

Research Article

Deterministic Annealing Approach to Fuzzy C-Means Clustering Based on Entropy Maximization

Makoto Yasuda

Department of Electrical and Computer Engineering, Gifu National College of Technology, Kamimakuwa 2236-2, Motosu, Gifu 501-0495, Japan

Correspondence should be addressed to Makoto Yasuda, yasuda@gifu-nct.ac.jp

Received 13 June 2011; Revised 26 August 2011; Accepted 27 August 2011

Academic Editor: Salvatore Sessa

Copyright © 2011 Makoto Yasuda. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper is dealing with the fuzzy clustering method which combines the deterministic annealing (DA) approach with an entropy, especially the Shannon entropy and the Tsallis entropy. By maximizing the Shannon entropy, the fuzzy entropy, or the Tsallis entropy within the framework of the fuzzy c -means (FCM) method, membership functions similar to the statistical mechanical distribution functions are obtained. We examine characteristics of these entropy-based membership functions from the statistical mechanical point of view. After that, both the Shannon- and Tsallis-entropy-based FCMs are formulated as DA clustering using the very fast annealing (VFA) method as a cooling schedule. Experimental results indicate that the Tsallis-entropy-based FCM is stable with very fast deterministic annealing and suitable for this annealing process.

1. Introduction

Statistical mechanics investigates the macroscopic properties of a physical system consisting of several elements. Recently, research activities that attempt to apply statistical mechanical models or tools to information science have become popular.

The deterministic annealing (DA) method [1] is a deterministic variant of the simulated annealing (SA) method [2]. DA characterizes the minimization problem of the cost function as the minimization of the free energy, which depends on temperature and tracks its minimum while decreasing the temperature, and thus it can deterministically optimize the cost function at each temperature. Hence, DA is more efficient than SA, but does not guarantee a global optimal solution.

There exists a strong relationship between the membership functions of the fuzzy c -means (FCM) clustering [3] with the maximum entropy or entropy regularization methods [4, 5] and the statistical mechanical distribution function. That is, FCM regularized with the Shannon entropy gives a membership function similar to the Boltzmann (or Gibbs) distribution function [1, 4], and FCM regularized with the fuzzy entropy [6] gives a membership function similar to the Fermi-Dirac distribution function [7]. These

membership functions are suitable for the annealing methods because they contain a parameter corresponding to the system temperature.

Tsallis [8] achieved nonextensive extension of the Boltzmann-Gibbs statistics. Tsallis postulated a generalization form of entropy with a generalization parameter q , which, in a limit of $q \rightarrow 1$, reaches the Shannon entropy. Later on, Ménard et al. [9] derived a membership function by regularizing FCM with the Tsallis entropy.

In this study, the membership function which takes the familiar form of the statistical mechanical distribution function is derived by maximizing the Shannon and fuzzy entropy within the framework of FCM. Similarly, the Tsallis entropy-based FCM membership function is derived [10, 11] by maximizing the Tsallis entropy. Then, the formulations of the free energy for these membership functions are calculated and examined from the statistical mechanical viewpoint.

On the other hand, there are some representative cooling schedules of the temperature for SA; for example, inversely proportional to a logarithmic function and inversely proportional to exponential function are well adopted. Rosen [12] proposed the more effective method for SA known as very fast annealing (VFA).

However, an applicability of VFA to DA is not known yet. In order to achieve good clustering by DA, a reliable annealing process is desirable. Therefore, by introducing VFA to DA, we formulate the Shannon- and Tsallis-entropy based FCMs as very fast DA clustering, to examine their reliabilities.

Experiments are performed on the numerical and iris data [13], and the obtained results indicate that Tsallis-entropy-based FCM clustering is suitable for very fast DA clustering because of its shape of the membership function.

2. Entropy Maximization Method

Let $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ ($\mathbf{x}_k = (x_k^1, \dots, x_k^p) \in R^p$) be a data set in the p -dimensional real space, which should be divided into c clusters. In addition, let $V = \{\mathbf{v}_1, \dots, \mathbf{v}_c\}$ ($\mathbf{v}_i = (v_i^1, \dots, v_i^p)$) be the centers of clusters, and let $u_{ik} \in [0, 1]$ ($i = 1, \dots, c; k = 1, \dots, n$) be the membership functions. Furthermore, let

$$J = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m d_{ik} \quad (m > 1) \quad (1)$$

be the objective function of FCM, where $d_{ik} = \|\mathbf{x}_k - \mathbf{v}_i\|^2$.

2.1. Shannon Entropy Maximization of FCM. First, we introduce the Shannon entropy into the FCM clustering. The Shannon entropy is given by

$$S = - \sum_{k=1}^n \sum_{i=1}^c u_{ik} \log u_{ik}. \quad (2)$$

Under the normalization constraint of

$$\sum_{i=1}^c u_{ik} = 1 \quad (\forall k), \quad (3)$$

and setting m to 1, the fuzzy entropy functional is given by

$$\delta S - \sum_{k=1}^n \alpha_k \delta \left(\sum_{i=1}^c u_{ik} - 1 \right) - \beta \sum_{k=1}^n \sum_{i=1}^c \delta(u_{ik} d_{ik}), \quad (4)$$

where α_k and β are the Lagrange multipliers and α_k must be determined so as to satisfy (3). The stationary condition for (4) leads to the following membership function

$$u_{ik} = \frac{e^{-\beta d_{ik}}}{\sum_{j=1}^c e^{-\beta d_{jk}}} \quad (5)$$

and the cluster centers

$$\mathbf{v}_i = \frac{\sum_{k=1}^n u_{ik} \mathbf{x}_k}{\sum_{k=1}^n u_{ik}}. \quad (6)$$

2.2. Fuzzy Entropy Maximization of FCM. We then introduce the fuzzy entropy into the FCM clustering.

The fuzzy entropy is given by

$$\hat{S} = - \sum_{k=1}^n \sum_{i=1}^c \{ \hat{u}_{ik} \log \hat{u}_{ik} + (1 - \hat{u}_{ik}) \log(1 - \hat{u}_{ik}) \}. \quad (7)$$

The fuzzy entropy functional is given by

$$\delta \hat{S} - \sum_{k=1}^n \alpha_k \delta \left(\sum_{i=1}^c \hat{u}_{ik} - 1 \right) - \beta \sum_{k=1}^n \sum_{i=1}^c \delta(\hat{u}_{ik} d_{ik}), \quad (8)$$

where α_k and β are the Lagrange multipliers [14]. The stationary condition for (8) leads to the following membership function:

$$\hat{u}_{ik} = \frac{1}{e^{\alpha_k + \beta d_{ik}} + 1} \quad (9)$$

and the cluster centers

$$\mathbf{v}_i = \frac{\sum_{k=1}^n \hat{u}_{ik} \mathbf{x}_k}{\sum_{k=1}^n \hat{u}_{ik}}. \quad (10)$$

In (9), β defines the extent of the distribution [7]. Equation (9) is formally normalized as

$$\hat{u}_{ik} = \frac{1}{e^{\alpha_k + \beta d_{ik}} + 1} / \sum_{j=1}^c \frac{1}{e^{\alpha_k + \beta d_{jk}} + 1}. \quad (11)$$

2.3. Tsallis Entropy Maximization of FCM. Let $\tilde{\mathbf{v}}_i$ and \tilde{u}_{ik} be the centers of clusters and the membership functions, respectively.

