On Semi-Supervised Fuzzy c-Means Clustering

ENDO Yasunori, HAMASUNA Yukihiro, YAMASHIRO Makito, MIYAMOTO Sadaaki

***Abstract*— We have two methods of pattern classification, one is supervised and the other is unsupervised. Unsupervised classification, which is called clustering and classifies data except external criteria, is very useful in the methods of pattern classification so that it has been applied in many fields.**

There are two types of clustering, one is hierarchical and the other is non-hierarchical. We often use hard c-means clustering (HCM) or fuzzy c-means clustering (FCM) as typical methods of non-hierarchical clustering.

By the way, supervised classification can achieve practical classification results but can't handle a lot of data. On the other hand, unsupervised classification can handle a lot of data but the method is complex and sometimes results look a bit of strange.

Therefore recently, study of semi-supervised classification has been studied. This classification has advantages of both of the above-mentioned methods, e.g., practical results, low costs and short calculation time.

In this paper, we propose new semi-supervised classification algorithms based on fuzzy c-means clustering in which some membership grades are given as supervised membership grade in advance.

I. INTRODUCTION

Recently, methods of pattern classification which classifies data automatically has been becoming more important because of complexity and growing size of the data. The pattern classification attracted interests of a lot of researchers in respects of not only mathematical background but also engineering applications.

We have two methods of pattern classification, one is supervised and the other is unsupervised. In supervised classification, e.g., K-nearest neighbor method, some data are given as supervised data in advance and the classification is done based on the supervised data. Because that supervised data are given, results of supervised methods become more natural than unsupervised methods and the algorithms are simplified. However, the more data increase, the less it is practical to give a lot of supervised data.

In unsupervised classification, which is also called clustering, data is classified into some clusters except external criteria. Because of except the criteria, the algorithms become complex. However, it can be said that how to construct algorithms of unsupervised classification is more significant than supervised one in the field of pattern classification from considering that we now have to handle a lot of data so that it is hard to give supervised data.

There are two types of clustering, one is hierarchical and the other is non-hierarchical. We can sometimes obtain more natural results by hierarchical method than non-hierarchical one, but the hierarchical one has a disadvantage that the calculation time more increases widely as the number of data more increases in comparison with the non-hierarchical one. Moreover, the methodology of not analysis theory but graph theory is necessary to approach the mathematical structure so that its theoretical development is difficult.

In comparison with hierarchical method, non-hierarchical one is more deeply studied and more widely applied in many fields because of small calculation time and interests in the mathematical background. We often use hard c-means clustering (HCM) [2] or fuzzy c-means clustering (FCM) [3] as typical methods of non-hierarchical clustering in many fields as well as engineering and medical science.

By the way, the concept of semi-supervised classification has been remarked [5], [6], [7], [8], [9]. Semi-supervised classification has both characters of supervised and semi-supervised classification, that is, both of labeled data and unlabeled data are classified in the semi-supervised classification and consequently, it becomes possible to achieve more practical results by semi-supervised classification than supervised or unsupervised classification.

In this paper, we present two new semi-supervised fuzzy c-means clustering algorithms. The algorithms are constructed by introducing some membership grades which are given as supervised membership grades in advance into fuzzy c-means clustering. Our proposed algorithms can achieve practical classification results and the algorithms are more simple than the conventional semi-supervised clustering [7], [9]

II. SEMI-SUPERVISED FUZZY c-MEANS CLUSTERING

In this section, we construct new fuzzy c-means clustering algorithms, which have the character of supervised classification, by introducing the supervised membership grade into fuzzy c-means clustering.

In semi-supervised fuzzy c-means clustering, some membership grades of a part of data to some clusters are given in advance. The membership grades are called supervised membership grades and denoted as $\bar{u}_{ki}$.

The supervised membership grade comes from regarding not data on the pattern space but membership grade of data to some cluster as supervisor. Giving $u_{ki} = 1$ can be substituted for setting a data x_k as a supervisor of a cluster G_i, however giving $u_{ki} = 0.6$ can't be substituted for "setting supervisor".

