Unsupervised Query-Based Learning of Neural Networks Using Selective-Attention and Self-Regulation

Ray-I Chang and Pei-Yung Hsiao

Abstract—**Query-based learning (QBL) has been introduced for training supervised network model with additional queried samples. Experiments demonstrated that the classification accuracy is further increased. Although QBL has been successfully applied to supervised neural networks, it is not suitable for unsupervised learning models without external supervisors. In this paper, an unsupervised QBL (UQBL) algorithm using selective-attention and self-regulation is proposed. Applying the selective-attention, we can ask the network to respond to its goal-directed behavior with self-focus. Since there is no supervisor to verify the self-focus, a compromise is then made to environment-focus with self-regulation. In this paper, we introduce UQBL1 and UQBL2 as two versions of UQBL; both of them can provide fast convergence. Our experiments indicate that the proposed methods are more insensitive to network initialization. They have better generalization performance and can be a significant reduction in their training size.**
"To teach students in accordance with their aptitude" Confucius

Index Terms— **Force-directed method, query-based method, selective attention, self-organizing maps, self regulation, unqupervised learning.**

I. INTRODUCTION

QUERY-BASED learning (QBL) algorithms have been applied to many machine-learning applications for providing correct classification output when presented with an input query [1], [2]. In the past, this algorithm has been presented for supervised learning with an additional oracle supervisor [3]–[5]. The supervisor could be the human expert, the data base of experimentation, or the computer simulation. In a supervised neural network, the correct classification output can easily be obtained after asking the external supervisor. Then, the queried data with proper input and output information can be applied to further refine the classification boundary, thereby increasing the classification accuracy [3]. Application of this supervised QBL (SQBL) algorithm has been presented to resolve various power system problems with great success [3], [7], [8]. Note that the query oracle presented in the SQBL model is a prespecified supervisor. Unfortunately, in an unsupervised neural network, i.e., Kohonen's self-organizing feature maps (SOM's) [15], and Carpenter and Grossberg's

adaptive resonance theory (ART) [45], there is no external supervisor to say what the output should be or whether the output is correct. Since the external supervisor is not valid, this SQBL algorithm cannot directly be applied to unsupervised neural networks. Thus, the development of an unsupervised QBL (UQBL) algorithm for the neural networks would be very interesting. In this paper, an UQBL algorithm based on the *behavior control theory* with *selective-attention* and *self-regulation* is proposed [9], [10].

Behavior control theory proposed by Powers in 1973 has already been shown to be of considerable value in designing artificial systems and constructing biological systems [10]. It provides the underlying basis for the elaborate computing machines that we all take for granted in human behavior. This theory suggests that our brain can realize our want and desire, so the nervous system will try to control external stimulus with selective-attention and direct to our internal desires under some self-regulation behaviors [11]. In which, selectively attending to information originating from within and concerning the internal self is referred to as *self-focus*. Selectively attending to information that originates from the external environment is termed as *environment-focus*. To sum this theory, human behavior is less static and stable than the behavior of lower animals. Thus, it is not only under the control of physiological factors, but also under the control of some psychological factors. In this paper, perception of ambiguous stimulus will rest on humans' expectations with selected *self-directive attention* [10]. The selective or directive nature is one of the well-known three aspects of the attention process presented in cognitive psychology [12]. The degree of vigilance, another aspect of the attention process, has already been successfully implemented in Carpenter and Grossberg's adaptive resonance theory (ART) model [46]. Besides, the distribution of observance, diffused or concentrated, is also one of the aspects of the attention process. Note that when a training goal is positive for the system, the distinction between *want* (self-focus from internal desire) and *need* (environment-focus from external stimulus) is collapsed. However, when a negative or threatening objective is involved, the distinction between want and need would be raised. In this, the self-focus and the environment-focus have different experiential or behavioral consequences. Thus, the internal-desired-samples with only self-focus cannot directly be used for neural-network training. In such cases, the self-regulation property of human

behavior that tries to make a compromise between need and want would be followed. In this paper, we apply the self-regulation property to construct the UQBL model.

With the behavior control theory, the proposed UQBL algorithm can be applied to generate the *real-used-samples* from the *external-input samples* (environment-focus) and the *internal-desired-samples* (self-focus). Considering a nervous system with the *goal-directed behavior* as shown in [10], we can ask the system to respond to its internal-desired-samples with goal-directed selective-attention. As Confucius said, "*To teach students in accordance with their aptitude.*" Because there is no external human expert to verify the correctness of the queried data, these internal-desired-samples cannot be used directly. They should be taken to make a compromise with self-regulation to the external-input-sample. Note that UQBL is not an anthropomorphic model that disregards the effect from the external stimulus. On the contrary, it tries to combine the effect from the external stimulus and the internal desires, and has shown that both the external stimulus and the internal desires are important and decisive for network training. By considering the internal-desired-sample, what is selected for attention may change from moment to moment and only depends on the system's current configuration [10]. The produced real-used-sample has a queried (or competitive) output label for selective-attention. Since the output label is not determined by an external supervisor but the network itself, the proposed algorithm is an unsupervised approach.

In this paper, two versions of UQBL are introduced, UQBL1 and UQBL2. An application example of the proposed UQBL algorithm to SOM is also demonstrated. Comparisons have been made with SOM. Experiments indicate that the proposed method can obtain faster convergence to a self-organized state and has a final result similar to that of the conventional approach. The obtained results are insensitive to different network initialization and can be a significant reduction in the training set. The organization of this paper is structured as follows. In Section II, the previous works of QBL are reviewed. The UQBL algorithm with selective-attention and self-regulation is described in Section III. An application example to design the query-based self-organizing feature map is demonstrated in Section IV. Experimental results and comparisons are shown in Section V with some discussions. Finally, Section VI gives our conclusion and future works.

II. REVIEW OF QUERY-BASED LEARNING METHODS

Define the training samples as pairs $[\boldsymbol{x}, o(\boldsymbol{x})]$, where $\boldsymbol{x}$ is an input vector and $o(\boldsymbol{x})$ is the target output vector of input $\boldsymbol{x}$. Assume that the source of training samples can be simply modeled as a query oracle. It can give the correct output when queried with an input. Thus, when a y is supplied, $o(y)$ would be told by the query oracle. This additional training sample $[y, o(y)]$ is called the *queried sample*. In 1991, Baum [4] had shown that the QBL paradigm corresponded more closely to the way humans learn. This method does not only look at the original training examples, but also utilizes queries to provide additional training samples and is then told which output the input vector is assigned. The presented superfluous query power is practical in many classification problems

where the algorithm can produce additional inputs and be instructed by an external supervisor to what classification outputs they correspond. Thus, the classification system can be further refined by the queried samples. Expansion of the QBL paradigm to include membership query was proposed by Valiant [1], and has been subsequently studied by Angluin [2]. In this paper, the problem of using queries to learn an unknown concept was considered.

Recently, the neural network community has focused on learning from input samples and queries [3]–[5], [7]. Considering multilayer propagation (MLP) with the backpropagation (BP) algorithm proposed by Rumelhart *et al.* [13], in which the output of neuron is binary trained to be either zero or one. Presenting a set of training samples with prespecified labels, the classification boundary of the neural system is defined as the set of points which produces an output of $(0+1)/2(=0.5)$. Using the points on the classification boundary, called *boundary points,* a set of conjugate input pairs with significant boundary information can easily be generated to refine the classification result. In [4], the boundary point that has maximum ambiguity $(=0.5)$ was simply produced by the interpolation process between positive and negative examples. Besides, the inversion method which allows a user to find one or more input vectors yields a specific output vector and can also be utilized to generate these boundary points [3], [7], [8]. Since the inversion algorithm is a time-consuming iterative procedure, a genetic-based approach used as a means of achieving neural network inversion was presented [5]. A simple example with two different input samples is shown in Fig. 1 to illustrate the operations of SQBL. As shown in Fig. 1(a), it can be found that the presented input samples A and B are linear-separable as the bold dash-line, but the real classification boundary is not (see the dot-line). Thus, although the training time is unlimited and the training error for these input samples is down to zero, the classification error for test data is still very large. As shown in Fig. 1(c), it can be found that the linear-separated classification boundary is refined as an S-type boundary when the boundary point and the two additional queried samples C and D shown in Fig. 1(b) are presented. Comparing the original classifier with a linear-separated boundary, the obtained classification accuracy is further increased.

