Contradiction Resolution between Self and Outer Evaluation for Supervised Multi-Layered Neural Networks

Ryotaro Kamimura
IT Education Center and Graduate School of Science and Technology
Tokai University, 1117 Kitakaname, Hiratsuka, Kanagawa, Japan
ryo@keyaki.cc.u-tokai.ac.jp

***Abstract*—In this paper, we propose a new type of information-theoretic method. We suppose that a neuron should be evaluated from different points of view to precisely discern its properties. In this paper, we restrict ourselves to two types of evaluation methods for neurons, namely, self and outer-evaluation. A neuron fires only as a result of evaluating itself, while the neuron can fire as a result of evaluation by all surrounding neurons. Self- and outer-evaluation should be equivalent to each other. When contradiction between two types of evaluation exists, the contradiction should be as small as possible. Contradiction between self- and outer-evaluations is realized in terms of the Kullback-Leibler divergence between two types of neurons. Contradiction between self- and outer-evaluation can be resolved by decreasing the contradiction ratio between the two types of evaluation in terms of KL divergence. This method is expected to extract the main features in input patterns, if those are shared by two types of evaluation. We applied the method to two data sets, namely, the logistic and dollar-yen exchange rate data. In both problems, experimental results showed that visualization performance could be improved, leading to clearer class structure for both problems. In addition, when visualization was improved, generalization performance did not necessarily degrade, showing the possibility of networks with better visualization and prediction performance.**

***Keywords*—contradiction resolution; self- and outer-evaluation; visualization; self-organizing maps**

I. Introduction

We here introduce the necessity of using multiple types of evaluation on a neuron to fully understand its main mechanism. We restrict ourselves to two types of evaluation for neurons, namely, self- and outer-evaluation for actual implementation. Then, the necessity is more concretely explained in terms of explicit class structure in the self-organizing maps and the interpretation of internal representations in neural networks.

A. Contradiction Resolution

In this paper, we suppose that a neuron can be evaluated from multiple points of view. For example, a neuron can be evaluated individually or as a member of a neuron group. The neuron may be seen differently when it is evaluated individually or as a member of a group. If evaluation as an individual neuron is contradictory to evaluation as a member of a neuron group, the neuron should take some action to solve this difficulty. This behavior cannot be fully explained without considering the effects of other neurons, meaning that a neuron cannot be understood from a single type of evaluation alone; we must evaluate a neuron from multiple points of view to fully understand the main mechanism of the neuron.

For the simplification and actual application of this necessity of multiple types of evaluation, we suppose two types of neuron evaluation [1]. One type of evaluation is realized by evaluating a neuron for itself without any consideration on other neurons. This means that a neuron fires to input patterns, taking into account a relation between the neuron and the input patterns. This evaluation is called ”self-evaluation”. On the other hand, the other one is obtained by evaluating a neuron by taking into account other neurons’ relations between the neurons and the input patterns. This means that the neuron’s firing rates are determined only by those of other neurons. This evaluation is called ”outer evaluation”. If the self-evaluation of a neuron is different from the outer-evaluation of the same neuron, the neuron has very specific characteristics which are not shared by other neurons. On the other hand, if the self-evaluation is equivalent to the outer-evaluation, the characteristics inherent to the neuron are shared by the other neurons. Our hypothesis is that the contradiction between the two types of evaluation should be reduced as much as possible. It is desirable that all neurons are in harmony with each other.

B. Explicit Class Structure

The contradiction resolution can be applied to the clarification of class structure in self-organizing maps. Self-organizing maps [2], [3], [4] are a well-known technique in neural networks. In particular, they have been used to visualize complex data on simpler form. However, it has been difficult to fully visualize knowledge obtained by the self-organizing maps. Thus, there have been many different kinds of visualization techniques so far developed for self-organizing maps [5], [6], [7], [8], [9], [10], [11], [12]. [13], [14], [15], [16], [17], [18], [19], [20], [21], [22]. However, we can say that those methods did not necessarily succeed in visualizing the knowledge obtained by conventional self-organizing maps.

We think that one of the main reasons for the difficulty in visualization is due to the characteristics of self-organizing maps. Self-organizing maps are based upon competition and cooperation between neurons. In particular, cooperation plays the most important role. This cooperation is based upon the hypothesis that neighboring neurons behave in the same way as a particular neuron. A winner neuron is selected by the process

of competition, and the neighboring neurons of the winner are updated in the same way, in proportion to distance from the winner neuron. Now, to clarify class structure, it is necessary to make class boundaries as clear as possible. However, self-organizing maps go through a process of making class boundaries unclear. This is because neighboring neurons must be as similar as possible, due to the process of cooperation. The class boundaries are naturally based upon dissimilarity between neurons. This shows the necessity to attenuate cooperation between neurons for explicit class structure.

In our contradiction resolution, cooperation between neurons corresponds to the outer-evaluation, because the results obtained by the outer-evaluation are ones realized only by all the other neurons. On the other hand, self-evaluation is a form of evaluation realized only within a neuron without considering any other neurons. The influence of self-evaluation can be used to attenuate the effect of cooperation between neurons and to clarify dissimilarity between neurons. This dissimilarity is related to the clarification of class boundaries in self-organizing maps.

C. Interpretation of Internal Representation

Our contradiction resolution is used to provide supervised neural networks with more interpretable internal representations. One of the most important things to do in neural networks is to explain why such neural mechanisms can be used to produce outputs from inputs through obtained internal representations [23]. Though there have been many attempts to interpret final representations [24], [25], [26], [27], [28] [29], [30], [31], the problem has remained unsolved. Furthermore, we are as yet unaware of the relationship between interpretable representations and generalization performance.

Our method aims to provide neural networks with interpretable internal representations in two ways, namely, through visualization performance due to self-organizing maps and the characteristics of self- and outer-evaluation. First, because our method is applied to the production of self-organizing maps, the outputs from our method are inherited from interpretable knowledge in the self-organizing maps. As above mentioned, the introduction of self- and outer-evaluation can be used to make class boundaries clearer and to make it easier to interpret final knowledge.

Second, the self- and outer-evaluation aim to enhance the characteristics shared by two types of evaluation. A neuron is evaluated for itself and by all the other ones. When the characteristics obtained by the self-evaluation are equivalent to those obtained by the outer-evaluation, those characteristics should be emphasized or enhanced. On the other hand, if the characteristics obtained by self-evaluation are different from those obtained by outer-evaluation, those characteristics should be weakened. Thus, we can expect that gradually, only important and shared characteristics will become stronger. Our contradiction resolution should finally produce interpretable knowledge by enhancing important characteristics.

D. Outline

In Section 2, we first explain the concept of contradiction between self and outer-evaluation. Then, we present how to compute the firing rates by the self and outer-evaluation. We introduce the Kullback-Leibler divergence between firing rates by self- and outer-evaluation. This Kullback-Leibler divergence is called the ”contradiction ratio” to show how self- and outer-evaluation are different from each other. Finally, we explain how to compute connection weights by reducing the contradiction ratio.