In this section, we propose two new algorithms of semi-supervised fuzzy c-means clustering with $\bar{u}_{ki}$ based on standard and entropy regularized fuzzy c-means clustering.

A. Preliminaries

The basic problem is how to construct the way to classify n numbers of data:

$$X = \{x_k \mid x_k = (x_{k1}, \dots, x_{kp})^T \in \Re^p,\ k = 1 \sim n\}$$

into c numbers of clusters:

$$\mathcal{C} = \{C_i \mid i = 1 \sim c\}$$

which is represented as cluster centers:

$$V = \{v_i \mid v_i = (v_{i1}, \dots, v_{ip})^T \in \Re^p,\ i = 1 \sim c\}$$

on a pattern space $\Re^p$. Euclidean norm is defined on the space, that is,

$$\|x_k - v_i\| = \sqrt{\sum_{j=1}^{p} (x_{kj} - v_{ij})^2}.$$

ENDO Yasunori is with Faculty of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan (phone & fax: +81 29 853 5135; email: endo@risk.tsukuba.ac.jp).

HAMASUNA Yukihiro is with Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan (phone & fax: +81 29 853 5250; email: yhama@soft.risk.tsukuba.ac.jp). He is also a research fellow of the japan society for the promotion of science.

YAMASHIRO Makito is with Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan (phone & fax: +81 29 853 5250; email: yamashiro@soft.risk.tsukuba.ac.jp).

MIYAMOTO Sadaaki is with Faculty of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan (phone: +81 29 853 5346; email: miyamoto@risk.tsukuba.ac.jp).

978-1-4244-3597-5/09/$25.00 ©2009 IEEE

$u_{ki} \in [0,1]$ represents membership grade that x_k belongs to C_i and the goal of this classification is to obtain the following U:

$$\begin{aligned} U =& \{u_{ki} \mid u_{ki} \in [0,1],\ \sum_{i=1}^{c} u_{ki} = 1, \\ & k = 1 \sim n,\ i = 1 \sim c\}. \end{aligned}$$

$\bar{u}_{ki} \in [0,1]$ represents supervised membership grade that x_k belongs to C_i and

$$\bar{U} = \{\bar{u}_{ki} \mid \bar{u}_{ki} \in [0,1],\ k = 1 \sim n,\ i = 1 \sim c\}$$

is given in advance. It is not necessary to determine $\bar{u}_{ki}$ between all x_k and C_i and the value $\bar{u}_{ki}$ is equals to zero in case that $\bar{u}_{ki}$ is not given. Thus, conditions for $\bar{u}_{ki}$ are as follows:

$$\sum_{i=1}^{c} \bar{u}_{ki} \leq 1. \quad (\forall k = 1, \dots, n) \tag{1}$$

B. Semi-Supervised Standard Fuzzy c-Means Clustering

In this paragraph, we try to construct an algorithm of semi-supervised standard fuzzy c-means clustering by introducing supervised membership grade $\bar{u}_{ki}$ into standard fuzzy c-means clustering. To do that, we consider to minimize the following objective function $J(U,V)$ which is derived by introducing $\bar{u}_{ki}$ into the objective function of standard fuzzy c-means clustering:

$$J(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} |u_{ki} - \bar{u}_{ki}|^m \|x_k - v_i\|^2, \tag{2}$$

where the constraints are as follows.

$$\sum_{i=1}^{c} u_{ki} = 1. \quad (\forall k = 1, \dots, n) \tag{3}$$

First, we derive an optimal solution v_i which minimizes $J(U,V)$. J is convex for v_i so that we can obtain the solution from $\frac{\partial J}{\partial v_i} = 0$. Thus,

$$\frac{\partial J}{\partial v_i} = -2 \sum_{k=1}^{n} |u_{ki} - \bar{u}_{ki}|^m (x_k - v_i) = 0.$$