In conventional QBL algorithms, the query oracle is defined as a prespecified supervisor. Although they have shown to be very efficient in SQBL, the prespecified external supervisor is not existent in unsupervised network models. Note that the unsupervised learning method has the advantage that it can automatically classify input vectors without specifying their output labels. The design of a UQBL algorithm which can learn from examples and queries without specifying their output labels would be a very attractive and significant research topic. In this paper, the behavior control theory that introduces the properties of selective-attention and self-regulation is presented to design a UQBL algorithm.

III. UNSUPERVISED QUERY-BASED LEARNING WITH SELECTIVE-ATTENTION AND SELF-REGULATION

Neural-network models can be divided into two principal categories, supervised and unsupervised, according to training
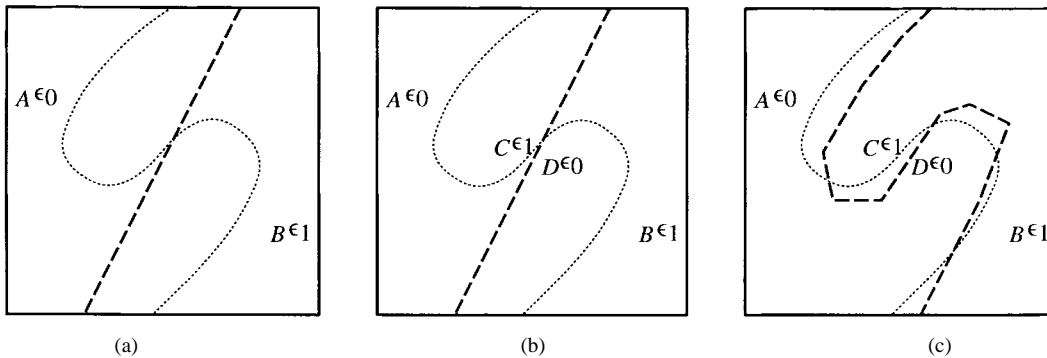
Fig. 1. A simple example with two different input samples is presented to demonstrate the operation flow of this SQBL method. (a) The original input samples (A in class 0 and B in class 1) are linear-separable, but the real classification boundary is not. Although the training error is already zero, the classification accuracy for the test set is still very low. (b) Two additional queried samples (C in class 1 and D in class 0) are produced. (c) The classification boundary is refined as an S-type boundary. The classification accuracy is further increased.

with or without an external supervisor. Unsupervised neural networks can automatically group input patterns into several clusters, such that without prespecifying the target output, each input pattern can be assigned to a unique cluster label with the presented adjustment rule. Assume that a set of vectors V which is drawn from some probability distribution $p(V)$ is defined as follows:

$$V = \{V_1, V_2, \cdots, V_n\} \tag{1}$$

where n is the number of training samples. The data point selectively attending to information of the external environment is called *environment-focus* [10]. In this paper, this input vector $V_i = (v_{i1}, v_{i2}, \cdots, v_{ir})$ is called the external-input-sample which comes from the r-dimensional sample space. Assume that the architecture of the neural-network model can be simply presented by a set of weight patterns as follows:

$$W = \{W_1, W_2, \cdots, W_m\} \tag{2}$$

where m is the number of neurons. The weight pattern $W_i = (w_{i1}, w_{i2}, \cdots, w_{ir})$ that comes from the connection weights between neuron i and its input neurons is called the state of neuron i. It changes from moment to moment during neural-network training.

In unsupervised learning, as there is no supervisor to say what the output should be, the output vector must be coded by weight patterns W and the input data V. Note that there are close connections between neural networks and standard statistical techniques of pattern classification and analysis [14]–[16]. Thus, each unsupervised learning system can be simply modeled by a well-defined quality $E = E(V, W)$ to be minimized. As the training set V is prespecified, we can rewrite E as $E(W)$. Assume that E is differentiable regarding W_i, using the iterative gradient descent method, the learning equation can be defined as follows:

$$W_i(t+1) = W_i(t) + \alpha_i(t) \times (-\partial E / \partial W_i) \tag{3}$$

where $\alpha_i(t)$ is the learning rate for neuron i. The well-defined quality E is sometimes called the cost function, the objective function, or the energy function in previous studies [13], [17], [18]. This optimization approach is closer to that of the statisticians.

Note that in the unsupervised learning model, the learning rate $\alpha_i(t)$ is usually set as a smaller value if neuron i is not the best-matched winner (or set as zero for the winner-take-all model), in which the winner neuron c is defined as the neuron that is the nearest to the input vector V_k. Considering the unsupervised competitive learning model, we have

$$\mathrm{dis}(W_c(t), Vk) = \min\{\mathrm{dis}(W_j(t), V_k), \text{ for } j = 1, 2, \cdots, m\} \tag{4}$$

where $\mathrm{dis}(\boldsymbol{x}, \boldsymbol{y})$ is a general distance measure between vector $\boldsymbol{x}$ and vector $\boldsymbol{y}$, such as $\|\boldsymbol{x} - \boldsymbol{y}\|$ or $(\boldsymbol{x} - \boldsymbol{y})^2$. Since $W_j(t)$ just represents the *current state* of neuron i at time t, this processing step is also called "*to compete with neurons' current states.*"

A. Presenting Internal Desires with Selective-Attention

In past years, many phenomena have been presented by psychologists concerning the regulation of internal states and perceptual experiences. This approach has been tried to model human behavior for the preprocessing of input stimulus from the external world by Newell and Simon [19]. A computer simulation of the way personality functions has been constructed by Loehlin [20]. In 1973, Berelson and Steiner [21] showed that people will tend to see or hear things as they want or desire to see or hear them. In their studies, hungry persons can report more food patterns in recognizing vague pictures than less hungry persons. In the same year, Powers [9] suggested that our brain can realize what we want and what we desire. The behavior control theory might be realized in the human nervous system as shown in [10], in which the nervous system is not only learned from environment-focus (external-input samples), but also self-focus (queries of internal desires).

Since the system could improve itself with selected *self-directive attention,* called the *goal-directed behavior* [10], we can ask the nervous system to respond to the question: "What do you want to learn?" (or "What do you want to be?"). The response of this question, which selectively attends to the information of the internal self, is called *self-focus.* It can be defined as a set of labeled data points U

$$U = \{(U_1, 1), (U_2, 2), \cdots, (U_m, m)\}. \tag{5}$$

The data point $U_i = (u_{i1}, u_{i2}, \cdots, u_{ir})$ is called the *internal-desired-sample* of neuron i. At each time step t, we can simply view that $W = W(t)$ as a constant matrix before network learning. Thus, the energy function E can be written as $E(V)$. Assume that E is differentiable regarding V_i (or U_i). By taking advantage of the duality between weights and the input vectors in minimizing E, the internal-desired-sample U_i can be simply defined as follows:

$$U_i(t+1) = U_i(t) + \beta_i(t) \times (-\partial E/\partial U_i) \qquad (6)$$

where $\beta_i(t)$ is the gain term for neuron i. In SQBL [3], this equation is applied to achieve *good* training samples. In this paper, we applied this idea to obtain the internal-desired-samples. The achieved vector U_i which represents "what the neuron i wants to learn" is the same as the answer of "what the neuron i wants to be." It can be also called "the *desired state* of neuron i" as the target desire of W_i.

B. Compromising Need and Want with Self-Regulation

In SQBL, the correctness of the queried samples can be checked and guaranteed by an external supervisor [3]. Unfortunately, as there is no external supervisor in UQBL, the correctness of the queried internal-desired-samples cannot be guaranteed. The distinction between want and need is collapsed if the goal is positive. However, when a negative or threatening objective is involved, called *negative feedback loop* [21], the distinction between want and need would be raised. Since these queried samples may have a negative objective to the original input samples, they cannot be directly applied for neural network training. The regulation procedure that finds a middle ground between want (internal-desired-sample) and need (external-input-sample) should be applied [10].

Although various paradigms can be designed for the self-regulation property of the UQBL, in this paper, the *real-used-samples* are simply defined as the prespecified external-input samples with their self-regulated labels. Note that in the unsupervised neural networks, the reliable information we can achieve is the external-input samples and the weight patterns only. These real-used-samples can be defined as follows:

$$V' = \{(V_1, o_1), (V_2, o_2), \cdots, (V_n, o_n)\} \qquad (7)$$

where input vector V_i corresponds to output label o_i. In other words, neuron o_i is active when V_i is input. In (6), we have applied the weight patterns to produce the internal-desired-samples. Since U_i is defined as the internal-desired-sample of neuron i, neuron i must be the winner when U_i is input. Now, presenting an external-input-sample V_k, we want to find the nearest internal-desired-sample U_{o_k} to provide the real-used-sample (V_k, o_k) for network training.