The Tsallis entropy is defined as

$$\tilde{S} = - \frac{1}{q-1} \left(\sum_{k=1}^n \sum_{i=1}^c \tilde{u}_{ik}^q - 1 \right), \quad (12)$$

where $q \in \mathbf{R}$ is any real number. The objective function is rewritten as

$$\tilde{U} = \sum_{k=1}^n \sum_{i=1}^c \tilde{u}_{ik}^q \tilde{d}_{ik}, \quad (13)$$

where $\tilde{d}_{ik} = \|\mathbf{x}_k - \tilde{\mathbf{v}}_i\|^2$.

Accordingly, the Tsallis entropy functional is given by

$$\delta \tilde{S} - \sum_{k=1}^n \alpha_k \delta \left(\sum_{i=1}^c \tilde{u}_{ik} - 1 \right) - \beta \sum_{k=1}^n \sum_{i=1}^c \delta(\tilde{u}_{ik}^q \tilde{d}_{ik}). \quad (14)$$

The stationary condition for (14) yields the following membership function:

$$\tilde{u}_{ik} = \frac{\{1 - \beta(1-q)\tilde{d}_{ik}\}^{1/(1-q)}}{\tilde{Z}}, \quad (15)$$

where

$$\tilde{Z} = \sum_{j=1}^c \{1 - \beta(1-q)\tilde{d}_{jk}\}^{1/(1-q)}. \quad (16)$$

In this case, the cluster centers are given by

$$\tilde{\mathbf{v}}_i = \frac{\sum_{k=1}^n \tilde{u}_{ik}^q \mathbf{x}_k}{\sum_{k=1}^n \tilde{u}_{ik}^q}. \quad (17)$$

In the limit of $q \rightarrow 1$, the Tsallis entropy recovers the Shannon entropy [8] and \tilde{u}_{ik} approaches u_{ik} in (5).

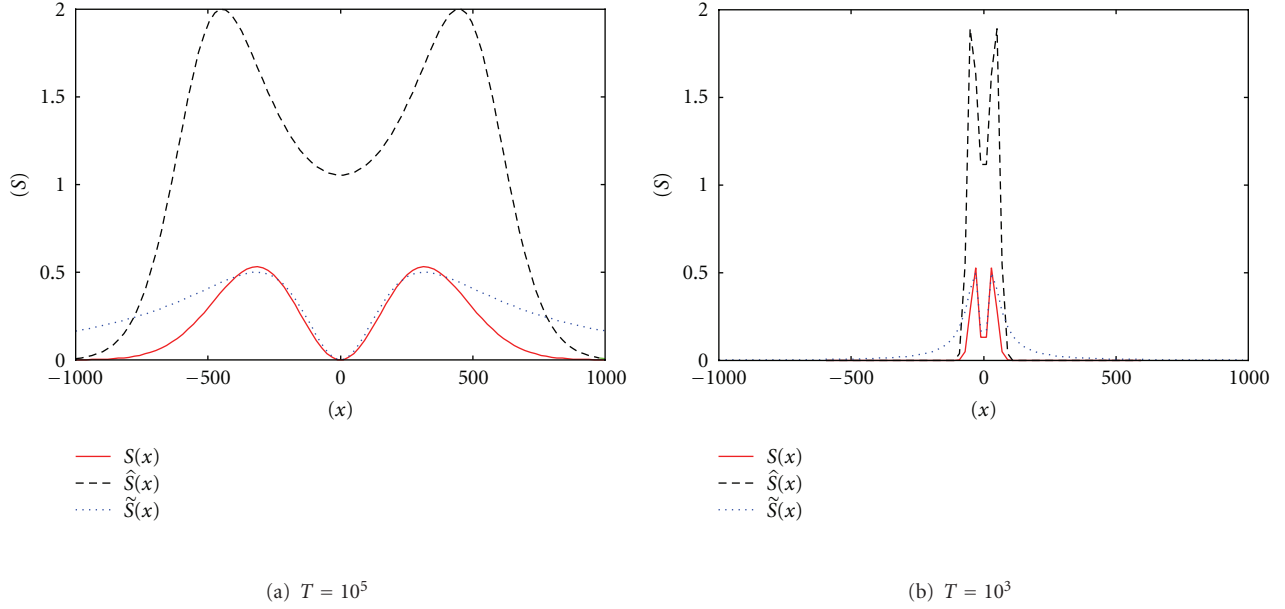


FIGURE 1: The plots of the entropy functions S , \bar{S} and \tilde{S} at (a) high and (b) low temperature ($n = 1, c = 2, q = 1.5, \alpha_k = -2$).