Therefore, we get

$$\sum_{k=1}^{n} |u_{ki} - \bar{u}_{ki}|^m v_i = \sum_{k=1}^{n} |u_{ki} - \bar{u}_{ki}|^m x_k.$$

Consequently, we obtain the following optimal solution v_i:

$$v_i = \frac{\sum_{k=1}^{n} |u_{ki} - \bar{u}_{ki}|^m x_k}{\sum_{k=1}^{n} |u_{ki} - \bar{u}_{ki}|^m}. \tag{4}$$

Next, we derive an optimal solution u_{ki} which minimizes $J(U,V)$. In this case, we have to consider two cases, $m > 1$ and $m = 1$. The reason is that J becomes piece-wise linear and J can't be differentiated on $u_{ki} = \bar{u}_{ki}$ in case of $m = 1$.

1) $m > 1$: Instead of (2), we consider the following objective function in which there is no absolute value:

$$J'(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ki} - \bar{u}_{ki})^m \|x_k - v_i\|^2, \tag{5}$$

where the constraints for u_{ki} and conditions for $\bar{u}_{ki}$ are (3) and (1), respectively.

To derive the optimal solution u_{ki} with the constraints (3), we introduce the following Lagrange function $L(U)$ as follows:

$$\begin{aligned} L(U) =& J'(U,V) + \sum_{k=1}^{n} \lambda_k \left(\sum_{i=1}^{c} u_{ki} - 1 \right) \\ =& \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ki} - \bar{u}_{ki})^m \|x_k - v_i\|^2 \\ & + \sum_{k=1}^{n} \lambda_k \left(\sum_{i=1}^{c} u_{ki} - 1 \right). \end{aligned}$$

J is convex for u_{ki} so that we can obtain the solution from $\frac{\partial L}{\partial v_i} = 0$, that is,

$$\frac{\partial L}{\partial u_{ki}} = m(u_{ki} - \bar{u}_{ki})^{m-1} \|x_k - v_i\|^2 + \lambda_k = 0.$$

Therefore,

$$u_{ki} = \bar{u}_{ki} + \left(-\frac{\lambda_k}{m} \cdot \frac{1}{\|x_k - v_i\|^2} \right)^{\frac{1}{m-1}}.$$

On the other hand, we get the following relation from the constraints (3):

$$\sum_{i=1}^{c} u_{ki} = \sum_{i=1}^{c} \left(\bar{u}_{ki} + \left(-\frac{\lambda_k}{m} \cdot \frac{1}{\|x_k - v_i\|^2} \right)^{\frac{1}{m-1}} \right) = 1,$$

that is,

$$\left(-\frac{\lambda_k}{m} \right)^{\frac{1}{m-1}} = \frac{1 - \sum_{j=1}^{c} \bar{u}_{kj}}{\sum_{j=1}^{c} \|x_k - v_j\|^{\frac{-2}{m-1}}}.$$

Hence, we get the following optimal solution:

$$u_{ki} = \bar{u}_{ki} + \left(1 - \sum_{j=1}^{c} \bar{u}_{kj} \right) \frac{\left(\frac{1}{d_{ki}} \right)^{\frac{1}{m-1}}}{\sum_{j=1}^{c} \left(\frac{1}{d_{kj}} \right)^{\frac{1}{m-1}}}, \tag{6}$$

$$d_{ki} = \|x_k - v_i\|^2. \tag{7}$$

Notice that

$$1 - \sum_{j=1}^{c} \bar{u}_{ki} \geq 0$$

and we know that $u_{ki} \geq \bar{u}_{ki}$. The solution is derived from (5), however it is also the optimal solution of (2) from $u_{ki} - \bar{u}_{ki} \geq 0$.

