Fifth International Workshop on Computational Intelligence & Applications
IEEE SMC Hiroshima Chapter, Hiroshima University, Japan, November 10, 11 & 12, 2009

Self-Organized Similarity based Kernel Fuzzy Clustering Model and Its Applications

Tomoyuki Kuwata
University of Tsukuba
Tsukuba, Ibaraki 305-8573
email:tkuwata@sk.tsukuba.ac.jp

Mika Sato-Ilic
University of Tsukuba
Tsukuba, Ibaraki 305-8573
email:mika@risk.tsukuba.ac.jp

***Abstract*—The purpose of this paper is to improve the performance of the kernel fuzzy clustering model by introducing a self-organized algorithm. A conventional kernel fuzzy clustering model is defined as a model which is an improved additive fuzzy clustering. The purpose of this conventional model is to obtain a clearer result by consideration of the interaction of clusters. This paper proposes a fuzzy clustering model based on the idea of self-organized dissimilarity between two objects.**

I. INTRODUCTION

Recently, in the area of data analysis there is a tremendous amount of interests in the analyses of noisy data. As a technique for analyzing these data, clustering techniques are well known. Clustering has several methods, including exploring data structure directly from the data, and another based on the model assumption. [1], [2], [3] This paper focuses on the later.

In this paper, we discuss a fuzzy clustering model using self-organized similarity. [4] The proposed model is an extension of a kernel fuzzy clustering model. [5] The kernel fuzzy clustering model classifies data, using similarity between two objects. This model estimates the degree of belongingness of objects to clusters in a mapped space by using kernel function. We define a self-organized similarity between objects from the classification structure under an assumption that similar objects have similar classification structures.

The proposed model estimates the degree of belongingness using the kernel fuzzy clustering model, calculates the self-organized similarity from the results of kernel fuzzy clustering model, and repeatedly estimates the degree of belongingness from the self-organized similarity.

The purpose of the kernel fuzzy clustering model [5] is to obtain a clear result by consideration of the interaction of clusters. However, if the data have noise, the result of the model tends to be uniformity. In order to solve this problem, we use the self-organized similarity for the kernel fuzzy clustering model and investigates the performance with several numerical examples.

II. KERNEL FUZZY CLUSTERING MODEL

The kernel fuzzy clustering model is defined as follows:

$$s_{ij} = \kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) + \varepsilon_{ij}. \tag{1}$$

s_{ij} is similarity between i-th object and j-th object. Degree of belongingness of i-th object is $\boldsymbol{u}_i = (u_{i1}, \cdots, u_{iK})$, and u_{ik} is a degree of belongingness of i-th object with respect to k-th cluster. u_{ik} satisfies the following conditions:

$$u_{ik} \geq 0, \ \sum_{k=1}^{K} u_{ik} = 1. \tag{2}$$

κ is a kernel function which satisfies the following conditions:

$$\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) = \kappa(\boldsymbol{u}_j, \boldsymbol{u}_i). \tag{3}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \kappa(\boldsymbol{u}_i, \boldsymbol{u}_j)\boldsymbol{u}_i\boldsymbol{u}_j \geq 0, \ \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in R^K. \tag{4}$$

When κ satisfies the condition shown in equation (5), the model (1) is a non-linear clustering model which includes the conventional additive fuzzy clustering model [3], when parameter d is 1.

$$\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) = \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle^d, \ d \geq 1. \tag{5}$$

In the additive fuzzy clustering model, sum of each cluster $(\sum_{k=1}^{K} u_{ik}u_{jk})$ explains similarity between objects. But, in the kernel fuzzy clustering model, not only sum of each cluster but also interaction of clusters $u_{ik}u_{jl}$ explains similarity between objects, when $d > 1$ in equation (5).

The additive fuzzy clustering model estimates u_{ik} in lower dimensional space (K dimensional space). In contrast, the kernel fuzzy clustering model estimates u_{ik} in higher dimensional space, when $d > 1$ in equation (5).