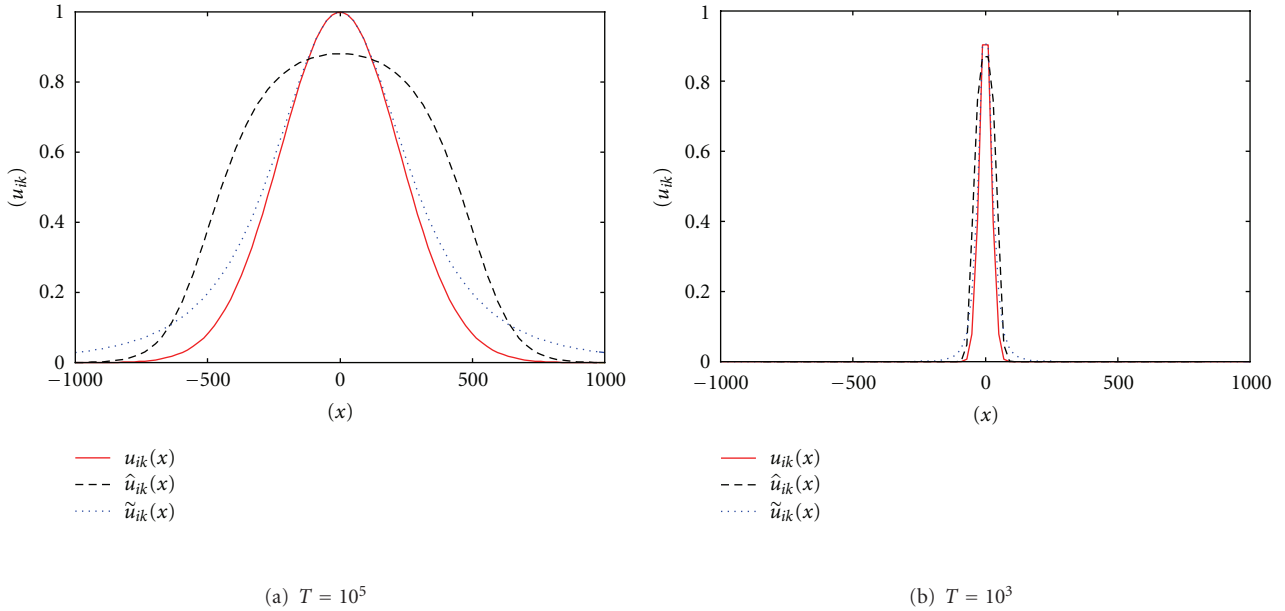


FIGURE 2: The plots of the membership functions u_{ik} , \bar{u}_{ik} and \tilde{u}_{ik} at (a) high and (b) low temperature ($n = 1, c = 2, q = 1.5, \alpha_k = -2$).

3. Statistical Mechanical Interpretation of Entropy-Based FCM

3.1. Shannon-Entropy-Based FCM Statistics. In the Shannon-entropy-based FCM, the sum of the states (the partition function) for the grand canonical ensemble of fuzzy clustering can be written as

$$Z = \prod_{k=1}^n \sum_{i=1}^c e^{-\beta d_{ik}}. \quad (18)$$

By substituting (18) for $F = -(1/\beta)(\log Z)$ [15], the free energy becomes

$$F = -\frac{1}{\beta} \sum_{k=1}^n \log \left\{ \sum_{i=1}^c e^{-\beta d_{ik}} \right\}. \quad (19)$$

Stable thermal equilibrium requires a minimization of the free energy. By formulating deterministic annealing as a minimization of the free energy, $\partial F / \partial \mathbf{v}_i = 0$ yields

$$\mathbf{v}_i = \frac{\sum_{k=1}^n u_{ik} \mathbf{x}_k}{\sum_{k=1}^n u_{ik}}. \quad (20)$$

This cluster center is the same as that in (6).

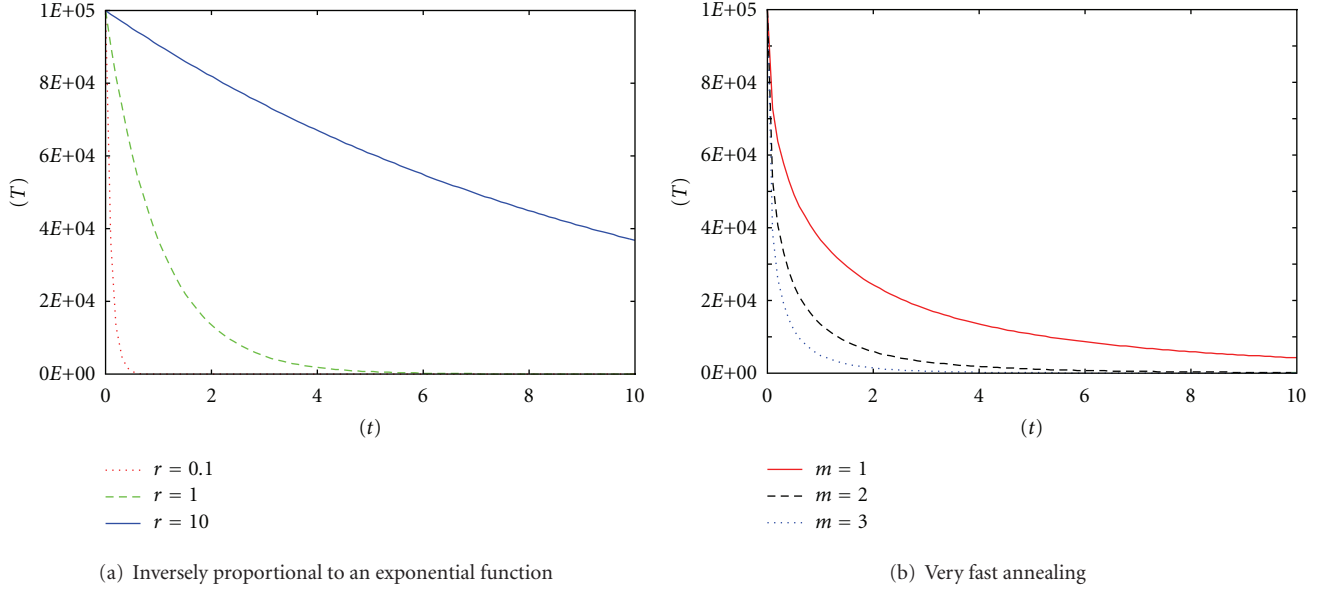


FIGURE 3: The plots of the cooling functions of (a) proportional to an exponential ($r = 0.1, 1, 10$) and (b) very fast annealing methods ($m = 1.0, 2.0, 3.0$).

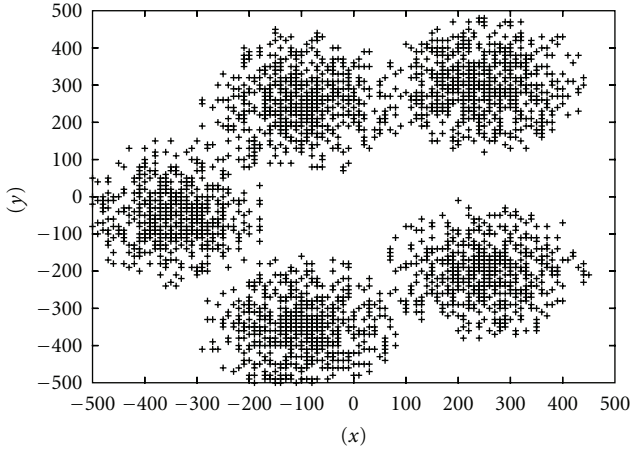


FIGURE 4: The numerical data.