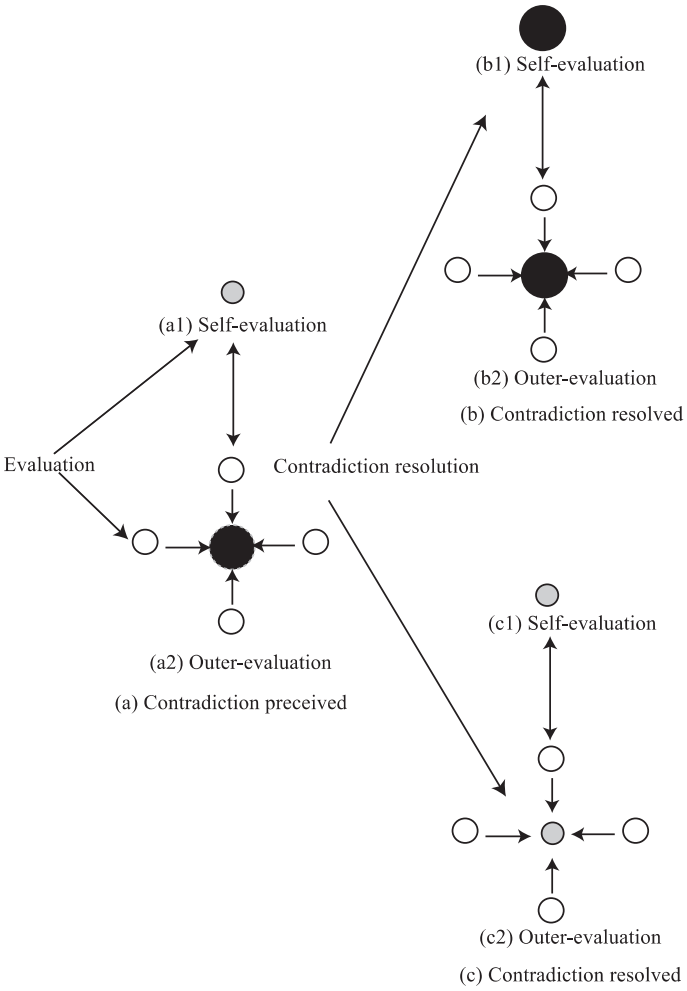

Fig. 1. Concept of contradiction resolution: perception (a) and resolution (b) and (c).

In Section 3, we present two experimental results, namely, the logistic data and dollar-yen exchange rate estimation. In the logistic data, we explain several evaluation measures and how to implement them, showing experimental results. While the results show that the conventional SOM failed to produce explicit class boundaries, the contradiction resolution produced from two to five class boundaries, depending on the parameter. In addition, generalization performance was the same as that by the conventional self-organizing maps and radial-basis function networks. In the dollar-yen exchange rate prediction, class structure became gradually complicated when the parameter was increased. The generalization error gradually decreased and seemed to reach a stable state when the parameter was increased. The generalization errors were sufficiently small compared to those generated by the conventional self-organizing maps and the radial-basis function networks.

II. Theory and Computational Methods

A. Concept of Contradiction Resolution

We suppose that a neuron can be evaluated from many different points of view to fully understand its main mechanism. The characteristics of a neuron can be changed by different types of evaluation. A neuron is defined by the unification of all the characteristics evaluated from different points of view. For simplification, we suppose that there are only two types of evaluation, namely, self- and outer-evaluation. In self-evaluation, a neuron is evaluated by itself and for itself, as represented in Figure 1 (a1). On the other hand, in outer-evaluation, a neuron is evaluated and viewed by all the other neurons as shown in Figure 1 (a2). It is desirable that the results of self-evaluation are equivalent to those of outer-evaluation. However, it happens that the results of self-evaluation are contrary to those by outer-evaluation. As shown in Figure 1 (a), the firing rate obtained via self-evaluation is small while that obtained via outer-evaluation is large. Because this contradiction between self- and outer-evaluation exists, we should make it as small as possible. In Figure 1 (b1), the firing rate obtained by self-evaluation becomes larger and closer to that obtained by outer-evaluation. This is a case where outer-evaluation is more influential. On the other hand, in Figure 1 (c), the effect of self-evaluation becomes apparent. The firing rate obtained by the outer-evaluation in Figure 1 (c2) becomes smaller by the contradiction resolution, because the effect of self-evaluation has some influence on the outer-evaluation.

By using this property of contradiction resolution, the characteristics shared by both self- and outer-evaluation can be enhanced, while those specific to each evaluation can be inhibited. Thus, we can expect that important characteristics are enhanced by the contradiction resolution.

B. Self and Outer Evaluation

Let us explain how to compute outputs from competitive neurons and input patterns in Figure 2. The sth input pattern of total S patterns can be represented by $\mathbf{x}^s = [x_1^s, x_2^s, \cdots, x_L^s]^T, \quad s = 1, 2, \cdots, S$. Connection weights into the jth competitive unit of total M units are computed by $\mathbf{c}_j = [c_{1j}, c_{2j}, \cdots, c_{Lj}]^T, \quad j = 1, 2, \ldots, M$. Now, we can compute the firing rates by self and outer-evaluation. A neuron's firing rates by self-evaluation can be defined by using the outputs from the neuron. Then, the jth competitive neuron output, without considering the other neurons, can be computed by

$$v_j^s = \exp\left(-\frac{\| \mathbf{x}^s - \mathbf{c}_j \|^2}{2\sigma_\beta^2}\right). \tag{1}$$

where $\mathbf{x}^s$ and $\mathbf{w}_j$ are supposed to represent L-dimensional input and weight column vectors, where L denotes the number of input units. The spread parameter σ_β is computed by $1/\beta$, where $\beta > 0$. Thus, the firing rate obtained by the self-evaluation can be computed by

$$p(j \mid s) = \frac{\exp\left(-\frac{\|\mathbf{x}^s - \mathbf{c}_j\|^2}{2\sigma_\beta^2}\right)}{\sum_{m=1}^{M} \exp\left(-\frac{\|\mathbf{x}^s - \mathbf{c}_m\|^2}{2\sigma_\beta^2}\right)}. \tag{2}$$

The firing rate of a neuron evaluated by outer-evaluation is determined by considering all the firing rates of all the other neurons. We approximate the firing rates by summing all firing rates of all neighboring neurons, excluding the target neuron. In addition, the rates are weighted by the distances between those neurons and the target neuron. Then, the outputs by the outer evaluation are defined by

$$y_j^s = \sum_{m=1}^{M} (1 - \delta_{jm}) \phi_{jm} p(m|s), \tag{3}$$

where ϕ_{jm} represent relations between the jth and mth neuron. The output obtained by the evaluation is the sum of all neighboring neurons' firing rates weighted by the relations between them. Then, the firing rate of the outer-evaluation can be defined by

$$q(j \mid s) = \frac{\sum_{m=1}^{M} (1 - \delta_{jm}) \phi_{jm} p(m|s)}{\sum_{r=1}^{M} \sum_{m=1}^{M} (1 - \delta_{rm}) \phi_{rm} p(m|s)}. \tag{4}$$

C. Contradiction Resolution

Contradiction resolution aims to reduce contradiction between self and outer-evaluation [32]. We use the Kullback-Leibler divergence to represent the difference between self- and outer-evaluation. Using the Kullback-Leiber divergence, the contradiction ratio is defined by

$$CR = \sum_{s=1}^{S} p(s) \sum_{j=1}^{M} p(j \mid s) \log \frac{p(j \mid s)}{q(j \mid s)}. \tag{5}$$

When the KL divergence is minimized, we have

$$p^*(j \mid s) = \frac{q(j \mid s) \exp\left(-\frac{\|\mathbf{x}^s - \mathbf{c}_j\|^2}{2\sigma_\beta^2}\right)}{\sum_{m=1}^{M} q(m \mid s) \exp\left(-\frac{\|\mathbf{x}^s - \mathbf{c}_m\|^2}{2\sigma_\beta^2}\right)}. \tag{6}$$