2) $m = 1$: As the above-mentioned, the objective function (2) can be differentiated on $u_{ki} = \bar{u}_{ki}$ when $m = 1$. If $\sum_{i=1}^{c} \bar{u}_{ki} = 1$, the optimal solution u_{ki} which minimizes $J(U, V)$ would become $u_{ki} = \bar{u}_{ki}$. However the constraints of $\bar{u}_{ki}$ are (1), thus the value of $1 - \sum_{i=1}^{c} \bar{u}_{ki}$ should be given to some u_{ki}. For minimizing $J(U, V)$, it is sufficient that $1 - \sum \bar{u}_{ki}$ is given to u_{ki} where $i = \arg\min_l d_{kl}$ for any k, that is,

$$u_{ki} = \begin{cases} \bar{u}_{ki} + 1 - \displaystyle\sum_{i=1}^{c} \bar{u}_{ki}, & (i = \arg\min_l d_{kl}) \\ \bar{u}_{ki}. & \text{(otherwise)} \end{cases} \tag{8}$$

3) Algorithm: Here, we construct an algorithm of semi-supervised standard fuzzy c-means clustering (sSFCM) using the optimal solutions derived above as follows. The algorithm obtains u_{ki} and v_i which minimize J by iterative optimization like as fuzzy c-means clustering.

Algorithm 1 (sSFCM):

sSFCM1 Give supervised membership grades $\bar{U}$ and set the initial values of V.
sSFCM2 Calculate U on fixing V by (6) ($m > 1$) or (8) ($m = 1$).
sSFCM3 Calculate V on fixing U by (4).
sSFCM4 If the stop criterion satisfies, the algorithm is finished. Otherwise, go back to **sSFCM2**.

C. Semi-Supervised Entropy Regularized Fuzzy c-Means Clustering

In this paragraph, we try to construct an algorithm of semi-supervised entropy regularized fuzzy c-means clustering by introducing supervised membership grade $\bar{u}_{ki}$ into entropy regularized fuzzy c-means clustering. To do that, we consider to minimize the following objective function $J(U, V)$ which is derived by introducing $\bar{u}_{ki}$ into the objective function of entropy regularized fuzzy c-means clustering:

$$\begin{aligned} J(U, V) =& \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ki} \|x_k - v_i\|^2 \\ &+ \lambda^{-1} \sum_{k=1}^{n} \sum_{i=1}^{c} (|u_{ki} - \bar{u}_{ki}| \log |u_{ki} - \bar{u}_{ki}|) \end{aligned} \tag{9}$$

under the constraints (3).

First, we derive an optimal solution v_i which minimizes $J(U, V)$. J' is convex for v_i so that we can obtain the solution from $\frac{\partial J}{\partial u_{ki}} = 0$. Thus,

$$\frac{\partial L}{\partial v_i} = \sum_{k=1}^{n} u_{ki}(2v_i - 2x_k) = 0.$$

Hence, we get the following optimal solution:

$$v_i = \frac{\displaystyle\sum_{k=1}^{n} u_{ki} x_k}{\displaystyle\sum_{k=1}^{n} u_{ki}}. \tag{10}$$

Next, we derive an optimal solution u_{ki} which minimizes $J(U, V)$. Instead of (9), we consider the following objective function in which there is no absolute value:

$$\begin{aligned} J'(U, V) =& \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ki} \|x_k - v_i\|^2 \\ &+ \lambda^{-1} \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ki} - \bar{u}_{ki}) \log(u_{ki} - \bar{u}_{ki}) \end{aligned} \tag{11}$$

To derive the optimal solution u_{ki} with the constraints (3), we introduce the following Lagrange function $L(U)$ as follows:

$$\begin{aligned} L =& J'(U, V) + \sum_{k=1}^{c} \gamma_k \left(\sum_{i=1}^{n} u_{ki} - 1 \right) \\ =& \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ki} \|x_k - v_i\|^2 \\ &+ \lambda^{-1} \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ki} - \bar{u}_{ki}) \log(u_{ki} - \bar{u}_{ki}) \\ &+ \sum_{k=1}^{c} \gamma_k \left(\sum_{i=1}^{n} u_{ki} - 1 \right) \end{aligned} \tag{12}$$