3.2. Fuzzy-Entropy-Based FCM Statistics. In the fuzzy-entropy based FCM, by analogy with statistical mechanics, the grand partition function for the grand canonical ensemble of fuzzy clustering can be written as

$$\hat{\Xi} = \prod_{k=1}^n \prod_{i=1}^c (1 + e^{-\alpha_k - \beta d_{ik}}), \quad (21)$$

because data can belong to any cluster. By substituting (21) for $\hat{F} = -(1/\beta)(\log \hat{\Xi} - \alpha_k \partial \log \hat{\Xi} / \partial \alpha_k)$ [15], the free energy becomes

$$\hat{F} = -\frac{1}{\beta} \sum_{k=1}^n \left\{ \sum_{i=1}^c \log(1 + e^{-\alpha_k - \beta d_{ik}}) + \alpha_k \right\}. \quad (22)$$

It should be noted that $J_{m=1} - T\hat{S}$, the Legendre transform of the fuzzy entropy, gives the same form for the free energy.

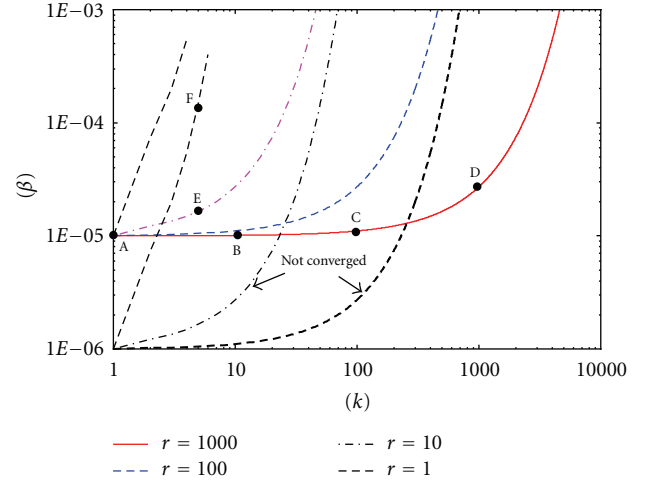


FIGURE 5: The inversely exponential cooling schedule of DA. The temperature decreases from T_{high} . (Inverse of temperature β increases from 1.0×10^{-5} or 1.0×10^{-6} to 1.0×10^{-3}) The curves are parameterized by the temperature reduction rate r .

3.3. The Tsallis-Entropy-Based FCM Statistics. On the other hand, \tilde{U} and \tilde{S} satisfy

$$\tilde{S} - \beta \tilde{U} = \sum_{k=1}^n \frac{\tilde{Z}^{1-q} - 1}{1 - q}, \quad (23)$$

which leads to

$$\frac{\partial \tilde{S}}{\partial \tilde{U}} = \beta. \quad (24)$$

Equation (24) makes it possible to regard β^{-1} as an artificial system temperature T [15]. Then, the free energy can be

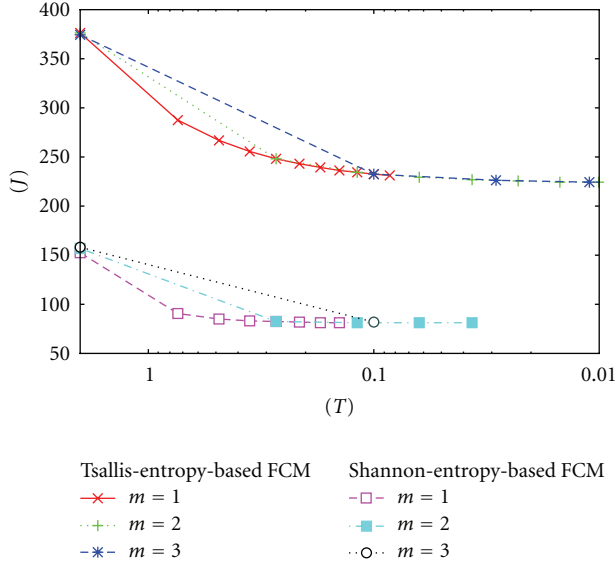


FIGURE 6: Reduction of the objective values for iris data with decreasing the temperature by VFA. The curves are parameterized by the temperature reduction rate m .

defined as

$$\tilde{F} = \tilde{U} - T\tilde{S} = -\frac{1}{\beta} \sum_{k=1}^n \frac{\tilde{Z}^{1-q} - 1}{1-q}. \quad (25)$$

\tilde{U} can be derived from \tilde{F} as

$$\tilde{U} = -\frac{\partial}{\partial \beta} \sum_{k=1}^n \frac{\tilde{Z}^{1-q} - 1}{1-q}. \quad (26)$$

$\partial \tilde{F} / \partial \tilde{v}_i = 0$ also gives

$$\tilde{v}_i = \frac{\sum_{k=1}^n \tilde{u}_{ik}^q \mathbf{x}_k}{\sum_{k=1}^n \tilde{u}_{ik}^q}. \quad (27)$$

4. Effects of Annealing Temperature

4.1. Dependency of Shapes of Membership Functions on Temperature. By reducing the temperature according to the annealing schedule, the deterministic annealing method achieves thermal equilibrium which minimizes the free energy. At absolute zero, the particle system settles down to the ground state, that is, the state of minimum energy. Figure 1 shows the forms of the entropy functions S , \hat{S} , and \tilde{S} . Figure 2 shows the forms of the membership functions u_{ik} , \hat{u}_{ik} , and \tilde{u}_{ik} .

In the deterministic annealing method, cluster distribution which minimizes the free energy is searched at the given temperature. At high temperature, the membership functions are widely distributed and clusters to which a data belongs are fuzzy. In case of \tilde{u}_{ik} with $q = 2$, the width of the distribution is roughly proportional to $\beta^{-0.5}$. At the limit of low temperature, on the other hand, fuzzy clustering reaches hard clustering. The relationship $F = U - TS$ suggests that

the higher temperature causes the larger entropy state, that is, chaotic state. This increase of the entropy is the result of the extent of the membership function.

In Figure 2, it can be seen that \hat{u}_{ik} has a flat peak, though both u_{ik} and \tilde{u}_{ik} have Gaussian forms. Also, it can be found that \tilde{u}_{ik} has a more gentle base slope than u_{ik} .