The kernel fuzzy clustering model defined in equation (1) consists of applying similarity data to equation (1) and estimating u_{ik} which minimize the sum of squared erros F as follows:

$$F = \sum_{i=1}^{n}\sum_{j=1}^{n}(s_{ij} - \sum_{k=1}^{K} \kappa(\boldsymbol{u}_i, \boldsymbol{u}_j))^2.$$

III. SELF-ORGANIZED KERNEL FUZZY CLUSTERING MODEL

We propose a self-organized kernel fuzzy clustering model which consists of a kernel fuzzy clustering model and self-organized similarity. The model is defined as follows:

$$\bar{s}_{ij} = \kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) + \bar{\varepsilon}_{ij}. \tag{6}$$

In this model, $\bar{s}_{ij}$ shows the self-organized similarity between i-th object and j-th object.

We obtain the self-organized similarity based on an assumption in which if dissimilarity of objects is similar to each other, classification structure obtained in solution space is also similar. We define the self-organized similarity as follows:

$$\bar{S} = UU^t, \ \bar{S} = (\bar{s}_{ij}). \tag{7}$$

$U = (u_{ik})$ is a matrix estimated by using the kernel fuzzy clustering model shown in equation (1).

Figure 1 shows the change of values of self-organized similarity shown in equation (7) by the change of the values of degree of belongingness of clusters, when the number of clusters is two.

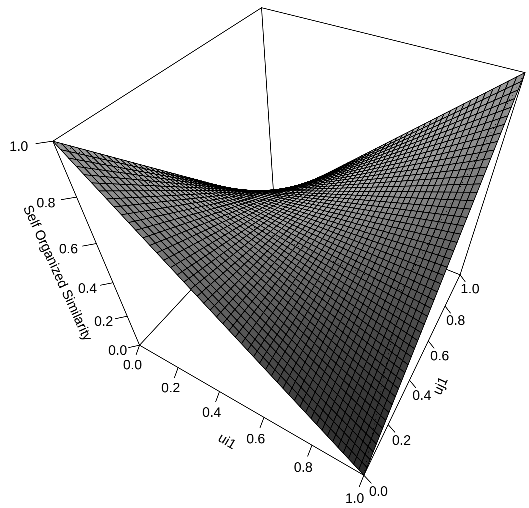

Fig. 1. Self-Organized Similarity between i-th Object and j-th Object

A value of the self-organized similarity between objects i and j

$$\bar{s}_{ij} = \sum_{k=1}^{2} u_{ik} u_{jk} \tag{8}$$

depends on u_{i1} and u_{j1} from the condition $\sum_{k=1}^{2} u_{ik} = 1$ shown in equation (2). Therefore, we show the value of change for $\bar{s}_{ij}$ with respect to the change of values of u_{i1} and u_{j1}.

In figure 1, the first axis shows the value of u_{i1}, the second axis shows the value of u_{j1} and the third axis shows the value of $\bar{s}_{ij}$ shown in equation (8).

The common line of the curved surface of $\bar{s}_{ij}$ and plane of $u_{i1} = u_{j1}$ is a downward convex line, and the minimum value of the line is 0.5 when u_{i1} and u_{j1} are 0.5. That is, if two objects have the same classification structures, then objects which have crisper classification structures have a larger value of the self-organized similarity.

Additionally, the common line of the curved surface of $\bar{s}_{ij}$ and plane of $u_{i1} + u_{j1} = 1$ is a convex rising line. And the maximum value of the line is 0.5 when u_{i1} and u_{j1} are 0.5, the minimum value is 0 when (u_{i1} is 1 and u_{j1} is 0) or (u_{i1} is 0 and u_{j1} is 1). That is, larger difference of classification structures between objects causes the smaller value of self-organized similarity between the objects.

Thus, the self-organized similarity measures similarity of the classification structures between two objects. Moreover, the self-organized similarity is a weighted similarity in which the weight becomes larger when the classification structure is crisper.

The proposed model defined in equation (6) consists of the following 5 steps.

1. Apply similarity data to equation (6), obtain the solution u_{ik}.
2. Using the obtained u_{ik}, calculate the self-organized similarity $\bar{S}$.
3. Apply $\bar{S}$ to equation (6), obtain solution $\bar{U} = (\bar{u}_{ik})$.
4. Calculate $\|\boldsymbol{U} - \bar{\boldsymbol{U}}\|$. $\|\cdot\|$ means norm.
5. If $\|\boldsymbol{U} - \bar{\boldsymbol{U}}\| < \varepsilon$ then stop. Otherwise calculate self-organized similarity by using $\bar{U}$, and repeat from step 3.