Since the input samples are unlabeled in an unsupervised learning model, neurons should be trained with the *competitive rule*. Competitive learning is the main tool for training without supervision. In this paper, we follow this competitive rule to query the output labels. Assume that an external-input-sample V_k is presented, then the corresponding active neuron o_k can be defined as

$$\text{dis}(U_{o_k}, V_k) = \min\{\text{dis}(U_j, V_k), \text{ for } j = 1, 2, \cdots, m\} \quad (8)$$

where U_j is the internal-desired-sample of neuron j. Cognitive research on this phenomenon has been presented by Duval and Wicklund [22]. It provided the demonstration that persons matched their internal behavior to specified external reference samples. As shown in previous sections, U_j is also called "the *desired state* of neuron j." Thus, the competitive equation described above can be viewed as "*to compete with neurons' desired states.*" It is different from the traditional algorithms to compete with neurons' current states [see (4)].

C. Two Versions of Unsupervised Query-Based Learning

The behavior control theory with selective-attention and self-regulation can be successfully applied for UQBL. The UQBL method with an iterative gradient descent method to produce the internal-desired-sample has been shown in (6). However, this iterative method will take more computation time and more memory space to store temporary results. In order to avoid these disadvantages, a small modification of the internal-desired-sample was made. Since the internal-desired-sample U_i is just the target desire of W_i, we can simply assume that $U_i(t) = W_i(t)$ without loss of generality. Thus, the definition of the internal-desired-sample can be simply rewritten as follows:

$$U_i(t+1) = W_i(t) + \beta_i(t) \times (-\partial E/\partial U_i). \qquad (9)$$

This version of the presented procedure, which produces the internal-desired-sample without using time-consuming iterative computations, is called UQBL1. Note that, if t is large and $\beta_i(t) = 0$ (said "the system is tired"), the proposed UQBL1 method is the same as the original unsupervised learning method as $U_i(t+1) = W_i(t)$ in (8). Thus, the characteristics of the original unsupervised learning method can be guaranteed.

Note that the "optimization" described in UQBL1 only means local gradient steps. When we apply the deepest descent in the continuous E landscape, the internal-desired-sample of neuron i can be defined as a vector U_i to fulfill the following conditions:

$$\partial E/\partial U_i = 0 \quad \text{and} \quad \partial E/\partial U_i \partial U_i > 0. \qquad (10)$$

UQBL with the deepest descent method to obtaining the internal-desired-sample is the second version of the proposed UQBL method, called UQBL2. Descriptions about the optimization methods which can be applied for the generation of the internal-desired-samples are shown in [23]. Fig. 2 shows the system architectures of the conventional unsupervised learning model and the proposed UQBL model. Note that the internal-desired-samples are generated with selective-attention. Then, these internal desires are trying to make a compromise to the external-input-sample with self-regulation. In this paper, a simple example is presented to design a QBL algorithm for Kohonen's SOM.

IV. AN APPLICATION EXAMPLE TO KOHONEN'S SELF-ORGANIZING MAPS

SOM proposed by Kohonen [15] is an ordered mapping of a high-dimensional input space which maps onto a low-dimensional discrete neuron topology. This dimensional re-
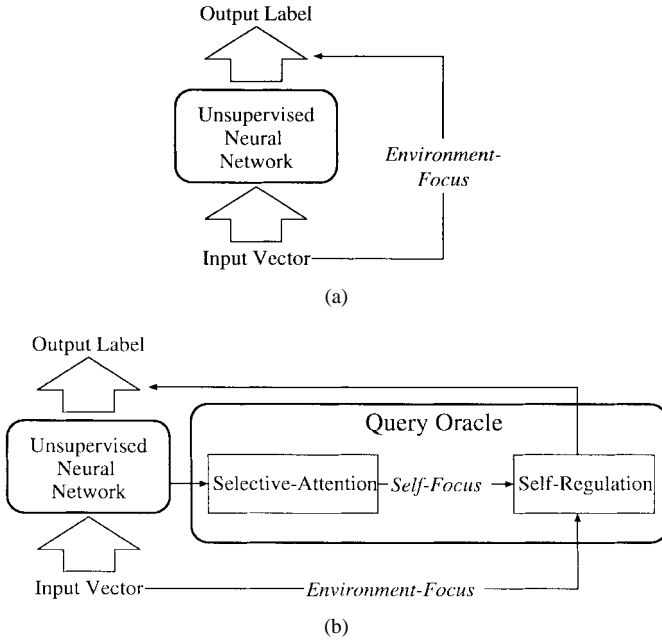
Fig. 2. (a) The system architecture of conventional unsupervised learning models. (b) The system architecture of the proposed UQBL model. The internal-desired-samples (neurons' self-focus) are first generated with selective-attention. Then, the winner neuron is obtained by compromising self-focus and environment-focus (external-input sample) with self-regulation. It is different from the conventional self-organization algorithm, which learns with only environment-focus.

duction allows us to easily visualize important relationships among the data. It is considered one of the most powerful methods that creates topographic maps. In the past, many different applications have been presented with great achievement [24]–[30]. In this paper, an application example of the proposed UQBL algorithm to SOM is presented. In addition, some particular applications of the proposed learning model to self-organize with only internal-desired-samples have also been illustrated.

A. Query-Based Self-Organizing Feature Maps

The objective of SOM is to competitively find the best matching winner and adapt the weights between input and output neurons, so that the output neurons become sensitive to different inputs in an organized manner. Let v be the data point in a set of training samples. The "winner" neuron c is defined as (4), in which the input vector v is presented. In [18], the global objective function of SOM has been defined as follows:

$$E = \int \sum_{j=1}^{m} \eta_{cj}(t) f(\mathrm{dis}(v, W_j)) p(v) \, dv \qquad (11)$$

where function $f(\cdot)$ of quantization error $\mathrm{dis}(v, W_j)$ is first weighted by the neighborhood function $\eta_{cj}(t)$ and then averaged. As the training samples are presented in a stochastic manner, we can rewrite the objective function E as the sample function $E(t)$ following the Robbins–Monro stochastic approximation method [31]. Assume that the input vector $v(t)$ presented at time t is V_k. The sample function $E(t)$ can be defined as follows:

$$E(t) = \sum_{j=1}^{m} \eta_{cj}(t) f(\mathrm{dis}(V_k, W_j(t))). \qquad (12)$$

For example, if $\mathrm{dis}(V_k, W_j(t))$ denotes the Euclidean distance between V_k and $W_j(t)$, and function $f(x) = x^2$. Using the iterative gradient descent method described in (3), the learning equation can be obtained as

$$W_i(t+1) = W_i(t) + \alpha_i(t) \times 2\eta_{ci}(t)(V_k - W_i(t)) \qquad (13)$$

where $\alpha_i(t)$ is the time-decreasing learning rate. For example, we can simply define $\alpha_i(t)$ as the variable $(\alpha_i(0) - t \times \varepsilon)$ where $\alpha_i(0)$ is a constant value between zero and one.

The above technique has already been studied with mathematical rigor, and, in particular, the convergence properties are known [18], [32]. In this paper, we apply the same method to generate the internal-desired-sample U_i for neuron i. Considering (9) and (12), the two versions of the internal-desired-sample U_i of SOM can be defined as follows.

For UQBL1

$$U_i(t+1) = W_i(t) + \beta_i(t) \times 2\sum_{j=1}^{m} \eta_{ij}(t)(W_j(t) - W_i(t)). \qquad (14)$$

For UQBL2

$$U_i(t+1) = \sum_{j=1}^{m} \eta_{ij}(t) W_j(t) \Big/ \sum_{j=1}^{m} \eta_{ij}(t) \qquad (15)$$

where $\beta_i(t)$ is a gain term. Since the gain term $\alpha_i(t)$ is called the learning rate for the learning equation (13), we can call the gain term $\beta_i(t)$ as the *querying rate* for the querying equation. Using the self-regulation technique described in (8), the real-used-sample V_k with output label o_k can be easily obtained. The step-by-step description of the proposed UQBL algorithm for the self-organizing feature map is shown as follows.

1) Initialize weights.
2) Generate internal-desired-samples.
3) Present a new external-input-sample.
4) Calculate distance from external-input-sample to all internal-desired-samples.
5) Select winner neuron with minimum distance.
6) Update weights of winner neuron and its neighbors.
7) If system is not convergent, then go to Step 2).