4.2. Cooling Schedule

4.2.1. Representative Annealing Methods. In SA, the temperature decreases according to a cooling schedule. The representative cooling schedules [16] for SA are

- (i) proportional to an exponential function

$$T = T_{\text{high}} r^t, \quad (28)$$

where T_{high} is a sufficiently high initial temperature, r is a parameter which defines a temperature reduction speed, and t is the number of iterations,

- (ii) inversely linear function

$$T = \frac{T_{\text{high}}}{t}, \quad (29)$$

- (iii) inversely proportional to a logarithmic function

$$T = \frac{T_{\text{high}}}{\ln(t)}, \quad (30)$$

- (iv) inversely proportional to exponential function

$$T = \frac{T_{\text{high}}}{e^{r^{-1}t}}. \quad (31)$$

4.2.2. Very Fast Annealing. Rosen proposed another inversely proportional to exponential function known as very fast annealing (VFA).

In VFA, T is given by

$$T = T_{\text{high}} e^{-mt^{(1/D)}}, \quad (32)$$

where m is a temperature reduction parameter and D is a dimension of a state space. Equations (31) and (32) are compared in Figure 3. It is observed that VFA initially decreases a temperature extremely.

In Section 6, we apply VFA as a cooling schedule of entropy based FCM clustering using DA.

5. Fuzzy C-Means as Clustering Algorithm Using Very Fast Annealing DA

The very fast deterministic annealing algorithm for the Tsallis-entropy-based FCM is given as follows.

- (1) Set the number of clusters c , the highest temperature T_{high} , the temperature reduction rate m , and the threshold of convergence test δ_1 and δ_2 ;
- (2) generate initial clusters at random positions and assign each data point to the nearest cluster. Set current temperature T to T_{high} ;

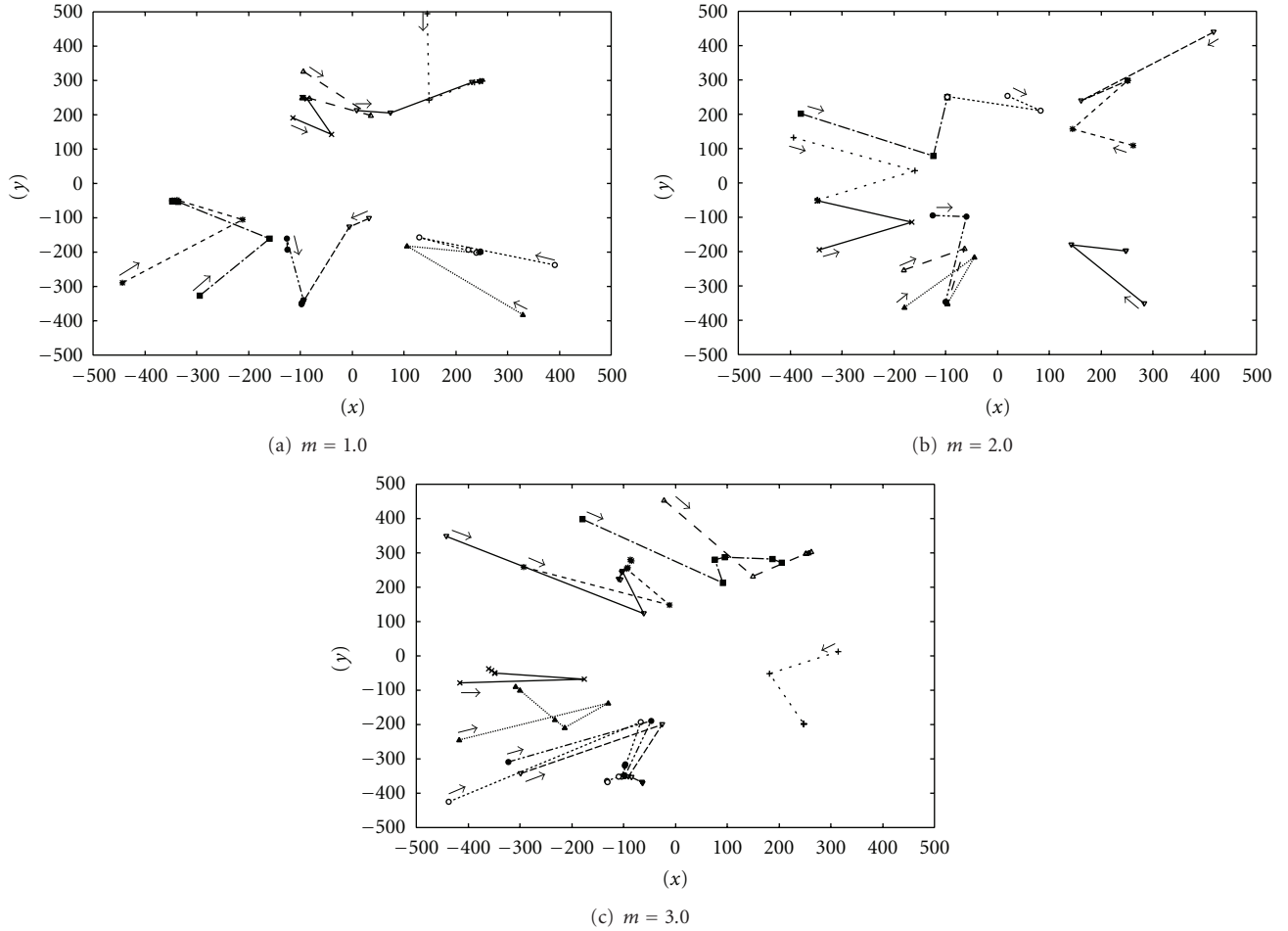


FIGURE 7: The shifts of cluster centers of Shannon entropy based clustering with decreasing the temperature by VFA ($m = 1.0, 2.0, 3.0$).

- (3) calculate \tilde{u}_{ik} by (15);
- (4) calculate cluster centers by (17);
- (5) compare the difference between the current centers and those obtained at the previous iteration $\hat{\mathbf{v}}_i$. If the convergence condition $\max_{1 \leq i \leq c} \|\mathbf{v}_i - \hat{\mathbf{v}}_i\| < \delta_1$ is satisfied, then go to (6), otherwise go back to (3);
- (6) if $\max_{1 \leq i \leq c} \|\mathbf{v}_i - \hat{\mathbf{v}}_i\| < \delta_2$ is satisfied, then stop. Otherwise decrease the temperature with (32) and go back to (3).

In case of Shannon-entropy-based FCM, (15) is replaced by (5) and (17) is replaced by (6), respectively.