By substituting this optimal firing rates $p^*(j \mid s)$ for $p(j \mid s)$, we have the free energy:

$$\begin{aligned} F &= -2\sigma^2 \sum_{s=1}^{S} p(s) \log \sum_{j=1}^{M} q(j \mid s) \\ &\quad \times \exp\left(-\frac{\| x^s - c_j \|^2}{2\sigma_\beta^2}\right). \end{aligned} \tag{7}$$

This equation can be expanded as

$$\begin{aligned} F &= \sum_{s=1}^{S} p(s) \sum_{j=1}^{M} p^*(j \mid s) \|\mathbf{x}^s - \mathbf{w}_j\|^2 \\ &\quad + 2\sigma_\beta^2 \sum_{s=1}^{S} p(s) \sum_{j=1}^{M} p^*(j \mid s) \log \frac{p(j \mid s)}{q(j \mid s)}. \end{aligned} \tag{8}$$

Thus, the free energy can be used to decrease KL divergence as well as quantization errors. By differentiating the free energy, we can obtain re-estimation formula:

$$\mathbf{w}_j = \frac{\sum_{s=1}^{S} p^*(j \mid s) \mathbf{x}^s}{\sum_{s=1}^{S} p^*(j \mid s)}. \tag{9}$$

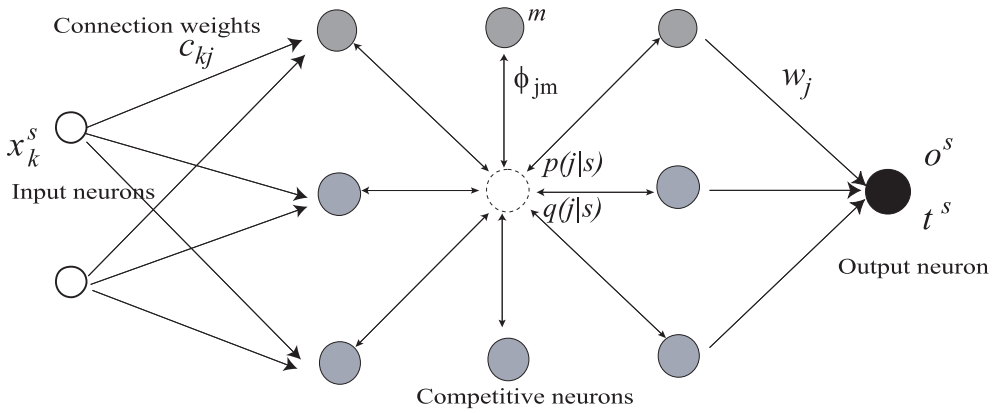

Fig. 2. A neural network for two types of evaluation where some connection weights are eliminated for simplification purposes.

III. Results and Discussion

We here present two experimental results by using the artificial data and the dollar-yen exchange rates. We used the logistic data as the artificial one, where in addition to the presentation of experimental results, we show how to compute several computational measures to evaluate the performance quantitatively and visually. In the dollar-yen exchange rate estimation, we in particular show how visualization performance is related to prediction performance. The network size was set to a much larger one for the intuitive interpretation of our results, namely, 20 by 15 neurons.

A. Logistic Function Identification

1) Experiment Outline: We first used artificial data generated by the logistic function $y = (1 - \exp(-x))/(1 + \exp(-x))$. As shown in Figure 3(a), the training data was generated by the function added to the normal random values with a standard deviation of 0.1. The number of training and testing patterns was 100 and 1,000, respectively. Figure 3(b) shows testing data (in black) and predicted values (in red) by the RBF network with forward selection with the Bayesian information criterion [33], [34], [35]. The input data increased linearly from -10 to 10, meaning that no boundaries inside could be identified. Thus, we tried to show how the contradiction resolution divides this linear data into classes. Then, we examine to what extent this classification is related to the prediction performance.

2) Quantitative Evaluation: We quantitatively evaluated the performance of contradiction resolution in terms of the property of topological maps and prediction performance. For the property of the self-organizing maps, we used the well-known and simple error measures for quantification evaluation, namely, quantization and topographic errors. This is because much importance was placed on the easy reproduction of our results. The quantization error is simply the average distance from each data vector to its BMU (best-matching unit). The topographic error is the percentage of data vectors for which the BMU and the second-BMU are not neighboring units [36]. In addition, to measure how the maps are organized, we computed mutual information (INF)

$$INF = \sum_{s=1}^{S} \sum_{j=1}^{M} p(s)p(j \mid s) \log \frac{p(j \mid s)}{p(j)}. \qquad (10)$$

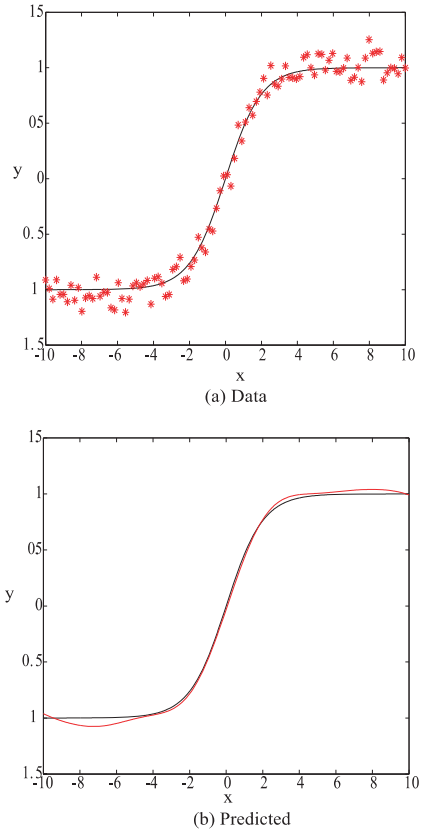

Fig. 3. Training data in red (a) and testing data in black with predicted values in red (b) for the logistic function $y = (1 - \exp(-x))/(1 + \exp(-x))$.

As this mutual information increases, the organization of the maps increases accordingly.

Table I shows the summary of the experimental results. The mean squared errors (MSE) between the targets and outputs on the output layer decreased from 0.11346 (β=1) to 0.00126 (β=5). Then, the mean squared errors (MSE) slightly increased and reached a stable value. The quantization errors (QE) decreased from 5.050 (β=1) to 0.073 (β=20). The topographic errors (TE) decreased from one (β=1) to 0.2 (β=30). Mutual information increased from 0 (β=1) to 1.654 (β=30). The correlation coefficients between mutual information and MSE, QE and TE were -0.962, -0.999 and -0.419, respectively.

TABLE I. MSE FOR THE TESTING DATA, QUANTIZATION ERRORS (QE), TOPOGRAPHIC ERRORS (TE) AND MUTUAL INFORMATION WHEN THE PARAMETER β WAS CHANGED FROM 1 TO 19 FOR THE 20 BY 15 MAP. THE SYMBOL CC IN THE FINAL ROW REPRESENTS THE CORRELATION COEFFICIENTS BETWEEN INFORMATION (INF) AND THE OTHER MEASURES.