The proposed network model is called query-based SOM (QBSOM). Examples demonstrating the produced internal-desired-samples for these two versions of QBSOM are shown in Fig. 3. Note that if t is large and $\beta_i(t) = 0$, QBSOM1 is the same as SOM. Thus, the beautiful system properties of SOM can be kept. However, the deepest descent QBSOM2 method cannot do this. The geometric meaning of the presented QBSOM is demonstrated in Fig. 4 with the example presented in Fig. 3. It can be found that the perception of input stimulus will rest on neurons' expectations as the way humans learn. In other words, the network is organized with neurons' desired states.
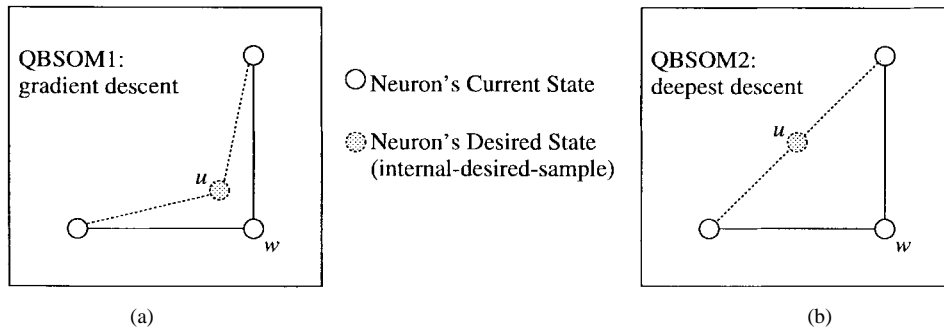
Fig. 3. (a) A simple example is presented to demonstrate the produced internal-desired-sample u of neuron w with QBSOM1. (b) The produced internal-desired-sample u of neuron w with QBSOM2 is shown. Note that QBSOM1 is the same as SOM if t is large. Thus, the beautiful properties of SOM can be kept. However, the deepest descent QBSOM2 method cannot.
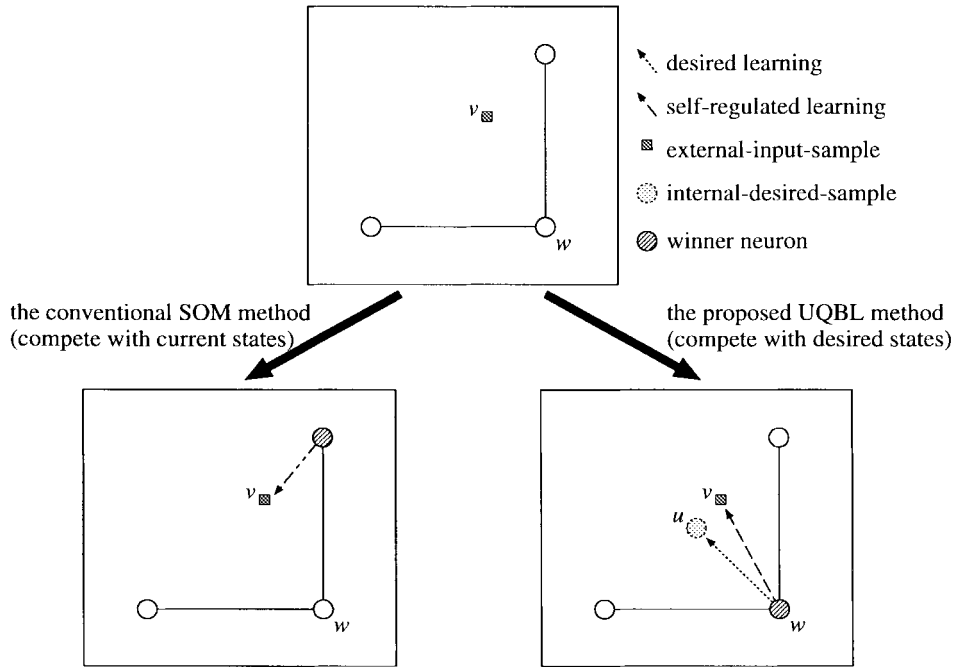

Fig. 4. The geometric meaning of the presented self-regulation property is demonstrated with the simple example presented in Fig. 3(b). It can be found that the perception of input stimulus v will rest on neurons' expectations u as the way humans learn.

B. Self-Organizing with Only Internal-Desired-Samples

Note that the proposed method self-organizes not only with external samples, but also with internal desire. Thus, the neural system could self-organize with its internal desire if the external samples are not valid. This self-organization power is really useful in many application problems where the external samples are not exist or are hard to obtain. When SOM is being trained without external samples, we can simply use the internal-desired-sample to correspond to the input training vector without loss of the generalization of behavior control theory. The original self-organizing algorithm cannot do this. Assume that the parameter $S_{ij}(t) = \beta_i(t) \times 2\eta_{ij}(t)$ is viewed as the string connected between element i and element j. It can be found that the query equation described in (14) is similar to the *force function* described in [33]–[35] for resolving the optimization problem

$$\sum_{j=1}^{m} S_{ij}(t) \times (W_j(t) - W_i(t)). \tag{16}$$

It represents the force between two spring-connected elements with positions W_j and W_i [36].

As the force-directed method has already been applied for many optimization problems, i.e., cell placement, scheduling, and robotics control [33]–[37], we can easily extend this force-directed QBSOM1 method to resolve these problems. For example, in cell placement, η_{ij} can be defined as a function of cm_{ij} (the number of wire connections between cells i and j) to represent the topology connectivity between neurons i and j. In such case, the objective function of the cell placement problem can be written as follows:

$$\sum_{i=1}^{m} \sum_{j=1}^{m} \eta_{ij}(cm_{ij}, t) f(\mathrm{dis}(W_i, W_j)). \tag{17}$$

Since the network query function shown in (14) is similar to the force-directed optimization function described in (16), this kind of query-based network model is called "force-directed self-organizing maps" [38]. The method presented can be

TABLE I
THE OBTAINED MSE AND TPG FOR SOM, QBSOM1, AND QBSOM2 USING INPUT
SAMPLES: U-SET AND G-SET. BOTH MESH AND HCUBE NETWORKS ARE TESTED

MSE		SOM	QBSOM1	QBSOM2
MESH	U-SET	8.7643×10^{-3}	8.5935×10^{-3}	9.0647×10^{-3}
		1.5159×10^{-2}	1.4728×10^{-2}	3.0464×10^{-2}
	G-SET	8.4693×10^{-3}	8.3245×10^{-3}	8.9242×10^{-3}
		1.0508×10^{-2}	1.0501×10^{-2}	2.8260×10^{-2}
TPG		SOM	QBSOM1	QBSOM2
HCUBE	U-SET	3.3498×10^{-1}	3.3635×10^{-1}	3.2918×10^{-1}
		1.5618×10^{-0}	1.5482×10^{-0}	1.3694×10^{-0}
	G-SET	2.6291×10^{-1}	2.6316×10^{-1}	2.5684×10^{-1}
		1.3537×10^{-0}	1.3432×10^{-0}	0.8505×10^{-0}

easily applied to the optimization problem which has the same type of objective function.

For example, consider a pattern taxonomy problem in which the similarity between patterns are given and the objective was to sequence the individuals on a line into a homogeneous group. Let cm_{ij} denote the similarity between patterns i and j. This pattern taxonomy problem with the objective function described in (17) can be resolved. Assume that the weight patterns of neuron networks are applied to represent the position of work centers in a job shop, and cm_{ij} is defined as the quantity of the job flow between work centers i and j. This job shop problem can also be resolved by the proposed QBSOM method. The application of the proposed method to different cell placement problems has been shown in [39] and [40]. Experiments show that the obtained results are better than those of the conventional self-organizing algorithms presented in [30], [41], and [42].