6. Experiments

6.1. Experiment 1. In experiment 1, we generated five randomly placed clusters composed of 2,000 data points shown in Figure 4. We set c to be 10, δ_1 to be 50, and δ_2 to be 2 (measured by the scale of Figure 4). We also set $T_{\text{high}} = 1.0 \times 10^6$ ($\beta = 1.0 \times 10^{-6}$) or $T_{\text{high}} = 1.0 \times 10^5$ ($\beta = 1.0 \times 10^{-5}$).

First, we have applied the inversely exponential scheduling method to the Tsallis-entropy-based FCM clustering. The cooling schedule is illustrated in Figure 5. The changes of β

are parameterized by the temperature reduction rate r : from 1 to 1000.

At the higher levels of T (Figure 5 (A)), clusters are created near the center of gravity of data because β is comparatively small and the membership function extends over the whole data area and is extremely uniform. As T is lowered from Figure 5 (B) to (C), the width of the membership functions becomes narrower; that is, the Tsallis entropy decreases, and the associations become less fuzzy. And finally, the desired result is obtained.

In case of $r = 10$ or $r = 1$ (Figure 5 (E) or (F)), it is observed that \tilde{u}_{ik} and $\tilde{\mathbf{v}}_i$ converge more rapidly. In case of $T_{\text{high}} = 1.0 \times 10^6$ ($\beta = 1.0 \times 10^{-6}$), however, the initial distribution of \tilde{u}_{ik} becomes too wide and the algorithm is not converged with $r = 100$ and $r = 10$ (indicated by “not converged” in Figure 5). Thus, it is important to set T_{high} and r values properly.

To examine the effectiveness of VFA as a cooling schedule of DA, we made numerical experiments of the Shannon- and Tsallis-entropy-based FCM clustering.

The shifts of cluster centers with decreasing temperature are illustrated in Figures 7 and 8.

Initially, clusters are located randomly. At the higher levels of T , clusters move to near the center of gravity of

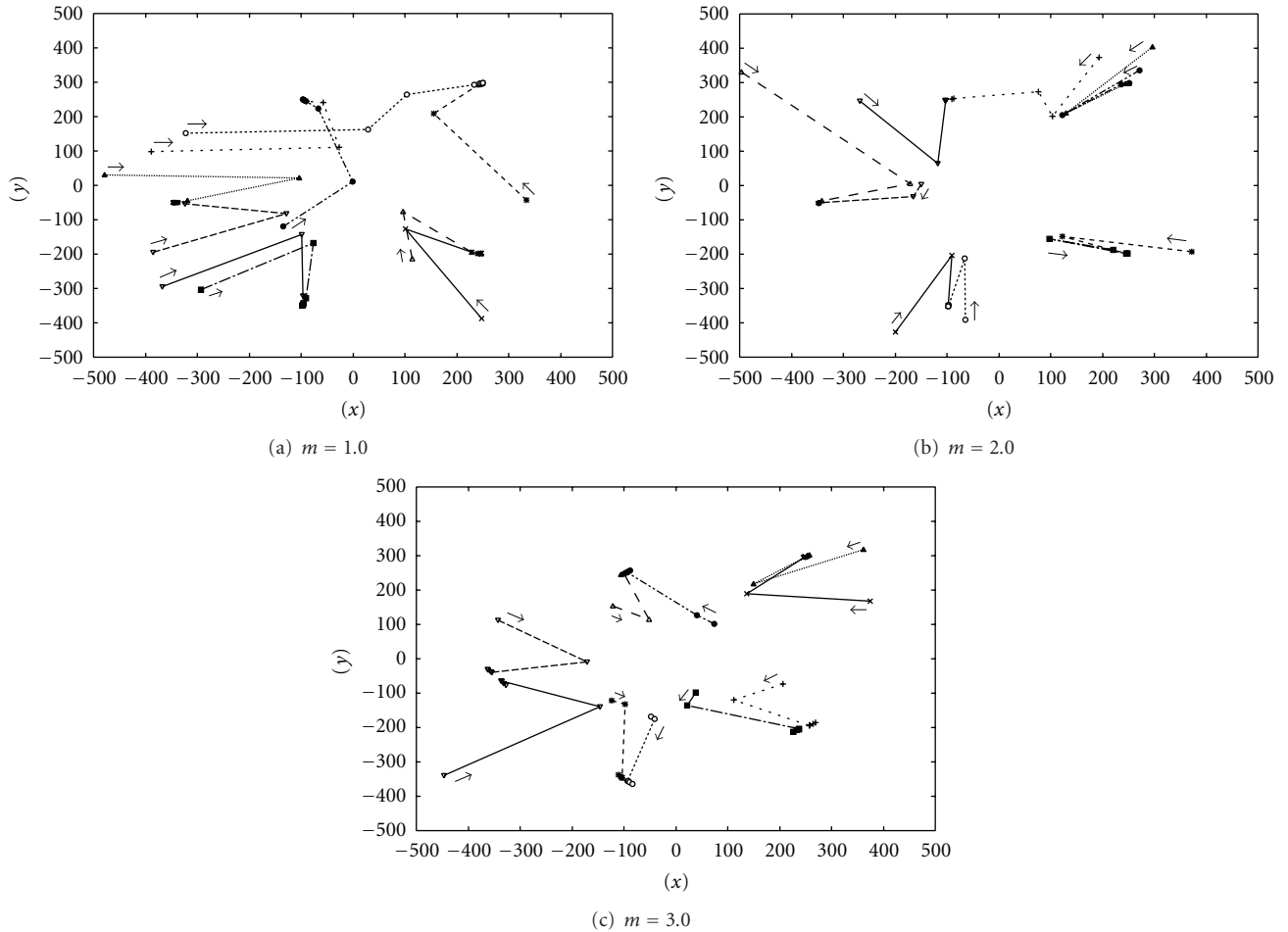


FIGURE 8: The shifts of cluster centers of Tsallis entropy based clustering with decreasing the temperature by VFA ($m = 1.0, 2.0, 3.0, q = 1.5$).

data because β is comparatively small and the membership function extends over the whole data area and is extremely uniform.

As T is lowered, the width of the membership functions becomes narrower and the associations of data become less fuzzy. In this process, in the Shannon-entropy-based FCM clustering, the clusters move to their nearest local data distribution centers. However, in the Tsallis-entropy-based FCM clustering, clusters can move a long distance to optimal positions because of their gentle base slopes.

