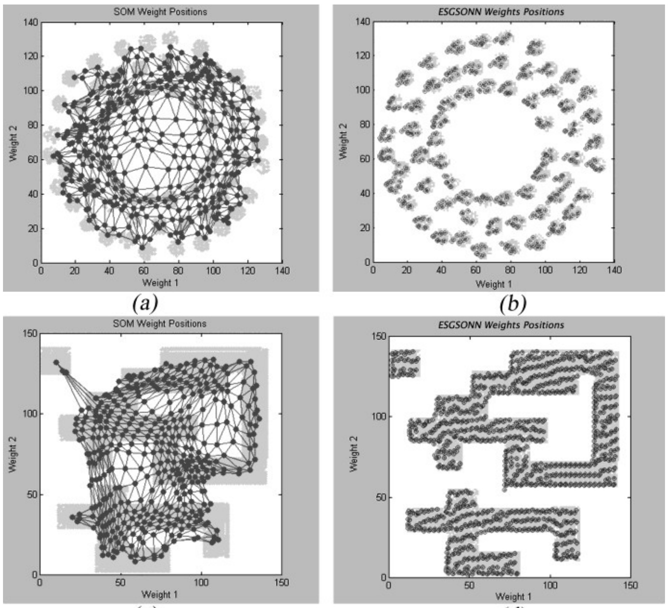
It is a general problem that if the input dataset represent a complex distribution, in this case, the network based on competitive learning algorithm will become unstable, such as SOFM's limitations. It often easily gets dead nodes that appear in the zero density-distributions of the input dataset and there exists more neurons around center than the edge. To check the stability of ESGSONN in this case, we use two groups of data which distribute in a complex structure. The first group of data is composed by 3243 samples distributing in a spherical band structure. (See Fig. 4 (a)). And the second group of data is composed by 9173 samples, which distribute as an irregular structure. (See Fig. 4 (b)).

In this experiment, we see that SFOM has lots of dead nodes outside the input data distribution. (See Fig. 5. (a), (c)) These dead nodes could hardly be activated. In addition, boundary effect is visible, there are more nodes around the center than the edges, because the nodes around the center could be easily activated and it can learn more than the nodes around the edge. However, ESGSONN could reflect the input data distribution more precisely. The experimental result of ESGSONN is shown in Fig. 5. (b), (d).



(a) (b)

Fig. 4. Distributions of two different input data. (a) The input represents as a spherical band. (b) The input distribution represents as an irregular structure.



(a) (b) (c) (d)

Fig. 5. The experiment result of stability to the complex input data distribution. (a) The result of SOFM by the group1 data. (b) The result of ESGSONN by the group1 data. (c) The result of SOFM by the group2 data. (d) The result of SOFM by the group2 data.

Since the weight of neurons in ESGSONN is updated by unit-vector in every learning-circle, there is no relationship of similarity between W_i and X_i . It is affected only by the topological structure and a few nodes in OGP will be adjusted in every learning-circle. Hence there are few dead nodes after learning.

4.2 Incremental learning

Incremental learning is a very important feature for SONNs. Usually the weights of neurons will change when they adapt new data. As a result, the network structure will be destroyed after further learning. In this experiment, we use three simple sequence groups of data and input them in order. The experimental result of SOFM is shown in Fig. 6. We see that in each incremental learning, the weight of previous network changed in every learning-circle.

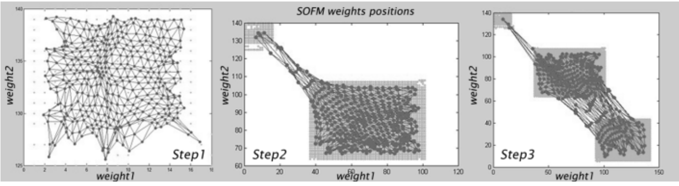


Fig. 6. The procedure of SOFM learning.

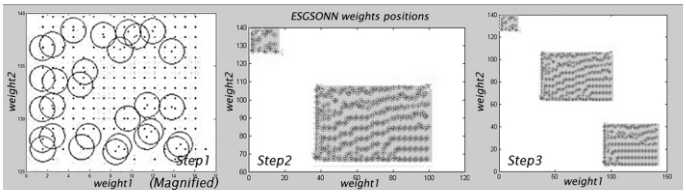


Fig. 7. The procedure of ESGSONN learning.