β	MSE	QE	TE	INF
1	0.11346	5.050	1.000	0.000
3	0.00159	1.580	1.000	1.146
5	0.00126	0.633	1.000	1.428
7	0.00127	0.317	1.000	1.533
10	0.00128	0.151	1.000	1.605
15	0.00126	0.094	0.720	1.638
20	0.00126	0.073	0.660	1.649
25	0.00126	0.075	0.420	1.653
30	0.00126	0.080	0.200	1.654
CC	-0.962	-0.999	-0.419	

Thus, the quantization errors and topographic errors could be decreased when mutual information was increased.

3) Visual Evaluation: We evaluated visualization performance in terms of a U-matrix and contradiction ratios. For visualization, we computed the contradiction ratio for each neuron by the average of the contradiction ratios over all input patterns

$$R_j = \sum_{s=1}^{S} p(s) \left| \log \frac{p(j \mid s)}{q(j \mid s)} \right|. \qquad (11)$$

We used the absolute values of the contradiction ratios only for better visualization, namely, to show a tendency similar to that of the U-matrix.

Figure 4 shows U-matrices (a) and contradiction ratios (b). When the parameter β was increased from 10 to 20 in Figure 4 (a1)-(a3), two class boundaries were gradually unfolded. When the parameter β was 25 in Figure 4 (a4), the number of classes increased to three. Finally, when the parameter was increased to 30 in Figure 4 (a5), the number of class boundaries increased to four. In addition, the contradiction ratios gradually detected five classes corresponding to those detected by the U-matrix in Figure 4 (b1)-(b5).

4) Interpretation: Figure 5 shows labels (1) and data with boundaries (2) when the parameter β was changed from 20 (a) to 30 (c). When the parameter β was 20, two class boundaries were detected and three classes were identified. As shown in Figure 5 (a), the logistic function was separated into flat higher and lower areas, and the intermediate areas between them. When the parameter β was increased to 25 in Figure 5 (b), we could see three classes boundaries dividing four classes, see Figure 5 (b1). As shown in Figure 5 (b2), in addition to two class boundaries in Figure 5 (b1), a class boundary in the middle of the U-matrix could be identified. Finally, when the parameter β was 30 in Figure 5(c), four class boundaries with five classes were identified, see Figure 5(c1). The slope between the lower and higher areas were further subdivided into three classes in Figure 5 (c2).

5) Using Conventional Methods: For comparison, we used the conventional SOM for the same logistic function. For producing the self-organizing maps, the well-known SOM toolbox of Vesanto et al. [37] was used, because the final results of SOMs have been very different given the small changes in implementation such as initial conditions. We have

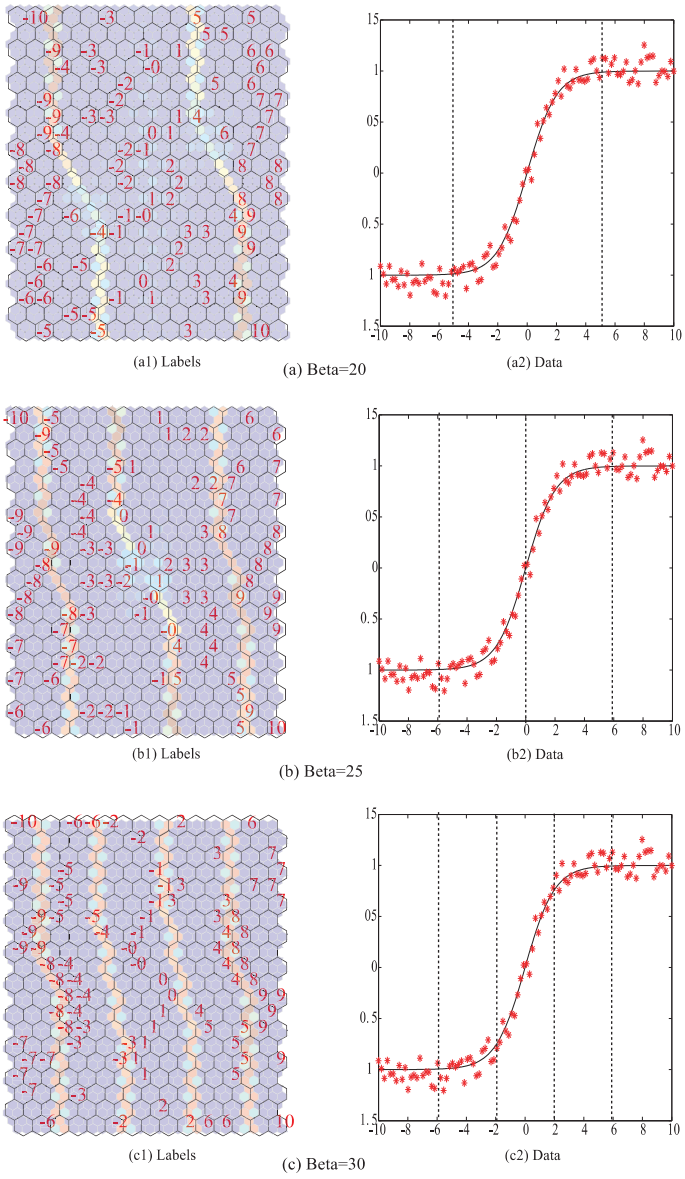

Fig. 5. U-matrices with 20 by 15 when the parameter β was changed from 20 (a) to 30 (c).

confirmed the reproduction of stable final results by using this package. In the SOMs, the Batch method was used, which has shown better performance than the popular real-time method in terms of visualization, quantization and topographic errors.

The mean squared error (MSE) was the same (0.00126) as that of the contradiction ratio in Table II. The quantization error was 0.048, see Table II. On the other hand, by the contradiction resolution, the lowest error was 0.073, as shown in Table I. Thus, the conventional SOM showed lower quantization errors. The topographic error was 0.580, shown in Table II. Thus, the topographic error by the contradiction ratio in Table I was much lower than that obtained by the conventional SOM. Mutual information was 1.635 by the SOM, while it increased to 1.654 via the contradiction resolution, see Table I. Thus, mutual information by contradiction resolution was slightly higher than that by the conventional SOM. The RBF networks with the Ridge regression and generalized cross validation