J' is convex for u_{ki} so that we can obtain the solution from $\frac{\partial L}{\partial u_{ki}} = 0$, that is,

$$\frac{\partial L}{\partial u_{ki}} = \|x_k - v_i\|^2 + \lambda^{-1}(\log(u_{ki} - \bar{u}_{ki} + 1)) + \gamma_k = 0.$$

Therefore,

$$u_{ki} = \bar{u}_{ki} + e^{-(\lambda(d_{ki} + \gamma_k) + 1)}. \tag{13}$$

On the other hand, we get the following relation from the constraints (3):

$$\sum_{i=1}^{c} u_{ki} = \sum_{i=1}^{c} \left(\bar{u}_{ki} + e^{-\lambda(d_{ki} + \gamma_k) + 1} \right),$$

that is,

$$e^{-\lambda \gamma_k} = \frac{1 - \displaystyle\sum_{i=1}^{c} \bar{u}_{ki}}{e^{-1} \displaystyle\sum_{i=1}^{c} e^{-\lambda d_{ki}}}.$$

Hence, we get the following optimal solution:

$$u_{ki} = \bar{u}_{ki} + \frac{e^{-\lambda d_{ki}}}{\displaystyle\sum_{j=1}^{c} e^{-\lambda d_{kj}}} \left(1 - \sum_{j=1}^{c} \bar{u}_{kj} \right). \tag{14}$$

Notice that $u_{ki} \geq \bar{u}_{ki}$ and we know u_{ki} which minimizes (11) also minimizes (9), i.e., .the u_{ki} is the optimal solution.

Here, we construct an algorithm of semi-supervised entropy regularized fuzzy c-means clustering (eSFCM) using the optimal solutions derived above as follows. The algorithm obtains u_{ki} and v_i which minimize J by iterative optimization like as the above-mentioned sSFCM.

Algorithm 2 (eSFCM):

eSFCM1 Give supervised membership grades $\bar{U}$ and set the initial values of V.
eSFCM2 Calculate U on fixing V by (14).
eSFCM3 Calculate V on fixing U by (10).
eSFCM4 If the stop criterion satisfies, the algorithm is finished. Otherwise, go back to **eSFCM2**.

III. Numerical Examples

In this section, we show some numerical examples through the following three proposed methods:

- Results by sSFCM ($m > 1$)
- Results by sSFCM ($m = 1$)
- Results by eFCM

Notice that, sSFCM on $m > 1$ and $\bar{u}_{ki} = 0.0$, sSFCM on $m = 1$ and $\bar{u}_{ki} = 0.0$ and eSFCM on $\bar{u}_{ki} = 0.0$ are equal to the conventional sFCM, the conventional HCM and the conventional eFCM, respectively.

A. Data and Supervised Membership Grades

We consider simple data for clarifying characters of our proposed algorithms. The data set consists of two clusters in which there are 10 data each (Table I). Moreover, we consider the following three cases for supervised membership grades.

TABLE I
Data Set $\{x_k \mid x_k \in \Re^2,\ k = 1 \sim 20\}$

k	(x_{k1}, x_{k2})	k	(x_{k1}, x_{k2})
1	$(0.00, 4.50)$	2	$(0.00, 5.50)$
3	$(1.75, 4.50)$	4	$(1.75, 5.50)$
5	$(3.50, 4.50)$	6	$(3.50, 5.50)$
7	$(5.25, 4.50)$	8	$(5.25, 5.50)$
9	$(7.00, 4.50)$	10	$(7.00, 5.50)$
11	$(9.00, 0.00)$	12	$(9.00, 2.50)$
13	$(9.00, 5.00)$	14	$(9.00, 7.50)$
15	$(9.00, 10.00)$	16	$(10.00, 0.00)$
17	$(10.00, 2.50)$	18	$(10.00, 5.00)$
19	$(10.00, 7.50)$	20	$(10.00, 10.00)$

Case 1 $\bar{u}_{ki} = 0.0$ for any k and i.
Case 2 $\bar{u}_{9,1} = \bar{u}_{10,1} = 0.3$. Otherwise $\bar{u}_{ki} = 0.0$.
Case 3 $\bar{u}_{9,1} = \bar{u}_{10,1} = 0.6$. Otherwise $\bar{u}_{ki} = 0.0$.