V. EXPERIMENTAL RESULTS

Our comparison below is based on training neural network models (including SOM, QBSOM1, and QBSOM2) with input samples until termination occurs; and then considering different factors to give a quantitative evaluation of the effectiveness of the proposed approach. These factors include generalization for training size requirement, robustness with initialization independence, and solution quality of the experimental results. In this paper, both the mean square error (MSE) and the degree of topographicity (TPG) [43] are computed. The function of MSE is defined as follows:

$$\text{MSE} = \sum_{i=1}^{n}\sum_{j=1}^{m}\phi_{ij}\times\text{dis}(V_i, W_j)^2/n \qquad (18)$$

where $\phi_{ij} = 1$ if input sample V_i is assigned to cluster W_j with the nearest prototype rule, otherwise $\phi_{ij} = 0$. The degree of topographicity, TPG, is defined as the mean of average distances from a neuron to its adjacent neurons as follows:

$$\text{TPG} = \sum_{i=1}^{m}\sum_{j=1}^{m}\varphi_{ij}\times\text{dis}(W_i, W_j)/m \qquad (19)$$

where $\varphi_{ij} = 1$ if $i = j$ or W_i is the adjacent neuron of neuron W_j, otherwise $\varphi_{ij} = 0$; it is applied to measure the quality of network map's self-organization.

The first training set that contains 5000 input vectors with a (pseudo) uniform rectangular distribution is called U-SET. To make it easier to understand, we suppose that inputs v_{i1} and v_{i2} for each two-dimensional (2-D) input sample $V_i = (v_{i1}, v_{i2})$ has uniform distributions from -1.0 to 1.0. The second training set that contains 5000 input vectors with a (pseudo) Gaussian distribution is called G-SET. In this data set, the inputs v_{i1} and v_{i2} of each sample V_i are mutually independent and have Gaussian density with zero mean and deviations of 0.3. Comparisons have been made with the conventional self-organizing algorithm using a simple 2-D 10×10 mesh-connected network model, called MESH, and a complicated six-dimensional $2 \times 2 \times 2 \times 2 \times 2 \times 2$ hyper-cube network model called HCUBE. Denote MT as the maximum iteration number applied. The learning rate $\alpha_i(t)$ is defined as follows:

$$\alpha_i(t) = \alpha_i(0) \times (1 - t/MT). \qquad (20)$$

The simulation programs [44] are written in C language on a Sun Sparc-10 workstation.

A. Solution Quality of Proposed UQBL Algorithms

In this paper, we first evaluate performance of the proposed UQBL algorithms through the MESH network and the HCUBE network with two different input distributions, U-SET and G-SET. The obtained MSE and TPG measurements for SOM, QBSOM1, and QBSOM2 are shown in Table I. Note that these tests all use the same initial weight pattern (between -1.0 and 1.0), and all execute with 5000 iterations. It can be found that QBSOM1 is better than SOM for both the simple MESH network and the complicated HCUBE network. Considering the test results for U-SET with the MESH network, the obtained MSE is smaller than that of SOM. Moreover, the organized network map obtained by QBSOM1 is closer to the optimal organized network map (TPG = 0.36). In this optimal case, neurons are evenly distributed on the sample space and the network map is
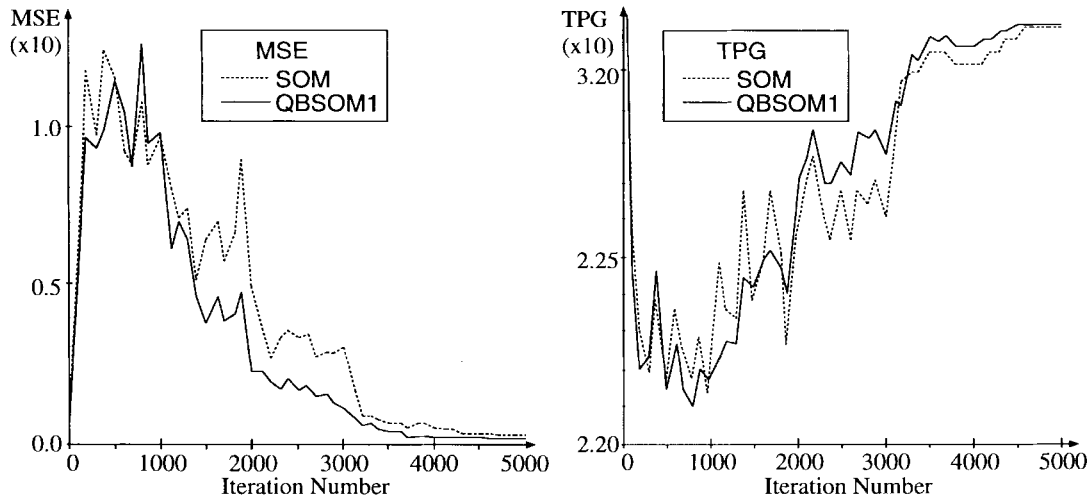
Fig. 5. The illustration of the obtained MSE and TPG versus the iteration numbers for the proposed QBSOM1. At the first learning stage, the network map will organize the network map as the self-focus is raised. In such case, the obtained MSE is increasing, but the presented TPG is decreasing. However, since the querying rate is a decreasing gain term, the environment-focus of the network map will be raised. In this stage, the obtained MSE is decreasing, but the presented TPG is increasing.

organized. The conventional SOM method also cannot obtain this optimal organized network map. Considering the test results for U-SET with the HCUBE network, the obtained MSE is also smaller than that of SOM with 3% decreases. When testing with QBSOM2, since neurons with the deepest descent learning method will shrink too much to the mean of input distribution, the network map needs more computation time to expand out. Thus, the obtained MSE for QBSOM2 and MESH network is larger than SOM (with 3% increases) and the obtained TPG is smaller than QBSOM1 (with 2% decreases). It also demonstrates that the external-input samples are important and decisive for network training. The overconfidence for the effect of internal-desired-samples, such as QBSOM2, usually misleads the pattern clustering process and introduces more classification error. Consider the obtained results for nonuniform distributed G-SET. The increasing rate of MSE obtained by SOM and QBSOM2 is larger than that obtained for the uniform distributed U-SET. It is due to the fact that the neurons at the boundary learn more than the center neurons do, called the boundary effect. It is the major disadvantage of QBSOM2.

The illustration of MSE and TPG obtained versus iteration numbers for QBSOM1 is shown in Fig. 5. Comparisons are made with SOM with the MESH network. Our experiments indicate that QBSOM1 is better than SOM for system convergence. In some of the learning process, the variations of MSE are high at the early learning stage as the neighboring range is large, and gradually converge to a small value as the neighboring range is small. The initial weights of the presented MESH network map are set as random values between $(-1.0, -1.0)$ and $(1.0, 1.0)$ as shown in Fig. 6(a). The unorganized network maps with and without connections are shown. At the first learning stage, the network map will try to organize the network map as the self-focus is raised. In such case, the obtained MSE is increasing, but the presented TPG is decreasing as illustrated in Fig. 5. However, since the querying

rate is a decreasing gain term in QBSOM1, the environment-focus of the network map will be raised. Then, the neurons will try to represent the input samples with output nodes. In this stage, the obtained MSE is decreasing, but the presented TPG is increasing. Fig. 6(b) shows the obtained results of QBSOM1 with and without network connections. It has smaller MSE than that of SOM. The obtained result with minimum TPG using the deepest descent QBSOM2 method is shown in Fig. 6(c). Fig. 7 shows the obtained MSE and TPG of the deepest descent QBSOM2 method versus iteration number. The obtained MSE and TPG of SOM are also illustrated with dot lines for the comparisons. Note that the obtained TPG for the deepest descent QBSOM2 method is decreasing very fast at the early learning stage, and needs more computation time to expand out the network map. Thus, the obtained MSE is larger than that of SOM.