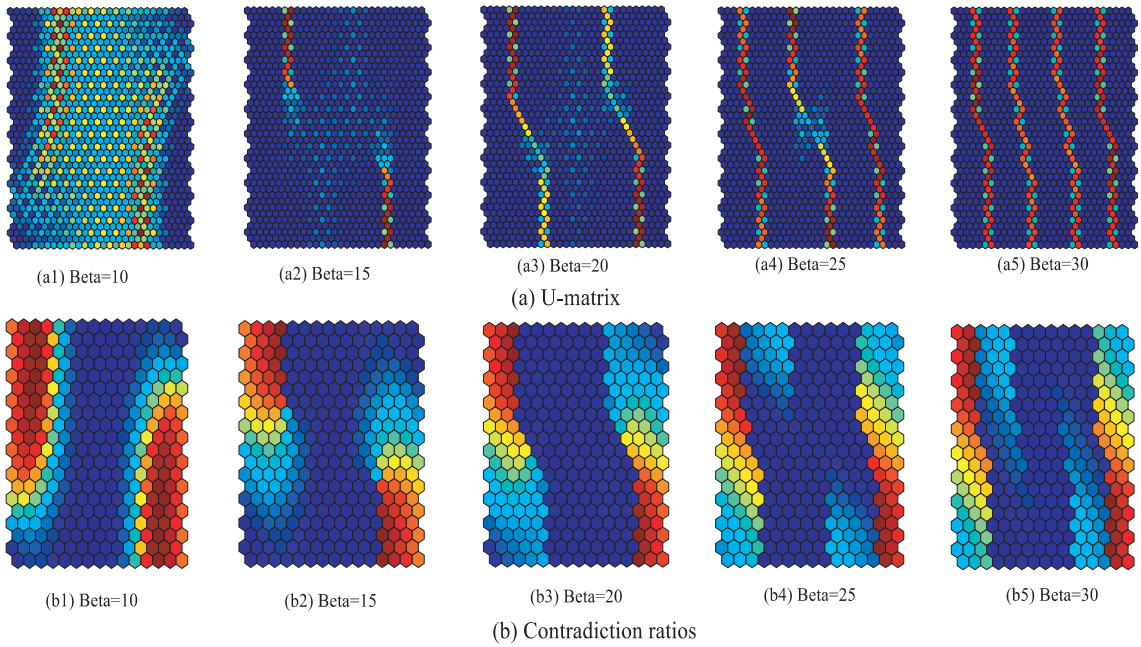

Fig. 4. U-matrices and contradiction ratios with 20 by 15 when the parameter β was changed from 10 (a) to 30 (c).

TABLE II. MSE FOR TESTING DATA, QUANTIZATION ERRORS (QE), TOPOGRAPHIC ERRORS (TE) AND MUTUAL INFORMATION (INF) BY THE SOM AND RADIAL-BASIS NETWORKS. THE SYMBOL "FS" DENOTES THE FORWARD SELECTION RBF WITH THE BAYESIAN INFORMATION CRITERION AND "RR" DENOTES THE RIDGE REGRESSION RBF WITH THE GENERALIZED CROSS VALIDATION.

Size or methods	MSE	QE	TE	INF
SOM	0.00126	0.048	0.580	1.635
FS(bic)	0.00142			
RR(gcv)	0.00124			

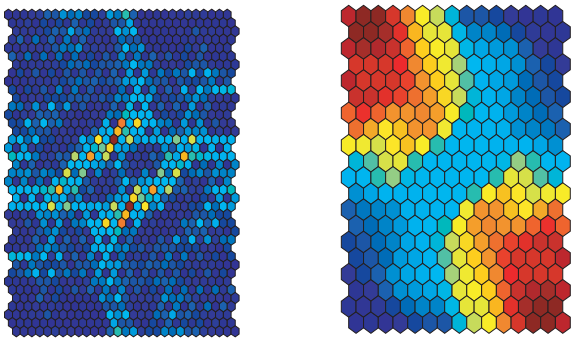

Fig. 6. U-matrices (a) and contradiction ratios (b) by the conventional SOM when the network size was 20 by 15.

[33], [34], [35] showed the best result of 0.00124 (MSE). However, by the RBF with forward selection and the Bayesian information criterion, the worst error of 0.00142 was obtained. By visual inspection, we could see that two class boundaries, which were rather weak, moved to the center of the U-matrix in Figure 6 (a). By the contradiction ratios, two classes on the upper right and lower left hand sides of the map became larger, as in Figure 6(b).

TABLE III. MSE FOR TESTING DATA, QUANTIZATION ERRORS (QE), TOPOGRAPHIC ERRORS (TE) AND MUTUAL INFORMATION WHEN THE TIME LAG WAS CHANGED FROM 1 TO 10 AND THE PARAMETER β WAS 10.

Lag	MSE	QE	TE	INF
1	0.068	0.058	0.776	0.794
2	0.077	0.286	0.385	0.802
3	0.066	0.370	0.304	0.800
4	**0.050**	0.450	0.194	0.793
5	0.058	0.447	0.133	0.794
6	0.060	0.495	0.257	0.794
7	0.061	0.536	0.174	0.791
8	0.064	0.559	0.181	0.789
9	0.063	0.571	0.130	0.790
10	0.065	0.608	0.097	0.786

B. Dollar-Yen Exchange Rate Prediction

1) Time Lag: In the dollar-yen exchange rate prediction, the rate at time t must be estimated by the previous q rates $x_{t-1,\ldots,t-q}$. The time lag q must be determined before the experiment. We determined the time lag based on the errors (MSE) for the testing data when the parameter β was ten. Table III shows the summary of the experimental results. The mean squared error (MSE) for the testing data decreased and increased from 0.068 (lag=1) to 0.065 (lag=10). When the time lag was four, the minimum value of 0.050 was obtained. The quantization errors (QE) increased from 0.058 to 0.608, while the topographic errors (TE) decreased from 0.776 to 0.097. Mutual information decreased slightly from 0.802 (lag=2) to 0.786 (lag=10). The experimental results showed that the mean squared error (MSE) was the lowest when the time lag was four. The quantization errors (QE) increased gradually when the lag increased, while the topographic errors decreased. In addition, mutual information remained almost unchanged. From these results, in our experiments, the time lag was determined to be four.

TABLE IV. MSE FOR THE TESTING DATA, QUANTIZATION ERRORS (QE), TOPOGRAPHIC ERRORS (TE) AND MUTUAL INFORMATION WHEN THE PARAMETER β WAS CHANGED FROM 1 TO 19 FOR THE 20 BY 15 MAP. . THE SYMBOL CC IN THE FINAL ROW REPRESENTS CORRELATION COEFFICIENTS BETWEEN INFORMATION (INF) AND THE OTHER MEASURES.

β	MSE	QE	TE	INF
1	5.024	3.977	0.000	0.000
3	0.108	0.861	0.161	0.715
5	0.155	0.611	0.300	0.787
7	0.109	0.573	0.294	0.794
9	**0.049**	0.508	0.317	0.796
11	0.049	0.401	0.161	0.796
13	0.055	0.326	0.194	0.797
15	0.054	0.298	0.244	0.797
17	0.055	0.281	0.228	0.797
19	0.054	0.263	0.206	0.795
CC	-0.973	-0.989	0.709	

2) Prediction Performance: Table IV shows the summary of the experimental results. The minimum value for MSE was 0.049 when the parameter β was nine and eleven. The quantization errors (QE) decreased gradually to 0.263 (β=19). The topographic errors (TE) fluctuated from zero (β=1) to 0.317(β=9). Mutual information immediately reached the level of 0.794 when the parameter β was seven. Then, mutual information seemed to be stable. The correlation coefficients between MSE, QE, TE and mutual information were -0.973, -0.989 and 0.709, respectively.

The experimental results showed that the mean squared errors for the testing data was the lowest when the parameter β was around 10. The quantization errors decreased gradually when the parameter β was increased. The topographic errors fluctuated when the parameter β was increased. In addition, mutual information had a strong negative correlation with the MSE and quantization errors. This means that if we want to decrease the MSE and quantization errors, we have only to increase mutual information, meaning that all we have to do given decreasing MSE and quantization errors is to increase the parameter β.

3) Visual Performance: Figure 7 shows the U-matrices and contradiction ratios when the parameter β was increased from 3 to 19. When the parameter β was increased from three in Figure 7 (a1) and (b1) to seven in Figure 7 (a3) and (b3), one clear class boundary could be seen and neurons with higher contradiction ratios could be seen on the upper side of the map. When the parameter β was increased to nine in Figure 7 (a4) and (b4), the class boundary on the upper side was twisted, and another boundary appeared at the bottom. Neurons with the contradiction ratios were also differentiated on the upper side, and a small group of higher contradiction ratios could be seen on the lower side of the map. When the parameter β was increased to 19 in Figure 7 (a5) and (b5), the class boundaries on the upper side of the map became further complicated, and other boundaries at the bottom appeared. The contradiction ratios with high values scattered to the four corners of the map, as in Figure 7 (b5).