However, because of the algorithm of ESGSONN, each learning-circle can be seen as an incremental learning. In other word, the whole procedure of ESGSONN learning is composed of small incremental learning-circles. (See Fig. 7).

5. CONCLUSIONS

The artificial neuron model is a fundamental element in the neural computation. In this paper, we proposed a new artificial neuron model called EAN model (Energy Artificial Neuron Model) based on the energy concept from the glial cells according to the recent achievements in the neuroscience field. EAN model can provide an energy-threshold during the network growing. Additionally, we suggest a way to demonstrate EAN model in mathematics. Based on the EAN model, we realized a self-growing and self-organizing neural network called ESGSONN, which has these features as below:

1) ESGSONN has no or less dead nodes after training and it has less boundary-effect than SOFM.

2) The algorithm of ESGSONN is simple. Each growing will generate a new NU structure, and only few nodes in OGP will be adjusted. The algorithm will terminate when all the data have been input into the network. And it less depend on the initialization of the network.

3) Each learning-circle for every input data can be considered as an incremental learning.

Future work will considerate on applying EAN model in different kinds of self-organizing neural networks and further verify the features of EAN model and ESGSONN in different fields of application.

REFERENCES

D. Alahakoon, S.K. Halgamuge, B. Srinivasan. (2000). Dynamic self-organizing maps with controlled growth for knowledge discovery. *IEEE Transactions on Neural Networks*, **11**(3): 601–614

D. J. Willshaw, C. Von Der Malsburg. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **194**(1117): 431-445

Doo-Il Choi, Sang-Hui Park. (1994). Self-creating and organizing neural networks. *IEEE Transactions on Neural Networks*, **5**(4): 561-575

Eric A. Newman. (2003). Glial Cell Inhibition of Neurons by Release of ATP. *The Journal of Neuroscience*, **23**(5): 1659

Fritzke B. (1994). Growing cell structures-A self-organizing network for unsupervised and supervised learning. *Neural Networks*, **7**(9): 1441-1460

Fritzke, B. (1995). A growing neural gas network learns topologies. *In G. Tesauro, DS Touretzky. and TK Leen, editors, Advances in Neural Information Processing System*, **7**: 625-632

J. Kangas, T. Kohnen, and J. Laaksonen. (1990). Variants of self-organizing maps. *IEEE Transactions on Neural Network*, **1**(1): 93-99

Joachim W. Deitmer. (2000). Glial strategy for metabolic shuttling and neuronal function. *BioEssays*, **22**(8): 747-752

Joachim W. Deitmer. (2001). Strategies for metabolic exchange between glial cells and neurons. *Respiration Physiology*, **129**(1-2): 71-81

Joachim W. Deitmer, Johannes Brockhaus, Diana Casel. (2006). Modulation of synaptic activity in Purkinje neurons by ATP. *The Cerebellum*. **5**(1): 49-54

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, **43**(1): 59-69

Pfrieger FW, Barres BA. (1997). Synaptic efficacy enhanced by glial cells in vitro. Science, **277**(5332): 1684–1687

Paola Bezzi, Andrea Volterra. (2001). A neuron–glia signalling network in the active brain. *Current Opinion in Neurobiology*, **11**(3): 387-394

Shen F, Hasegawa O. (2006). An incremental network for on-line unsupervised classification and topology learning. *Neural Networks*, 19(1): 90–106

Slezak M. Pfrieger FW. (2006). Soltys Z. Synaptic plasticity, astrocytes and morphological homeostasis. *Physiol*, **99**: 84-91

Song H, Stevens CF, Gage FH. Astroglia (2002). induce neurogenesis from adult neural stem cells. *Nature*, **417**: 39-44

Tommy W.S, Chow, Sitao Wu. (2004). Cell-splitting grid: a self-creating and self-organizing neural network. *Neurocomputing*, **57**: 373-387

Ullian EM, Stephanie K, Sapperstein, et al. (2001). Control of synapse number by glial. *Science*, **291**: 657-661