Figures 9 and 10 illustrate the three-dimensional plots of u_{ik} and \tilde{u}_{ik} in the progress of very fast DA clustering.

At the higher temperature, roughness of \tilde{u}_{ik} is smaller than that of u_{ik} . After that, the shapes of both membership functions do not change greatly, because VFA reduces the temperature extremely only at the early annealing stage.

Consequently, because the Tsallis-entropy-based FCM has gentle slope in the region far from the origin, clusters can move long distance to optimal positions stably and the temperature can be reduced rapidly. This feature makes it possible to use VFA as a cooling schedule of DA for the Tsallis-entropy-based FCM. On the other hand, final cluster positions obtained by the Shannon-entropy-based FCM tend to depend on their initial positions.

TABLE 1: Comparison of minimum, maximum, and average values of misclassified iris data (100 trials).

Tsallis-entropy-based FCM				Shannon-entropy-based FCM			
m	Min.	Max.	Ave.	m	Min.	Max.	Ave.
1.0	14	14	14.00	1.0	15	16	15.01
2.0	14	15	14.01	2.0	11	13	12.97
3.0	14	15	14.68	3.0	11	14	13.59

6.2. *Experiment 2.* In experiment 2, the iris data set [13] consisting of 150 four-dimensional vectors of iris flowers are used. Three clusters of flowers detected are Versicolor, Virginia, and Setosa. Each cluster consists of 50 vectors.

The Shannon- and Tsallis-entropy-based FCM with DA are examined. VFA is used as a cooling schedule of DA. We set the parameters as follows: $c = 3$, $T_{\text{high}} = 2$, $\delta_1 = 0.1$, $\delta_2 = 0.01$, and $q = 1.5$.

The minimum, maximum, and average values of misclassified data of 100 trials are summarized in Table 1. The Shannon-entropy-based FCM gives slightly better results than the Tsallis-entropy-based FCM. However, it is found that the Tsallis-entropy-based FCM gives the best results

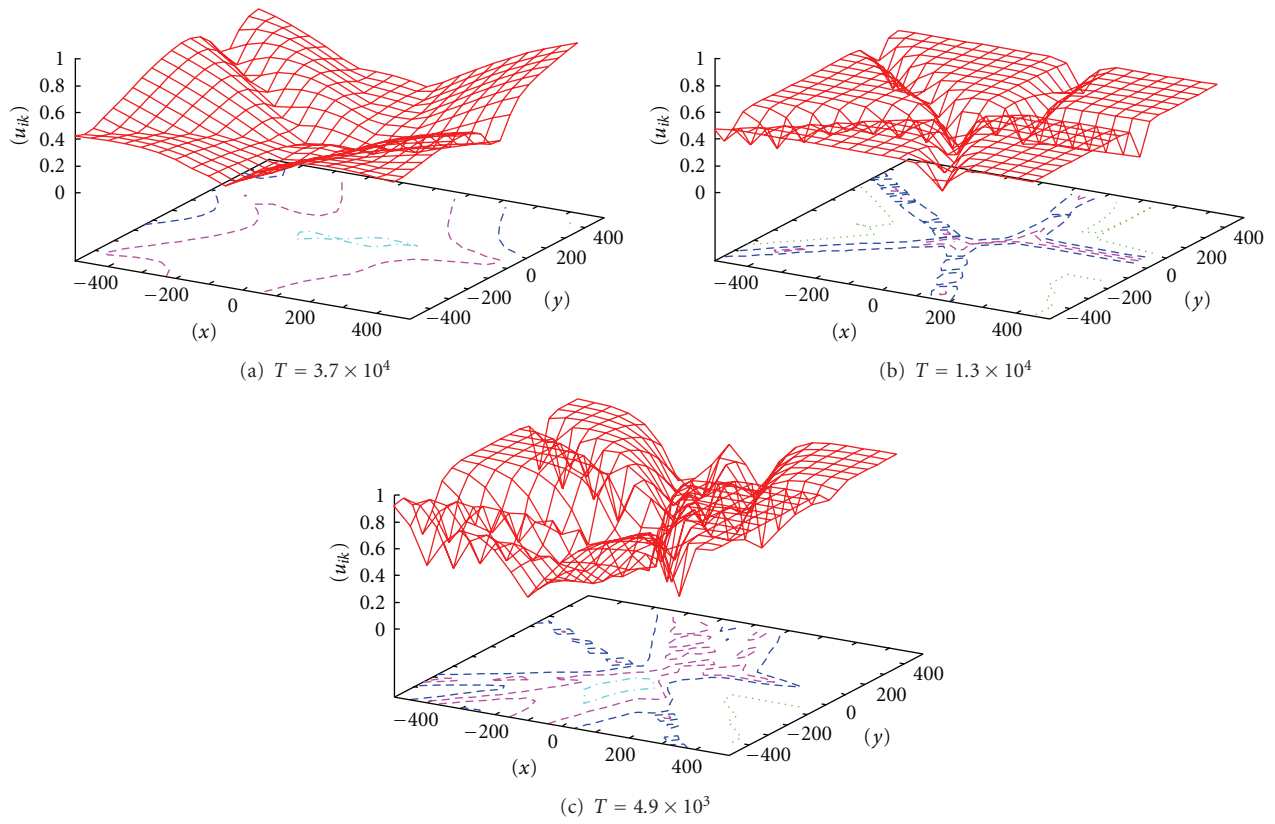


FIGURE 9: The changes of the landscape of u_{ik} with decreasing the temperature.