IV. NUMERICAL EXAMPLES

The data is artificial data which consists of 100 objects, and each object has 2 variables. The values of 50 objects are obtained as random values from the normal distribution with $\boldsymbol{\mu} = (4, 4)^t, \ \sigma = 2I$. The values of the other 50 objects are values from the normal distribution with $\boldsymbol{\mu} = (-4, -4)^t, \ \sigma = 2I$. Figure 2 shows the artificial data.

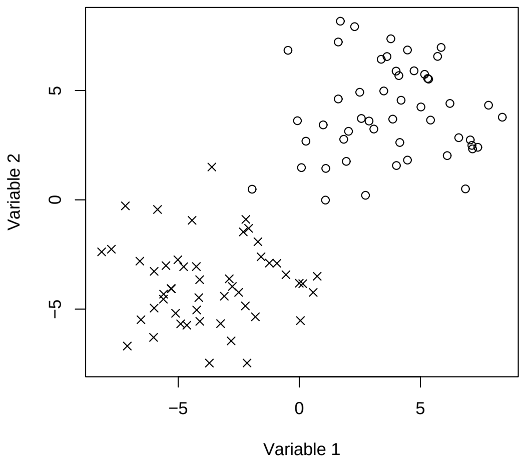

Fig. 2. Artificial Data

Figure 3 shows a result by using the kernel fuzzy clustering model on the artificial data. In this figure, the abscissa shows the value of variable 1, and the ordinate shows the degree of belongingness of cluster 1. We applied the proposed model to the artificial data. The first time, the model did not converge, so we repeated from step 3 once more. This result is shown in figure 4. Figure 5 shows a result which is repeated once more from the result of figure 4. Figure 6 shows a converged result. From these results, we can see that the degree of belongingness using the self-organized similarity is gradually crisper than the result of the kernel fuzzy clustering model shown in figure 3.

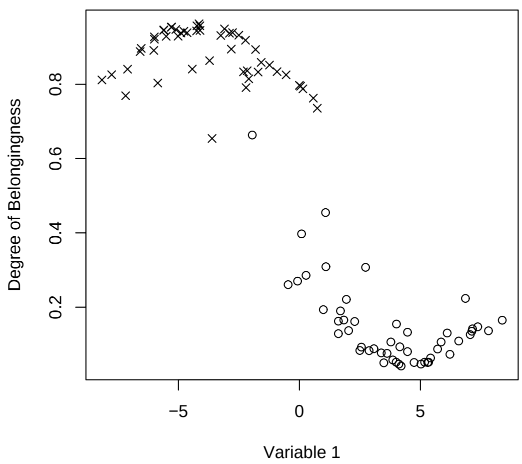

Fig. 3. Result of Kernel Fuzzy Clustering

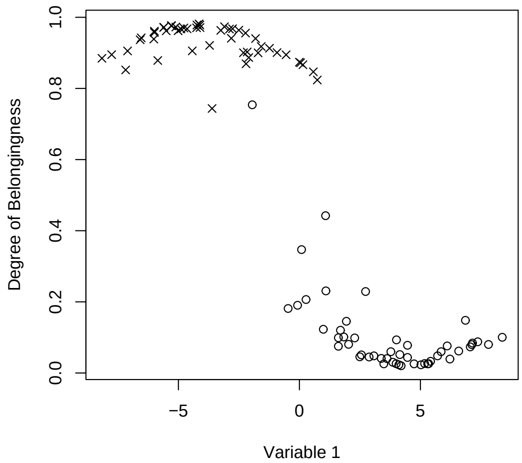

Fig. 4. Result of Self-Organized Kernel Fuzzy Clustering (After One Iteration)

The next data is transport data of Japanese freight [6] which was surveyed in 2005. The value of the data shows the amount of freight moving among prefectures in Japan. The data is shown by a 47 $\times$ 47 matrix. Row shows starting point, column shows destination. x_{ij} shows amount of freight from a prefecture i to a prefecture j.