B. Numerical Results

1) Results by sSFCM ($m = 2$): Here, we show numerical results by sSFCM ($m = 2$). Fig. 1, Fig. 2 and Fig. 3 show the results in Case 1, Case 2 and Case 3, respectively. Table II shows membership grades u_{ki} in each case.

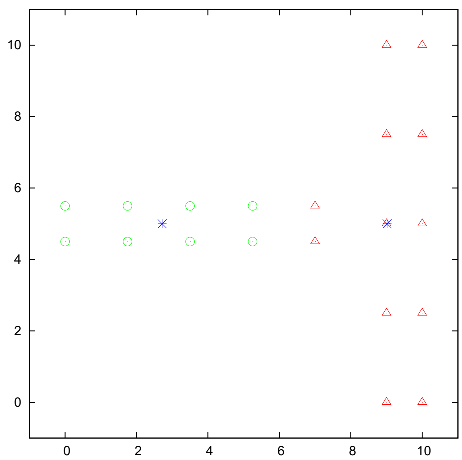

Fig. 1. Numerical Result by sSFCM in Case 1 ($m = 2$), i.e., sFCM

2) Results by sSFCM ($m = 1$): Here, we show numerical results by sSFCM ($m = 1$). Fig. 4, Fig. 5 and Fig. 6 show the results in Case 1, Case 2 and Case 3, respectively. Table III shows membership grades u_{ki} in each case.

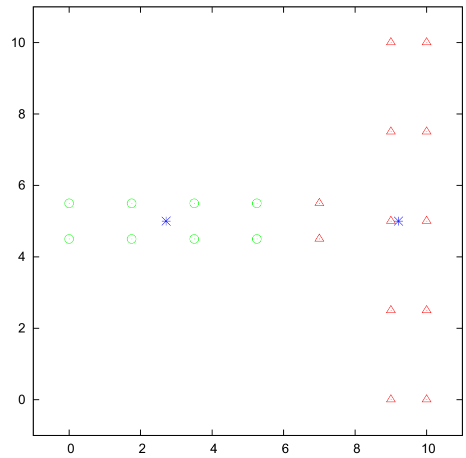

Fig. 2. Numerical Result by sSFCM in Case 2 ($m = 2$)

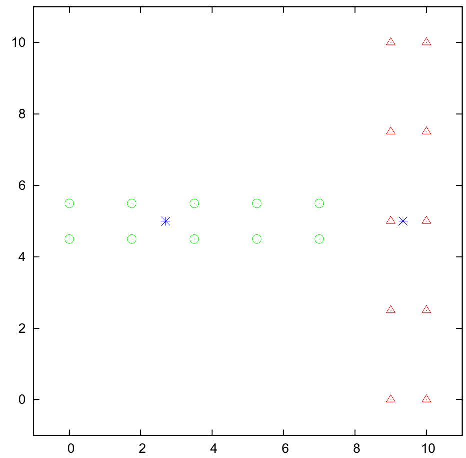

Fig. 3. Numerical Result by sSFCM in Case 3 ($m = 2$)

3) Results by eSFCM: Here, we show numerical results by eSFCM. Fig. 7, Fig. 8 and Fig. 9 show the results in Case 1, Case 2 and Case 3, respectively. Table IV shows membership grades u_{ki} in each case.

C. Consideration

Through the above-mentioned numerical examples, we can show that incorrect classified data belongs to the correct cluster after giving supervised membership grades so that we believe that our proposed is effectiveness.

Here, we consider characters of the algorithms. We can find a similar point between the case that $m = 2$ and $m = 1$ for sSFCM. In case that $m = 1$, final calculated results of membership grade of data given a supervised membership grade doesn't change as well as the other data. Moreover, in case that $m = 2$, final calculated results and initial values for membership grades are different, however the difference is small. We can also say the same for eSFCM.