4) Interpretation: We tried to interpret how contradiction resolution classified the entire period. We interpreted the maps when the parameter β was fixed so as to minimize the MSE for the testing data. Figure 8 shows maps with labels based on the U-matrices when the network size was 20 by 15 (c)

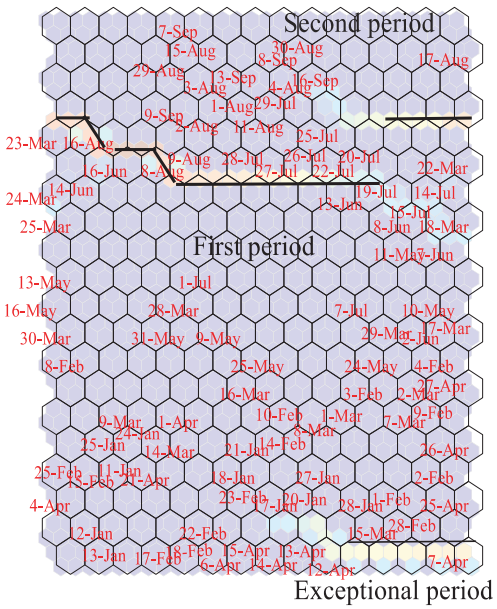

Fig. 8. Map with labels and class boundaries based on the U-matrices and with 20 by 15 neurons for the dollar-yen exchange rates.

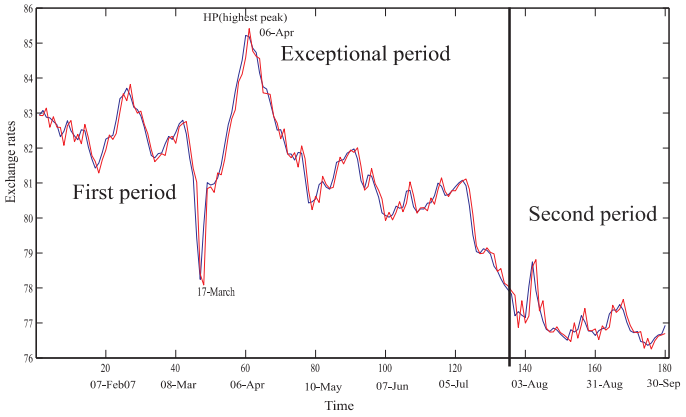

Fig. 9. Dollar-yen exchange rates during 2011.

and the parameter β was nine. The entire period seemed to be divided into three. The first period saw the rates gradually decreasing from January to July in Figure 9. The second period saw the dollar-yen rates remain relatively stable during August and September, as in Figure 9. In addition, a period with the highest rates of April was separated by the class boundary on the lower side.

5) Conventional Methods: For comparison, we used the conventional SOM and RBF networks. Table V shows the summary of experimental results when we used the conventional SOM and the radial basis networks. By the SOM, the MSE was 0.052, which was higher than that obtained by the contradiction ratio in Table IV.

Using the RBE network with forward selection and the Bayesian information criterion, the obtained error was 0.055. When the RBF network with Ridge regression was used, the MSE increased to 0.063, which was larger than that obtained by our method in Table IV. Thus, our method showed the possibility of better prediction performance. For the other measures, the quantization and topographic errors were smaller

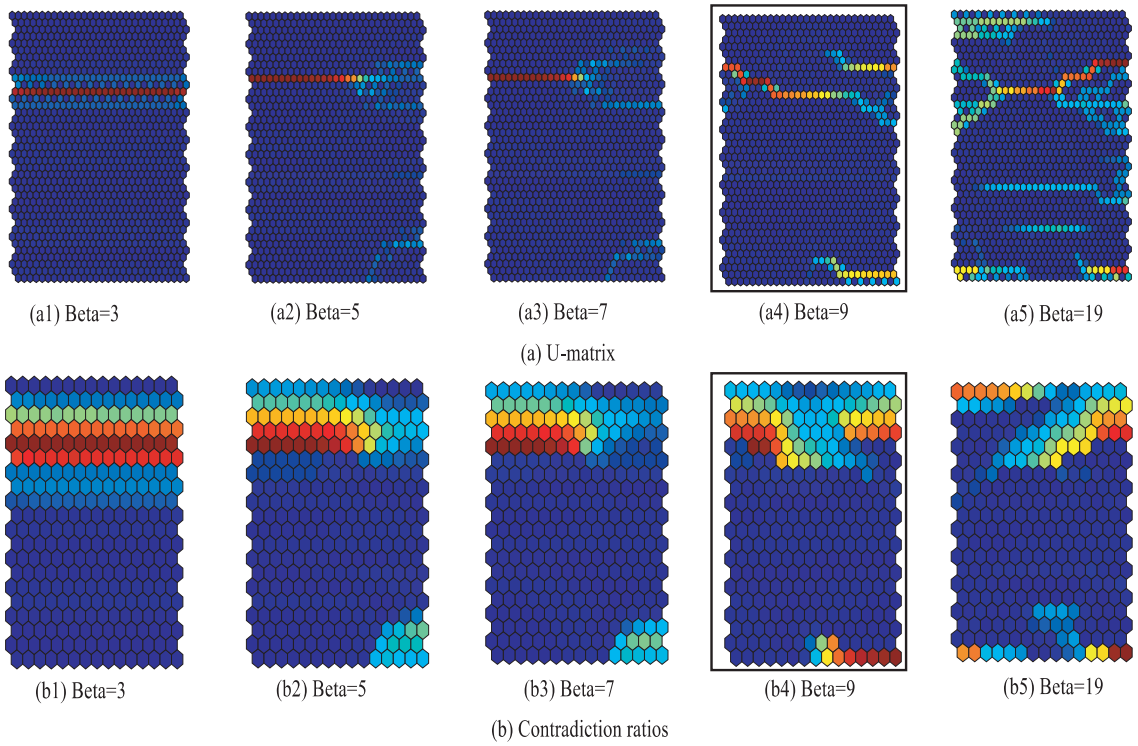

Fig. 7. U-matrices with 20 by 15 when the parameter β was increased from 3 to 19.

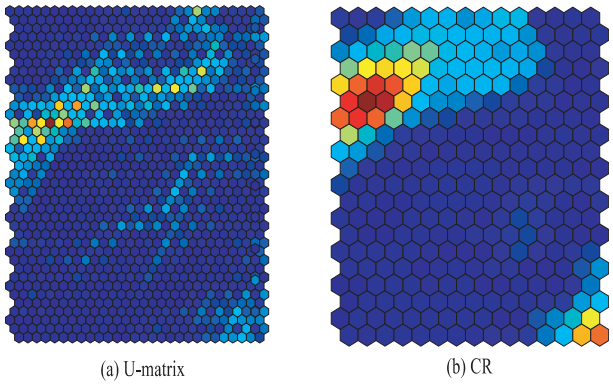

Fig. 10. Contradiction ratios and U-matrices by the conventional methods.