We apply this data for the kernel fuzzy clustering model and the proposed model. The number of clusters is assumed to be 3. Since the data is obtained as an asymmetric matrix, in order to apply these models, we transform the asymmetric data to a symmetric data.

1. Calculate averages for i th row and j th column, which are r_i and c_j, and obtain normalized matrix $\hat{X} = (\hat{x}_{ij})$ as follows:

$$\hat{x}_{ij} = \frac{x_{ij}}{r_i \times c_j}.$$

2. Calculate symmetric data Y by using the following equation:

$$Y = \frac{\hat{X} + t(\hat{S})}{2}.$$

3. For obtaining similarity matrix $S = (s_{ij})$ in which each element is scaled in an interval [0,1], calculate following equation:

$$s_{ij} = 1 - \frac{1}{exp(\frac{y_{ij}}{\mu})},$$

where μ is mean of $y_{ij}, \forall i, j$.

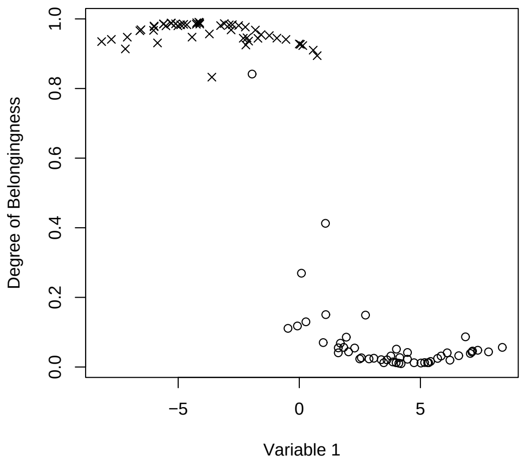

Fig. 5. Result of Self-Organized Kernel Fuzzy Clustering (After Two Iteration)

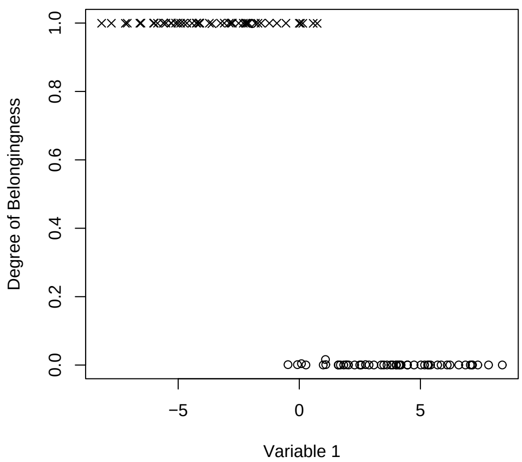

Fig. 6. Result of Self-Organized Kernel Fuzzy Clustering (After Converged)

Figures 7 - 9 show results of the kernel fuzzy clustering model. Figure 7 shows a result of cluster 1, figure 8 shows a result of cluster 2, and figure 9 shows a result of cluster

3. Darker colors in these figures show a larger degree of belongingness of clusters. Figures 10 - 12 show results of the proposed model after one iteration. Figures 13 - 15 show results of the proposed model after convergence.

From these results of the kernel fuzzy clustering model shown in figures 7 - 9, it can be seen that objects of the Kinki district mainly belong to cluster 1, objects of the Kanto and Tohoku districts mainly belong to cluster 2, and objects of the Kyushu and Tohoku districts mainly belong to cluster 3.

From the result of the proposed model (after one iteration), it can be seen that we obtain result in which objects mainly classified into 2 clusters. And, from the result of the proposed model (after converged), it can be seen that all objects belong to cluster 1 which is not good. However, we can see the tendency of the clustering model result becoming crisper. For practical use, we need to stop before we obtain the results shown in figures 13 - 15.

From this point, selecting the adaptable stopping threshold ε is an important issue. By selecting the adaptable stopping threshold, an adaptable number of clusters can be automatically selected, since the result shows that the proposed model tends to eliminate unexplainable and noisy clusters.