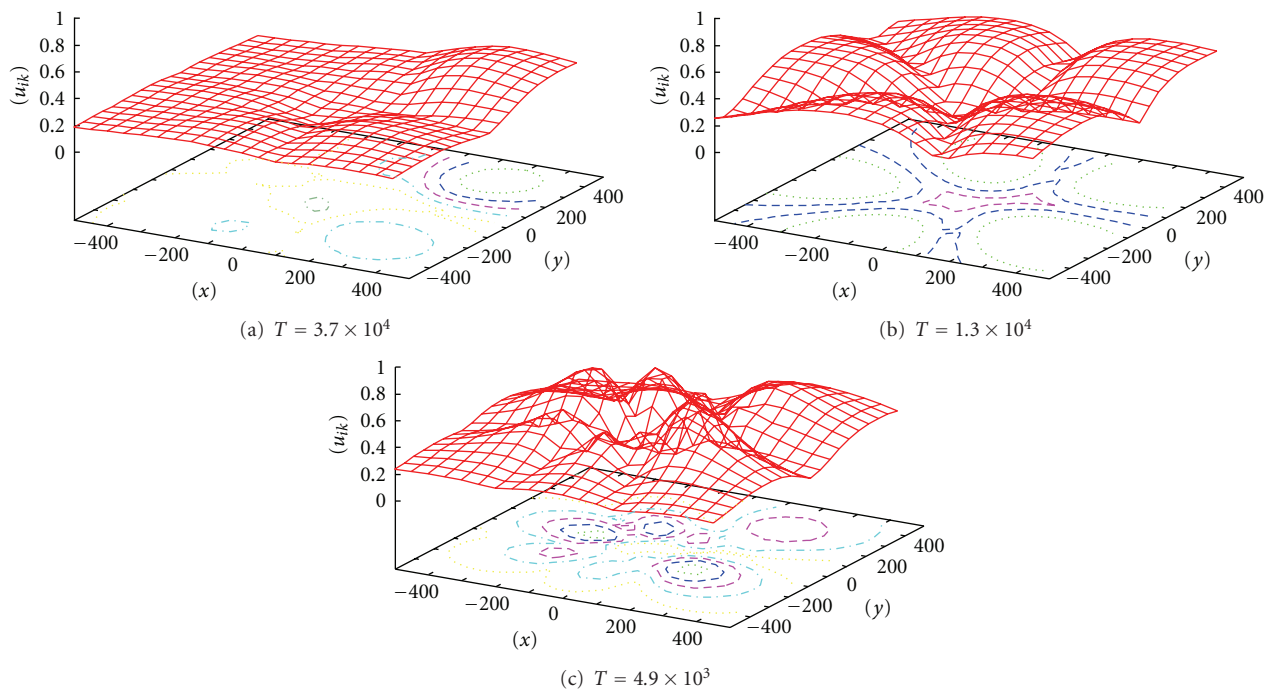


FIGURE 10: The changes of the landscape of \tilde{u}_{ik} with decreasing the temperature ($q = 1.5$).

when the temperature reduction rate $m = 1.0$ or 2.0 , though the best results for the Shannon-entropy-based FCM are obtained only when $m = 2.0$. Furthermore, variances of the Tsallis-entropy-based FCM are smaller than those of the Shannon-entropy-based FCM. These features indicate that a wide range of m values are applicable to Tsallis-entropy-based FCM.

Figure 6 shows the reduction of the objective values of the Tsallis- and Shannon-entropy-based FCM with decreasing the temperature by VFA. The Shannon-entropy-based FCM does not converge properly when $T = 0.023$ for $m = 2.0$ and $T = 0.029$ for $m = 3.0$. That is, with larger m values, the Shannon-entropy-based FCM becomes unstable.

7. Conclusion

By maximizing the Tsallis-entropy, the membership function of the Tsallis-entropy-based FCM is formulated. It has a more gentle base slope in the region far from the origin than that of the Shannon- and fuzzy-entropy-based FCMs. This feature allows clusters to move long distance and the temperature can be reduced rapidly in the Tsallis-entropy-based FCM.

Next, the deterministic annealing (DA) method using very fast annealing (VFA) as its cooling schedule is applied to the Tsallis-entropy-based FCM. VFA initially decreases a temperature extremely, and experimental results showed that the Tsallis-entropy-based FCM was suitable for DA combined with VFA.

Our future works include convergence and computational time test under various conditions (temperatures and parameters, especially q -value) of the Tsallis-entropy-based FCM using very fast deterministic annealing. They also include experiments and examinations of its applications.

Acknowledgment

This research has been supported by the Kayamori Foundation of Informational Science Advancement.

References

- [1] K. Rose, E. Gurewitz, and B. G. Fox, "A deterministic annealing approach to clustering," *Pattern Recognition Letters*, vol. 11, no. 9, pp. 589–594, 1990.
- [2] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [3] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Prentice-Hall, New York, NY, USA, 1981.
- [4] R.-P. Li and M. Mukaidono, "A maximum-entropy approach to fuzzy clustering," in *Proceedings of the 4th IEEE International Conference on Fuzzy Systems, (FUZZIEEE/IFES '95)*, pp. 2227–2232, March 1995.
- [5] S. Miyamoto and M. Mukaidono, "Fuzzy c-means as a regularization and maximum entropy approach," in *Proceedings of the 7th International Fuzzy Systems Association World Congress*, vol. II, pp. 86–92, 1997.
- [6] A. De Luca and S. Termini, "A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory," *Information and Control*, vol. 20, no. 4, pp. 301–312, 1972.
- [7] M. Yasuda, T. Furuhashi, M. Matsuzaki, and S. Okuma, "Fuzzy clustering using deterministic annealing method and its statistical mechanical characteristics," in *Proceedings of the 10th IEEE International Conference on Fuzzy Systems*, pp. 797–800, December 2001.
- [8] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *Journal of Statistical Physics*, vol. 52, no. 1-2, pp. 479–487, 1988.
- [9] M. Ménard, V. Courboulay, and P. -A. Dardignac, "Possibilistic and probabilistic fuzzy clustering: unification within the framework of the non-extensive thermostatics," *Pattern Recognition*, vol. 36, no. 6, pp. 1325–1342, 2003.
- [10] M. Yasuda, "Entropy maximization and deterministic annealing approach to fuzzy c-means clustering," in *Proceedings of Joint 4th International Conference on Soft Computing and Intelligent Systems and 9th International Symposium on Advanced Intelligent Systems (SCIS&ISIS '08)*, 2008.
- [11] M. Yasuda, "Entropy maximization and very fast deterministic annealing approach to fuzzy c-means clustering," in *Proceedings of the Joint 5th International Conference on Soft Computing and Intelligent Systems and 11th International Symposium on Advanced Intelligent Systems*, 2008.
- [12] B. E. Rosen, "Function optimization based on advanced simulated annealing," in *Proceedings of the IEEE Workshop on Physics and Computation, (PhysComp '92)*, pp. 289–293, 1992.
- [13] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, part 2, pp. 179–188, 1936.
- [14] M. Yasuda, T. Furuhashi, and S. Okuma, "Entropy based fuzzy c-means clustering: analogy with statistical mechanics," *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, vol. 17, no. 4, pp. 468–476, 2005.
- [15] L. E. Reichl, *A Modern Course in Statistical Physics*, John Wiley & Sons, New York, NY, USA, 1998.
- [16] H. Szu and R. Hartley, "Fast simulated annealing," *Physics Letters A*, vol. 122, no. 3-4, pp. 157–162, 1987.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