TABLE V. MSE FOR TESTING DATA, QUANTIZATION ERRORS (QE), TOPOGRAPHIC ERRORS (TE) AND MUTUAL INFORMATION BY THE CONVENTIONAL SOM AND RBF NETWORKS. "FS" DENOTES THE FORWARD SELECTION RBF WITH THE BAYESIAN INFORMATION CRITERION AND "RR" DENOTES THE RIDGE REGRESSION RBF WITH THE GENERALIZED CROSS VALIDATION.

Methods	MSE	QE	TE	INF
SOM	0.052	0.235	0.117	0.773
FS(bic)	0.055			
RR(gcv)	0.063			

than those obtained by our method in Table IV. In addition, mutual information was also smaller than that obtained by the contradiction ratio in Table IV.

Figures 10 (a) and (b) show the contradiction ratios and U-matrices obtained by the conventional SOM. Class boundaries on the upper side of the matrix seemed to be present, but were weaker than those obtained by our method, as shown in Figure 7. In addition, a group of neurons with higher contradiction ratios was smaller and weaker compared to those present in our method, see Figure 7.

C. Discussion

1) Validity of Methods and Experimental Results: In this paper, we proposed a new information-theoretic method to resolve contradiction in neural networks. We supposed that a neuron can be evaluated from multiple points of view. To fully understand the characteristics of the neuron, we should examine the characteristics obtained from these different types of evaluation. For simplification, we suppose that a neuron can be viewed from two forms of evaluation. These two types of evaluation are self- and outer-evaluation, respectively. For the self-evaluation, a neuron's firing rate is determined by evaluating only relations between the neuron and the incoming input patterns. For the outer-evaluation, a neuron's firing rate is determined by evaluating relations between the incoming input patterns and all its surrounding neurons. If contradiction between self- and outer-evaluation exists, this contradiction should be reduced as much as possible. In this way, we expect that the characteristics shared by self- and outer-evaluation will be enhanced, and eventually, that important characteristics can be intensified. Thus, the contradiction resolution can be expected to improve visualization and prediction performance by extracting important characteristics in input patterns.

First, the experimental results confirmed that interpretation by visualization could be improved. In the logistic data, weak class boundaries were produced in terms of the U-matrix by the conventional SOM, as in Figure 6. By contradiction resolution,

depending on the parameter β, we were able to explicitly identify three to five classes in the logistic data in Figure 4. For the dollar-yen exchange rates, the number of class boundaries differed depending on the parameter β, see Figure 7. The conventional SOM failed to produce explicit class structure in terms of the U-matrix in Figure 10. Compared with the class structure obtained by the conventional SOM, the class structure obtained by contradiction resolution was clearer, as in Figure 7.

As mentioned in the introduction section, many attempts have been made to visualize final representations in self-organizing maps [5], [6], [7], [8], [9], [10], [13], [14], [16]. However, it has been difficult to visualize SOMs' knowledge at the present of stage of techniques. Our method of contradiction resolution succeeded in producing explicit class structure in two experimental results. These experimental results showed that distinction between self- and outer-evaluation had effects to clarify class structure.

However, the quality of visualized maps was not necessarily improved. For the logistic data, quantization errors were not necessarily smaller than those obtained by the conventional SOM in Table I and II. For the dollar-yen exchange rates, quantization and topographic errors were not smaller than those obtained by the conventional SOM in Tables IV and V. Only for the topographic errors of the logistic data, it was possible that the errors were smaller than those by obtained the conventional SOM, see Tables I and II. Though improved predication performance was obtained, there is the possibility that quantization and topographic errors were scarified. Thus, careful attention should be paid to final representations for interpretation.

Favorable results were obtained for prediction performance. For the logistic data, contradiction resolution showed that the MSE for the testing data was almost equivalent to that obtained by the conventional SOM and RBF networks, as in Tables I and II. In the dollar-yen exchange rates, the MSE by the contradiction reduction was the lowest in Tables IV and V. In both cases, clearer class structures could be identified as in Figures 7 and 10. This shows that contradiction resolution could be used to improve visualization while keeping generalization errors low.

Because the evaluation criteria between our method and the conventional RBF were different from each other, we could not say definitely that contradiction resolution showed better prediction performance. However, the experimental results showed the possibility that prediction performance by contradiction resolution is not necessarily contrary to visualization performance.

2) Limitation of the Method: One of the main problems is that we have not yet determined criteria to choose optimal representations for visualization and prediction. For visualization, the experimental results showed that when the parameter β was increased, class structure became complicated. Because we do not yet have criteria to measure the clarity of class structure, all we have to do is to visually inspect the final representations and to choose the best possible representations for visualization. Thus, it is apparent that we need objective criteria to choose optimal representations for better visualization.

Objective criteria have not existed to choose the optimal representations for prediction performance either. One of the ways to solve this problem is to use mutual information. In the two experimental results, mutual information was correlated with quantization errors and MSE. The correlation coefficients were close to one for the two data sets in Tables I and IV. This means that quantization and MSE can be decreased by increasing mutual information. Thus, one way to obtain optimal representations is to monitor mutual information. Then, the optimal point should be where mutual information ceases to increase, namely, where mutual information reaches a stable point.

3) Possibility of the Method: The possibility of the method was explained in terms of a new SOM with a new visualization method for self-organizing maps and an extension to multiple types of evaluation. First, our method is related to a new type of self-organizing maps which produce more explicit class structure with a new visualization method. As mentioned in the introduction section, it has been very difficult to visualize SOM knowledge because final representations are usually not easily interpretable. There have been many attempts to visualize SOM knowledge more clearly. However, even if sophisticated visualization methods are used, it is still difficult to interpret SOM knowledge. In our experiments, by using the conventional SOM, weak boundaries were detected both for the logistic data in Figure 6 and for the dollar-yen exchange rates in Figure 10. However, even if class boundaries were weak in terms of a U-matrix, clear contradiction ratios could be obtained in the above figures. In addition, the contradiction ratios for each neuron showed clear class structure for both cases in Figures 5 and 8. These results show the possibility of a new SOM with a visualization method by using the contradiction ratios.

Second, we have restricted ourselves to two types of evaluation, namely, self- and outer-evaluation. This is because they are easily implemented when we suppose the two types of evaluation. However, as mentioned in the introduction section, we think that a neuron should be evaluated from as many different viewpoints as possible. Contradiction should be computed among many different types of evaluation, and this contradiction should be reduced as much as possible. If it is possible to take into account many types of evaluation for neurons, the characteristics of the neurons can be more exactly interpreted.

IV. Conclusion

In this paper, we proposed a new type of information-theoretic method called "contradiction resolution." We supposed that a neuron should be evaluated differently and those different types of evaluation should be unified. For simplification, we restricted ourselves to only two types of evaluation, namely, self- and outer-evaluation. In the self-evaluation, a neuron is evaluated for itself, while in the outer-evaluation, a neuron is evaluated by all its neighboring neurons. Contradiction between the two types of neurons is represented by the Kullback-Leibler divergence. Contradiction ratio in terms of the Kullback-Leibler divergence is reduced as much as possible.

We applied the method to the logistic and dollar-yen exchange rates. For the logistic data, experimental results confirmed that several explicit class boundaries which could

not be detected by the conventional self-organizing maps were detected by our method. In addition, better generalization performance was obtained without significantly degrading topographic preservation. In the dollar-yen exchange rates, class structure obtained by contradiction resolution was much better than that obtained by the conventional self-organizing maps. The best MSE was obtained by using contradiction resolution, though different criteria were used for comparison. At the least, the experimental results showed a possibility of better prediction performance. However, this improved performance was obtained by sacrificing quantization and topographic errors.