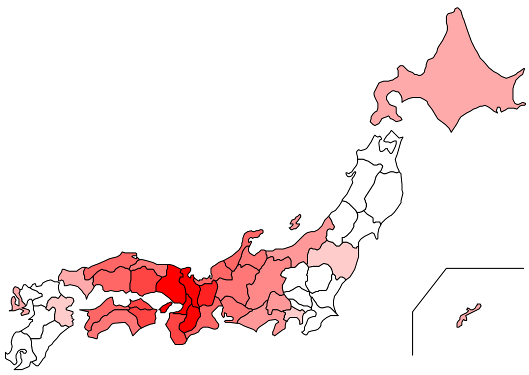

Fig. 7. Result of Cluster 1 Kernel Fuzzy Clustering

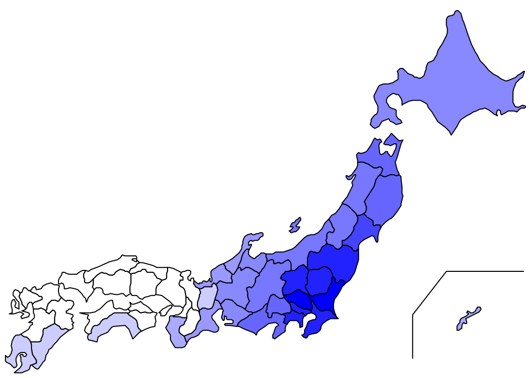

Fig. 8. Result of Cluster 2 Kernel Fuzzy Clustering

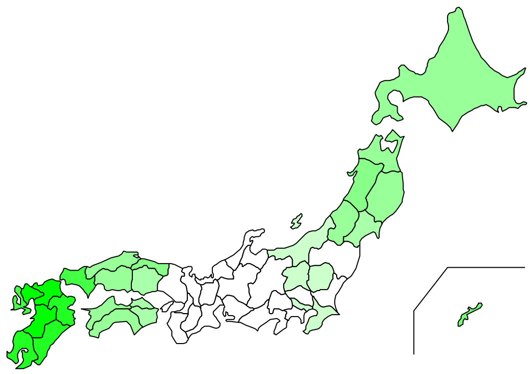

Fig. 9. Result of Cluster 3 Kernel Fuzzy Clustering

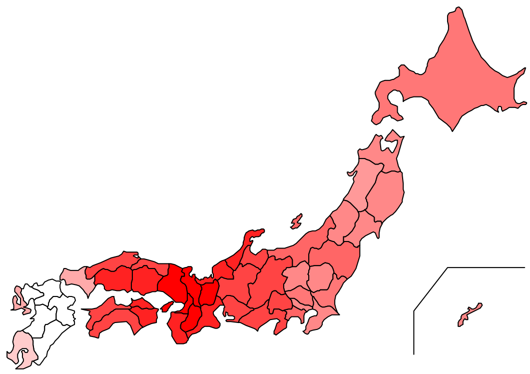

Fig. 10. Result of Cluster 1 Self-Organized Kernel Fuzzy Clustering (After One Iteration)

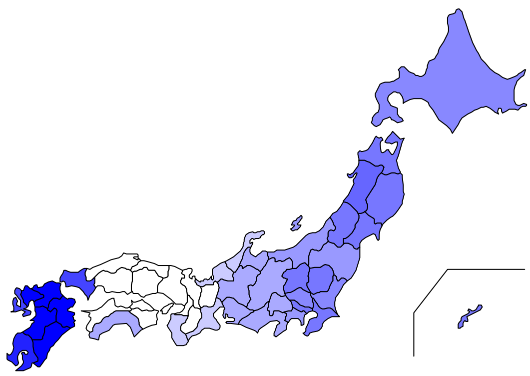

Fig. 11. Result of Cluster 2 Self-Organized Kernel Fuzzy Clustering (After One Iteration)

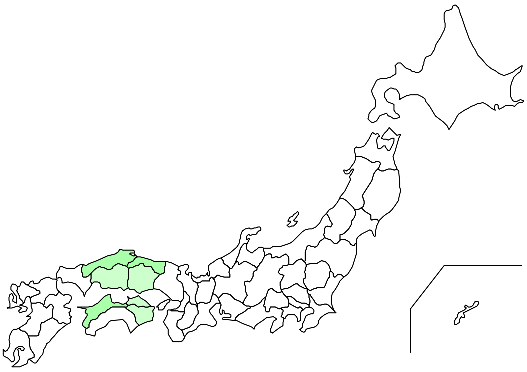

Fig. 12. Result of Cluster 3 Self-Organized Kernel Fuzzy Clustering (After One Iteration)