B. Robustness for Different Network Initialization

Stability of competitive learning algorithms largely depends on the initial weight pattern of the network model. In this paper, we have tested SOM, QBSOM1, and QBSOM2 with 100 randomly generated weight patterns to examine the robustness of the proposed UQBL approaches. Throughout these investigations, the training set U1000 and the network model MESH are presented with various initialization states for investigating the effect of different initializations in which U1000 is a subset of U-SET with 1000 input samples. For each of the initialization states, 5000 iterations are executed. The obtained results are shown in Fig. 8, in which the measurements of MSE and TPG are represented by the 2-D points (TPG, MSE). Our experiments indicate that both QBSOM1 and QBSOM2 are insensitive for different initialization of weight patterns. It can be found that all the obtained results of QBSOM1 are terminated at a small range of (TPG, MSE) points, whereas SOM obtains different TPG and MSE for different network initializations. Note that the obtained (TPG, MSE) results for
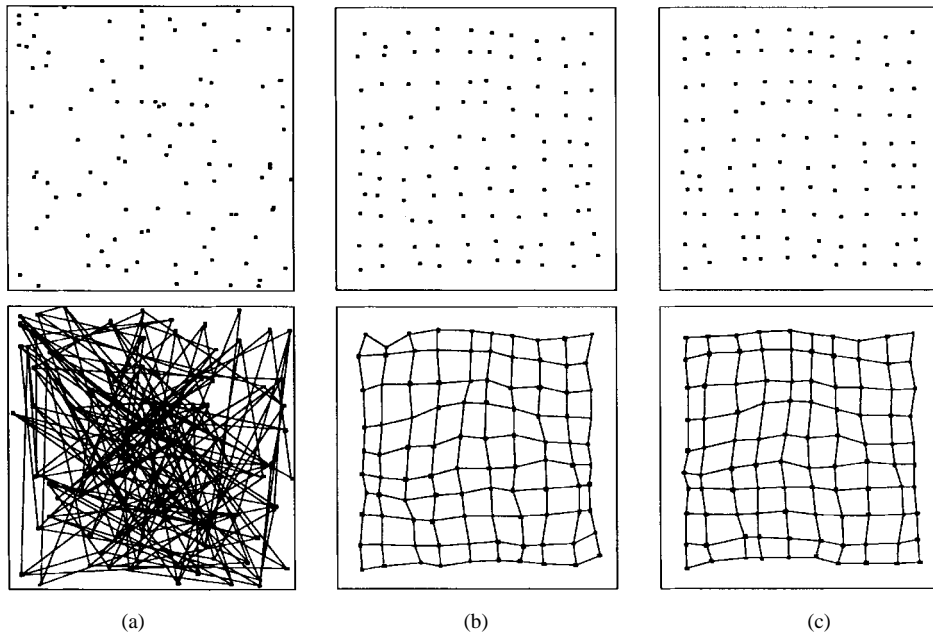
Fig. 6. (a) The initial weights of the presented MESH network map are set as random values between $(-1.0, -1.0)$ and $(1.0, 1.0)$. (b) The obtained learning results with minimum MSE (by QBSOM1). (c) The obtained learning results with minimum TPG (by QBSOM2). The network maps, with and without network connections, are shown.
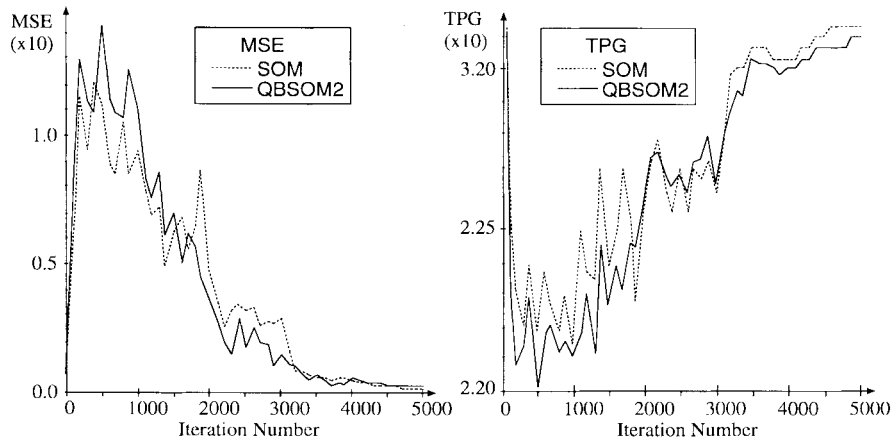

Fig. 7. The obtained MSE and TPG of the deepest descent QBSOM2 method versus iteration number is illustrated. Comparisons are also made with the MSE and TPG obtained for SOM. Note that the obtained TPG for the deepest descent QBSOM2 method is decreasing very fast at the early learning stage, and needs more computation time to expand out the network map.

QBSOM2 are all eliminated at nearly the same learning result. However, since the effect of self-focus is too large to make a compromise with the original environment-focus learning (like a spoiled child), the obtained MSE is larger than that obtained with SOM. Table II shows the mean and the variance of the obtained MSE and TPG for three various learning models using the MESH network in which G1000, a subset of G-SET with 1000 input samples, is also tested. It can be found that the variance of the obtained results is smaller than those obtained by SOM. We have also initialized the network weights at each one of the corners of the sample space $[(1, -1), (1, 1), (-1, -1),$ and $(-1, 1)]$. They are far away from the center of the input distribution and outside the convex hull of the input samples. However, the obtained results of the proposed query-based methods are nearly the same. Since the proposed UQBL algorithm adapts the network map not only with the external-input samples but also with the internal-desired-samples, the obtained results will not depend only on the external-input samples as those of SOM. It is robust for different network initialization.

C. Generalization Performance for Various Training Size

Among the many competitive learning algorithms, SOM is considered to be powerful in the sense that it not only clusters the input patterns adaptively, but also organizes the output neuron topographically. It is different from the classical pattern recognition problem which has no topological relations to be organized. In SOM, long computational time is necessary for

TABLE II
THE MEAN AND THE VARIANCE OF THE OBTAINED MSE AND TPG FOR SOM, QBSOM1,
AND QBSOM2 USING THE MESH NETWORK AND THE INPUT SAMPLES: U1000 AND G1000

			SOM	QBSOM1	QBSOM2
U1000	MSE	mean	8.9094×10^{-3}	8.7737×10^{-3}	9.3293×10^{-3}
		variance	1.8423×10^{-7}	7.8576×10^{-16}	7.7684×10^{-17}
	TPG	mean	3.2941×10^{-1}	3.3054×10^{-1}	3.2531×10^{-1}
		variance	2.5275×10^{-5}	4.5685×10^{-14}	2.2206×10^{-15}
G1000	MSE	mean	7.5431×10^{-3}	7.0769×10^{-3}	7.8259×10^{-3}
		variance	3.2164×10^{-5}	2.4376×10^{-15}	1.2673×10^{-15}
	TPG	mean	2.5955×10^{-1}	2.6139×10^{-1}	2.5488×10^{-1}
		variance	4.3541×10^{-4}	3.5564×10^{-13}	3.5593×10^{-13}

TABLE III
THE GENERALIZATION PERFORMANCE OF THE AVERAGE DMSE OBTAINED FOR SOM, QBSOM1,
AND QBSOM2 USING U-SET AND G-SET. BOTH MESH AND HCUBE ARE TESTED

DMSE		SOM	QBSOM1	QBSOM2
MESH	U-SET	0.9252×10^{-3}	0.9158×10^{-3}	0.8943×10^{-3}
	G-SET	1.1095×10^{-3}	1.0692×10^{-3}	1.0267×10^{-3}
HCUBE	U-SET	1.2980×10^{-3}	1.0814×10^{-3}	1.0406×10^{-3}
	G-SET	1.3654×10^{-3}	1.2634×10^{-3}	1.2212×10^{-3}

large sizes of a training set to organize the network map. If the training size is too small or the training time is too short, the obtained result might be a twist network map. Moreover, the generalization performance of the network map that trains with a smaller data set to test by a larger data set is not acceptable. In this paper, we assume DMSE as *"the difference of MSE"* between the training results and the testing results

$$DMSE = (MSE \text{ for the testing set})$$
$$- (MSE \text{ for the training set}). \quad (21)$$

It is defined as a kind of measurement for the generalization performance of the network map. The illustrations of DMSE for different training sizes (from 300 to 5000 input samples) and different training methods are shown in Fig. 9. Throughout these experiments, the sample set U-SET with 5000 sample points is tested. It can be found that the obtained DMSE of the proposed query-based methods is inversely proportional to the size of training set. However, it is smaller than that obtained by SOM. Furthermore, the obtained DMSE for the deepest descent QBSOM2 method is smaller than that of the gradient descent QBSOM1 method. Table III shows the generalization performance for the average DMSE obtained for three various learning models. In this paper, both the MESH network and the HCUBE network are tested. It can be found that the proposed UQBL algorithm has good generalization properties for neural network training. However, note that small DMSE does not mean that the obtained MSE is also small. For example, QBSOM2 has the largest MSE although its obtained DMSE is the smallest one.
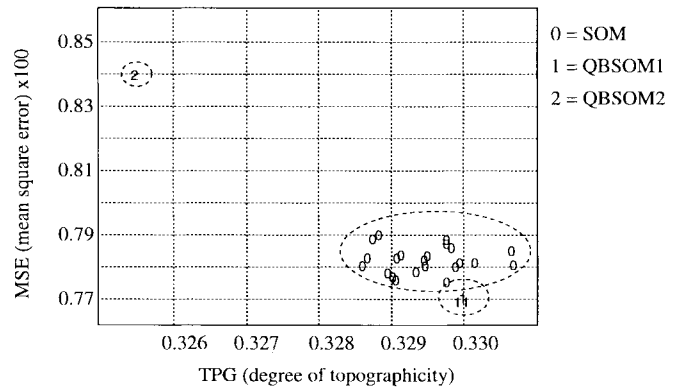

Fig. 8. The obtained results with 100 randomly generated weight patterns are presented to examine the robustness of the proposed UQBL approaches. The measurements of MSE and TPG are represented by the 2-D points (TPG, MSE). Our experiments indicate that the proposed UQBL approaches, QBSOM1 and QBSOM2, are insensitive for different initialization of weight patterns.