Though there are several problems in our method, such as the selection of optimal representations, it still, according to the experimental results, has shown its potential for visualizing data without sacrificing prediction performance. This shows that we can develop neural networks with better interpretation while keeping better predication performance.

REFERENCES

[1] R. Kamimura, "Contradiction resolution and its application to self-organizing maps," in *SMC, The 2012 International Joint Conference on*, IEEE, 2012.

[2] T. Kohonen, *Self-Organizing Maps*. Springer-Verlag, 1995.

[3] T. Kohonen, "The self-organization map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[4] T. Kohonen, "Self-organized formation of topological correct feature maps," *Biological Cybernetics*, vol. 43, pp. 59–69, 1982.

[5] J. W. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Transactions on Computers*, vol. C-18, no. 5, pp. 401–409, 1969.

[6] A. Ultsch and H. P. Siemon, "Kohonen self-organization feature maps for exploratory data analysis," in *Proceedings of International Neural Network Conference*, (Dordrecht), pp. 305–308, Kulwer Academic Publisher, 1990.

[7] A. Ultsch, "U*-matrix: a tool to visualize clusters in high dimensional data," Tech. Rep. 36, Department of Computer Science, University of Marburg, 2003.

[8] J. Vesanto, "SOM-based data visualization methods," *Intelligent Data Analysis*, vol. 3, pp. 111–126, 1999.

[9] S. Kaski, J. Nikkila, and T. Kohonen, "Methods for interpreting a self-organized map in data analysis," in *Proceedings of European Symposium on Artificial Neural Networks*, (Bruges, Belgium), 1998.

[10] I. Mao and A. K. Jain, "Artificial neural networks for feature extraction and multivariate data projection," *IEEE Transactions on Neural Networks*, vol. 6, no. 2, pp. 296–317, 1995.

[11] C. De Runz, E. Desjardin, and M. Herbin, "Unsupervised visual data mining using self-organizing maps and a data-driven color mapping," in *Information Visualisation (IV), 2012 16th International Conference on*, pp. 241–245, IEEE, 2012.

[12] S. Shieh and I. Liao, "A new approach for data clustering and visualization using self-organizing maps," *Expert Systems with Applications*, 2012.

[13] H. Yin, "ViSOM-a novel method for multivariate data projection and structure visualization," *IEEE Transactions on Neural Networks*, vol. 13, no. 1, pp. 237–243, 2002.

[14] M.-C. Su and H.-T. Chang, "A new model of self-organizing neural networks and its application in data projection," *IEEE Transactions on Neural Networks*, vol. 123, no. 1, pp. 153–158, 2001.

[15] S. Wu and T. Chow, "Prsom: A new visualization method by hybridizing multidimensional scaling and self-organizing map," *Neural Networks, IEEE Transactions on*, vol. 16, no. 6, pp. 1362–1380, 2005.

[16] L. Xu, Y. Xu, and T. W. Chow, "PolSOM-a new method for multidimentional data visualization," *Pattern Recognition*, vol. 43, pp. 1668–1675, 2010.

[17] Y. Xu, L. Xu, and T. Chow, "Pposom: A new variant of polsom by using probabilistic assignment for multidimensional data visualization," *Neurocomputing*, vol. 74, no. 11, pp. 2018–2027, 2011.

[18] L. Xu and T. Chow, "Multivariate data classification using polsom," in *Prognostics and System Health Management Conference (PHM-Shenzhen), 2011*, pp. 1–4, IEEE, 2011.

[19] M. Bogdan and W. Rosenstiel, "Detection of cluster in self-organizing maps for controlling a prostheses using nerve signals," in *9th European Symposium on Artificial Neural Networks. ESANN' 2001. Proceedings. D-Facto, Evere, Belgium*, pp. 131–6, 2001.

[20] D. Brugger, M. Bogdan, and W. Rosenstiel, "Automatic cluster detection in kohonen's som," *Neural Networks, IEEE Transactions on*, vol. 19, no. 3, pp. 442–459, 2008.

[21] T. Haraguchi, H. Matsushita, and Y. Nishio, "Community self-organizing map and its application to data extraction," in *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*, pp. 1107–1114, IEEE, 2009.

[22] Z. Li, R. Wang, and L. Chen, "Extracting community structure of complex networks by self-organizing maps," in *Proceedings of the third international symposium on optimization and systems biology (OSB' 09)*, pp. 48–56, 2009.

[23] D. E. Rumelhart, G. E. Hinton, and R. Williams, "Learning internal representations by error progagation," in *Parallel Distributed Processing* (D. E. Rumelhart and G. E. H. et al., eds.), vol. 1, pp. 318–362, Cambridge: MIT Press, 1986.

[24] L. I. Nord and S. P. Jacobsson, "A novel method for examination of the variable contribution to computational neural network models," *Chemometrics and Intelligent Laboratory Systems*, vol. 44, pp. 153–160, 1998.

[25] A. Micheli, A. Sperduti, and A. Starita, "Analysis of the internal representations developed by neural networks for structures applied to quantitative structure-activity relationship studies of benzodiazepines," *Journal of Chemical Information and Computer Sciences*, vol. 41, pp. 202–218, 2001.

[26] G. G. Towell and J. W. Shavlik, "Extracting refined rules from knowledge-based neural networks," *Machine learning*, vol. 13, pp. 71–101, 1993.

[27] M. Ishikawa, "Structural learning with forgetting," *Neural Networks*, vol. 9, no. 3, pp. 509–521, 1996.

[28] M. Ishikawa, "Rule extraction by successive regularization," *Neural Networks*, vol. 13, pp. 1171–1183, 2000.

[29] P. Howes and N. Crook, "Using input parameter influences to support the decisions of feedforward neural networks," *Neurocomputing*, vol. 24, p. 1999, 1999.

[30] R. Feraud and F. Clerot, "A methodology to explain neural network classification," *Neural Networks*, vol. 15, pp. 237–246, 2002.

[31] R. Setiono, W. K. Leow, and J. M. Zurada, "Extracting of rules from artificial neural networks for nonlinear regression," *IEEE Transactions on Neural Networks*, vol. 13, no. 3, pp. 564–577, 2002.

[32] R. Kamimura, "Self-enhancement learning: target-creating learning and its application to self-organizing maps," *Biological cybernetics*, pp. 1–34, 2011.

[33] M. Orr *et al.*, "Introduction to radial basis function networks," *University of Edinburg*, 1996.

[34] M. Orr, "Recent advances in radial basis function networks," *Relatório técnico, Centre for Cognitive Science, University of Edinburgh*, 1999.

[35] M. Orr, "Regularization in the selection of radial basis function centers," *Neural computation*, vol. 7, no. 3, pp. 606–623, 1995.

[36] K. Kiviluoto, "Topology preservation in self-organizing maps," in *In Proceedings of the IEEE International Conference on Neural Networks*, pp. 294–299, 1996.

[37] J. Vesanto, J. Himberg, E. Alhoniemi, and J. Parhankangas, "SOM toolbox for Matlab," tech. rep., Laboratory of Computer and Information Science, Helsinki University of Technology, 2000.

AUTHOR PROFILE

Ryotaro Kamimura is currently a professor of IT Education Center and Graduate School of Science and Technology of Tokai University in Japan. His research interests are information-theoretic approach to neural computing.