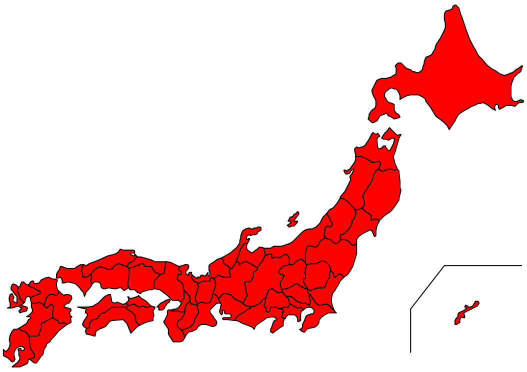

Fig. 13. Result of Cluster 1 Self-Organized Kernel Fuzzy Clustering (After Converged)

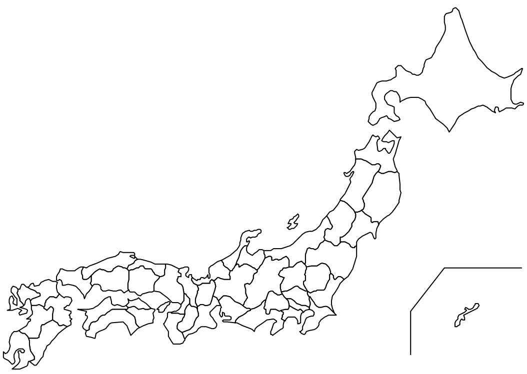

Fig. 14. Result of Cluster 2 Self-Organized Kernel Fuzzy Clustering (After Converged)

V. CONCLUSION

We propose a self-organized kernel fuzzy clustering model. This model includes self-organized similarity which calculates similarity of the classification structures between two objects.

The kernel fuzzy clustering model tends to obtain a crisper result for noisy data when compared with a conventional fuzzy clustering model. However, if the data have uniformity struc-

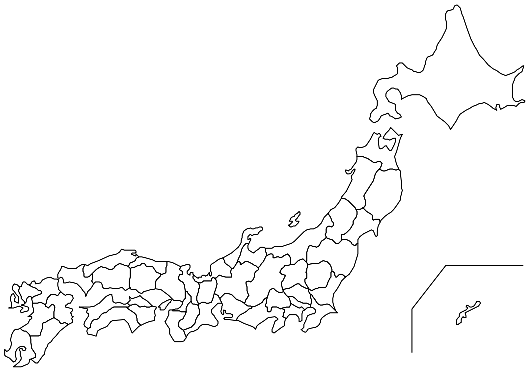

Fig. 15. Result of Cluster 3 Self-Organized Kernel Fuzzy Clustering (After Converged)

ture, the two models tend to obtain similar results. In order to solve this problem, we use the self-organized similarity for the kernel fuzzy clustering model.

Numerical examples show a better performance.

ACKNOWLEDGMENT

This research was supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Scientific Research (C) (20500252).

REFERENCES

[1] L. Fahrmeir, *Multivariate Statistical Modeling Based on Generalized Linear Models*, Springer, 1994.

[2] T.J. Hastie and R.J. Tibshirani, *Generalized Additive Models*, Champman & Hall, 1996.

[3] M. Sato, Y. Sato and L.C. Jain, *Fuzzy Clustering Models and Applications*, Springer, 1997.

[4] M. Sato-Ilic, Self Organized Fuzzy Clustering, *Intelligent Engineering Systems through Artificial Neural Networks*, Vol. 14, pp. 579-584, 2004.

[5] M. Sato-Ilic and S. Ito, Kernel Fuzzy Clustering Model, *24th Fuzzy System Symposium*, pp. 153-154, 2008, (in Japanese).

[6] http://www.mlit.go.jp/index.html, Ministry of Land, Infrastructure, Transport and Tourism