D. Self-Organizing with Different Iteration Numbers

In this paper, the maximum iteration number MT is pre-specified as 5000. As shown in Fig. 5, the MSE obtained by QBSOM1 is smaller than 0.25 with only 2000 iteration times. However, it will take over 3000 iteration times using SOM. In this section, a small hyper-cube network with four-dimensional $2 \times 2 \times 2 \times 2$ neurons has been tested with different definitions of MT. Fig. 10(a) shows the experimental results for defining $MT = 50\,000$. It can be found that
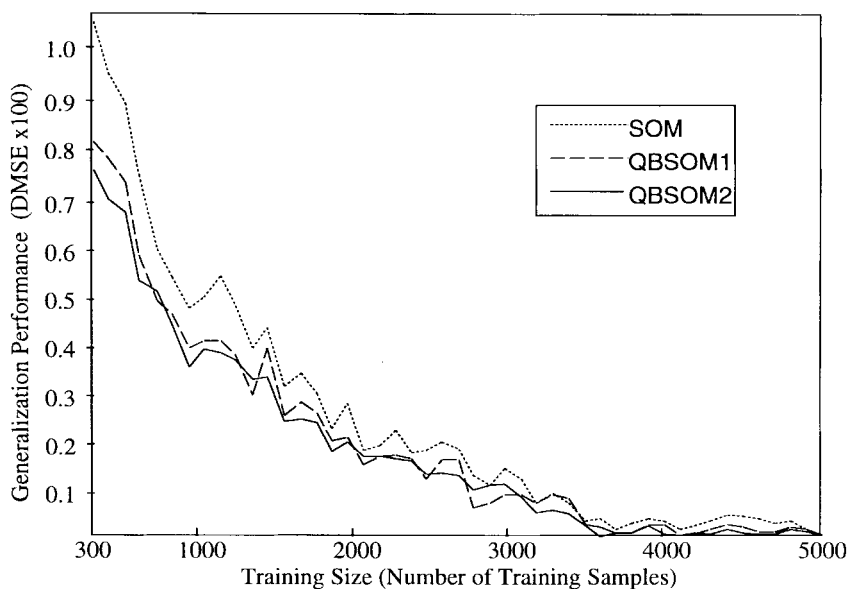
Fig. 9. The illustrations of DMSE for different training sizes (from 300 input samples to 5000 input samples) and different training methods (SOM, QBSOM1, and QBSOM2) are demonstrated to show the generalization performance of the proposed UQBL methods. The measurement DMSE is defined as the difference of MSE between the training results and the testing results.
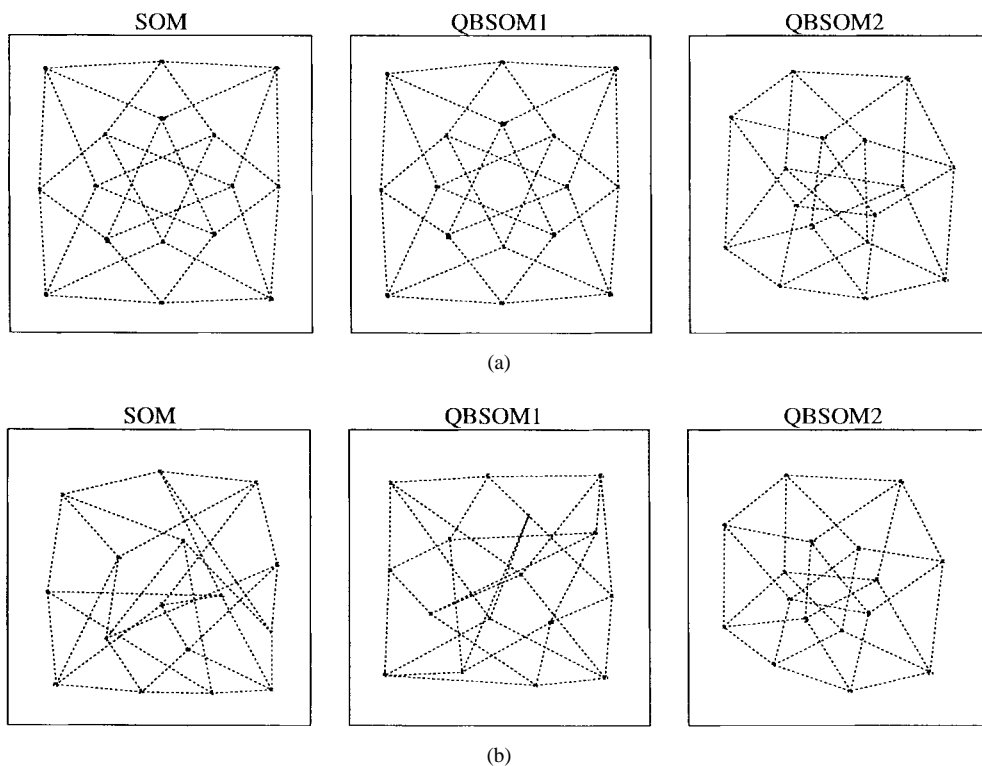

Fig. 10. The network maps obtained with (a) 5000 iteration times, and (b) with 2500 iteration times. It can be found that although QBSOM2 has better topographicity features, it has shrunk the network map too much and provided large MSE.

the obtained results for SOM and QBSOM1 are nearly the same if the applied maximum iteration number is really large. In order to demonstrate the effect for the small maximum iteration number, we have tested the same network map with $MT = 2500$. The obtained results are shown in Fig. 10(b). It can be found that QBSOM1 has obtained better results than that of SOM. Note that the obtained network map for the deepest descent QBSOM2 method is really similar to the network map shown in Fig. 10(a). However, this shrunken network provides large MSE and needs more computation times to expand out to the input sample space.

VI. CONCLUSION

In this paper, a novel UQBL algorithm based on the behavior control theory is proposed. The network model is queried to respond to its internal-desired samples. It is

difficult to learn by both the internal-desired sample and the external-input sample. In this paper, a compromise is made with an intuitive and sound conjecture from human behavior called self-regulation. Although the queried internal-desire is aimless and meaningless to minimize MSE, it has a tendency to improve network topographicity. As many researchers presented, good initial weights for network topography will be favorable to the system convergence. The proposed approach can improve the learning performance by using the internal-desire to obtain better network topography. We have introduced UQBL1 and UQBL2 as two versions of UQBL methods. Experiments have been applied to Kohonen's SOM. The obtained results show that QBSOM has better MSE and TPG, is more insensitive to network initialization, and can be a significant reduction in the training set. Our future work is to extend this learning procedure to other neural-network models.

ACKNOWLEDGMENT

The author would like to thank the anonymous referees and Prof. J. G. Taylor from King's College, London, U.K., who have provided many valuable comments and suggestions about this research. Finally, thanks also go to Y.-L. J. Chiu for her discussions about this paper.

REFERENCES

[1] L. G. Valiant, "A theory of the learnable," *Comm. ACM,* vol. 11, pp. 1134–1142, 1984.
[2] D. Angluin, "Queries and concept learning," *Machine Learning,* vol. 4, pp. 319–342, 1988.
[3] J. N. Hwang, J. J. Choi, S. Oh, and R. J. Marks, II, "QBL applied to partially trained multilayer perceptrons," *IEEE Trans. Neural Networks,* vol. 1, pp. 131–136, 1991.
[4] E. B. Baum, "Neural-net algorithms that learn in polynomial time from examples and queries," *IEEE Trans. Neural Networks,* vol. 1, pp. 5–19, 1991.
[5] R. C. Eberhart, "The role of genetic algorithms in neural network QBL and explanation facilities," in *Proc. Int. Wkshp. Combinations of GA and NN,* 1992, pp. 169–183.
[6] A. Linden and J. Kindermann, "Inversion of multilayer nets," in *Proc. Int. Joint Conf. Neural Networks,* 1989, pp. 425–430.
[7] S. Oh, R. J. Marks, II, and M. A. El-Sharkawi, "QBL in multilayered perceptrons in the presence of data jitter," in *Proc. Int. Forum Applications of NN to Power Systems,* 1991, pp. 72–75.
[8] M. A. El-Sharkawi and S. S. Huang, "Application of QBL to power system static security assessment," in *Proc. Int. Forum Applications of NN to Power Systems,* 1993, pp. 111–117.
[9] W. T. Powers, *Behavior: The Control of Perception.* Chicago: Aldine, 1973.
[10] C. S. Carver and M. F. Scheier, *Attention and Self-Regulation: A Control-Theory Approach to Human Behavior.* New York: Springer-Verlag, 1981.
[11] W. Glasser, *Stations of the Mind: New Directions for Reality Therapy.* Cambridge: Harper and Row, 1981.
[12] L. E. Bourne, R. L. Dominowski, and E. F. Loftus, *Cognitive Processes.* Englewood Cliffs, NJ: Erlbaum, 1980.
[13] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing (PDP): Exploration in the Microstructure of Cognition*, D. E. Rumelhart and G. E. Hinton, Eds. Cambridge, MA: MIT Press, 1986, ch. 8.
[14] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE Acoust., Speech, Signal Processing Mag.,* pp. 4–22, 1987.
[15] T. Kohonen, *Self-Organization and Associative Memory.* New York: Springer-Verlag, 1989.
[16] S. Y. Kung, *Digital Neural Networks.* Singapore: Prentice-Hall, 1993.
[17] J. J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems," *Biol. Cybern.,* vol. 52, pp. 141–152, 1985.
[18] V. V. Tolat, "An analysis of Kohonen's self-organizing maps using a system of energy functions," *Biol. Cybern.,* vol. 64, pp. 155–164, 1990.
[19] A. Newell and H. A. Simon, "Computer simulation and human thinking," *Sci.,* vol. 134, pp. 2011–2017, 1961.
[20] J. C. Loehlin, *Computer Models of Personality.* New York: Random House, 1968.
[21] B. Berelson and G. A. Steiner, *Human Behavior: An Inventory of Scientific Findings.* Orlando, FL: Harcourt Brace World, 1973.
[22] S. Duval and R. A. Wicklund, *A Theory of Objective Self-Awareness.* New York: Academic, 1972.
[23] E. Kryszig, *Advanced Engineering Mathematics.* New York: Wiley, 1983.
[24] T. Kohonen, "Clustering, taxonomy, and topological maps of patterns," in *Proc. Int. Conf. Pattern Recognition,* 1982, pp. 182–185.
[25] ——, "The 'neural' phonetic typewriter," *IEEE Trans. Comput.,* vol. 8, pp. 11–22, 1988.
[26] B. Angeliol, G. De La Vaubols, and J. Y. Le Texier, "Self-organizing feature maps and the travelling salesman problem," *Neural Networks,* vol. 1, pp. 289–293, 1988.
[27] H. Ritter, T. Martinetz, and K. Schulten, "Topology conserving maps for learning visuomotor-coordination," *Neural Networks,* vol. 2, pp. 159–168, 1988.
[28] R.-I Chang, K. Y. Huang, and H. T. Yen, "Self-organizing neural network for picking of seismic horizons," in *Proc. Symp. Telecommunications,* 1990, pp. 431–434.
[29] D. S. Bradburn, "Reducing transmission error effects using a self-organizing networks," in *Proc. Int. Joint Conf. Neural Networks,* 1989, pp. 531–538.
[30] C. X. Zhang and D. A. Mlynski, "Neural somatotopical mapping for VLSI placement optimization," in *Proc. Int. Joint Conf. Neural Networks,* 1991, pp. 863–868.
[31] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.,* vol. 22, pp. 400–407, 1951.
[32] T. Kohonen, "Things you haven't heard about the self-organizing map," in *Proc. IEEE Int. Conf. Neural Networks,* 1993, pp. 1147–1156.
[33] N. R. Quinn and M. A. Breuer, "A force directed component placement procedure for print circuit boards," *IEEE Trans. Circuits Syst.,* vol. CAS-3, pp. 377–388, 1979.
[34] P. G. Paulin and J. P. Knight, "Force-directed scheduling in automatic data path synthesis," in *Proc. 24th ACM/IEEE Design Automat. Conf.,* 1987, pp. 195–202.
[35] ——, "Force-directed scheduling for the behavioral synthesis for ASIC's," *IEEE Trans. Computer-Aided Design,* vol. 8, pp. 661–679, 1989.
[36] K. Shahookar and P. Mazumder, "VLSI cell placement techniques," *ACM Computing Surveys,* vol. 2, pp. 143–220, 1991.
[37] K. M. Hall, "An r-dimensional quadratic placement algorithm," *Management Sci.,* vol. 17, pp. 219–229, 1970.
[38] R.-I Chang and P. Y. Hsiao, "Force directed self-organizing maps and its application to VLSI cell placement," in *Proc. IEEE Int. Conf. Neural Networks,* 1993, pp. 103–109.
[39] ——, "VLSI cell placement for custom-chip design using force directed self-organizing maps," *IEEE Trans. Neural Networks*, to be published.
[40] ——, "VLSI circuit placement with rectilinear modules using three-layer force directed self-organizing maps," *IEEE Trans. Neural Networks*, to be published.
[41] A. Hemani and A. Postula, "Cell placement by self-organization," *Neural Networks,* vol. 3, pp. 377–383, 1990.
[42] S. S. Kim and C. M. Kyung, "Circuit placement in arbitrarily-shaped regions using the self-organization principle," *IEEE Trans. Computer-Aided Design,* vol. 11, pp. 844–854, 1992.
[43] D. I. Choi and S. H. Park, "Self-creating and organizing neural networks," *IEEE Trans. Neural Networks,* vol. 4, pp. 561–575, 1994.
[44] T. Kohonen, J. Kangas, J. Laaksonen, and K. Torkkola, "LVQ_PAK: A program package for the correct application of learning vector quantization algorithms," in *Proc. Int. J. Conf. Neural Networks,* 1992, pp. I-725–730.
[45] G. A. Carpenter and S. Grossberg, "ART2: Self-organization of stable catagory recognition codes for analog input patterns," *Appl. Opt.,* vol. 26, pp. 4919–4930, 1987.
[46] Y. H. Pao, *Adaptive Pattern Recognition and Neural Networks.* Reading, MA: Addison-Wesley, 1989.
[47] R. G. Geen and J. J. Gange, "Drive theory of social facilitation: Twelve years of theory and research," *Psych. Bull.,* vol. 84, pp. 1267–1288, 1977.
[48] R.-I Chang and P. Y. Hsiao, "UQBL algorithm and its application to Kohonen's self-organizing maps," *IEEE Int. Conf. Neural Networks,* to be published.

Ray-I Chang was born on February 7, 1967 in Taichung, Taiwan, R.O.C. In 1989, he received the B.S. degree in computer and information science from the National Chiao Tung University, Hsinchu, Taiwan. Since 1990 he has been studying toward the Ph.D. degree.

He has published over 20 scientific papers. His research interests include neural networks, fuzzy logic, genetic algorithms, image processing, VLSI/CAD, real-time systems, and multimedia applications.

Pei-Yung Hsiao was born on January 5, 1957 in Taiwan, R.O.C. He received the B.S. degree in chemical engineering from Tung Hai University, Taiwan, in 1980, and the M.S. and Ph.D. degrees in electrical engineering from the National Taiwan University, Taiwan, in 1987 and 1990, respectively.

Since August 1990, he has been an Associate Professor in the Department of Computer and Information Science at the National Chiao Tung University, Hsinchu, Taiwan. Since November 1991, he also has been a Technique Consultant in the Piping Engineering Department, CTCI Corporation, Taipei, Taiwan, R.O.C. He was a Visiting Senior Fellow in the National University of Singapore from July to September 1993. He was the past President of Taiwan-IGUG for Intergraph Corporation from 1992 to 1994. His main research interests are artificial intelligence, neural networks, and expert system applications, VLSI-CAD, and piping engineering design automation. He has published more than 70 articles in conferences and authoritative journals.

Dr. Hsiao ranked second in the 1985 Electronics Engineering Award Examination conducted by the R.O.C. Government for studying abroad, and was awarded the 1990 Acer Long Term Ph.D. Dissertation Award.