On the other hand, we can see the difference between sSFCM and eSFCM for the position of cluster centers. The reason is that cluster centers by (4) have inference on supervised membership grade $\bar{u}_{ki}$ while the centers by (10) does not include $\bar{u}_{ki}$.

Last, we have to notice that we can not always get correct results of classification if only we give incorrect

TABLE II
MEMBERSHIP GRADE (u_{k1}, u_{k2}) OF EACH DATA BY SSFCM $(m = 2)$

k	$\bar{u}_{ki} = 0.0$	$\bar{u}_{ki} = 0.3$	$\bar{u}_{ki} = 0.6$
1	$(0.91, 0.09)$	$(0.92, 0.08)$	$(0.92, 0.08)$
2	$(0.91, 0.09)$	$(0.92, 0.08)$	$(0.92, 0.08)$
3	$(0.98, 0.02)$	$(0.98, 0.02)$	$(0.98, 0.02)$
4	$(0.98, 0.02)$	$(0.98, 0.02)$	$(0.98, 0.02)$
5	$(0.97, 0.03)$	$(0.97, 0.03)$	$(0.97, 0.03)$
6	$(0.97, 0.03)$	$(0.97, 0.03)$	$(0.97, 0.03)$
7	$(0.68, 0.32)$	$(0.70, 0.30)$	$(0.72, 0.28)$
8	$(0.68, 0.32)$	$(0.70, 0.30)$	$(0.72, 0.28)$
9	$(0.19, 0.81)$	$(0.45, 0.55)$	$(0.69, 0.31)$
10	$(0.19, 0.81)$	$(0.45, 0.55)$	$(0.69, 0.31)$
11	$(0.28, 0.72)$	$(0.28, 0.72)$	$(0.28, 0.72)$
12	$(0.12, 0.88)$	$(0.12, 0.88)$	$(0.12, 0.88)$
13	$(0.00, 1.00)$	$(0.00, 1.00)$	$(0.00, 1.00)$
14	$(0.12, 0.88)$	$(0.12, 0.88)$	$(0.12, 0.88)$
15	$(0.28, 0.72)$	$(0.28, 0.72)$	$(0.28, 0.72)$
16	$(0.25, 0.75)$	$(0.25, 0.75)$	$(0.25, 0.75)$
17	$(0.11, 0.89)$	$(0.10, 0.90)$	$(0.10, 0.90)$
18	$(0.02, 0.98)$	$(0.01, 0.99)$	$(0.01, 0.99)$
19	$(0.11, 0.89)$	$(0.10, 0.90)$	$(0.10, 0.90)$
20	$(0.25, 0.75)$	$(0.25, 0.75)$	$(0.25, 0.75)$

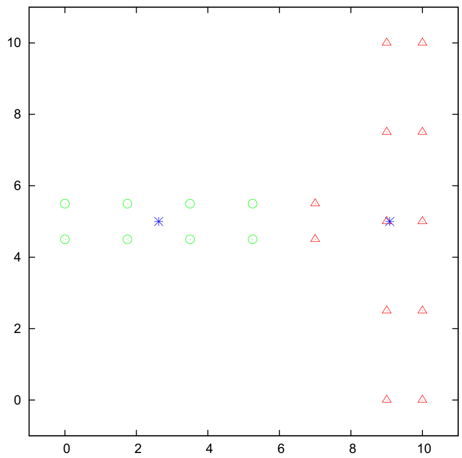

Fig. 4. Result by sSFCM in Case 1 ($m = 1$), i.e., HCM

classified data a supervised membership grade because each membership grade have influence on calculation of values of other membership grades or cluster centers.

IV. CONCLUSION

In this paper, we have presented two new semi-supervised FCM. Our proposed algorithms are constructed by introducing some membership grades which are given as supervised membership grades in advance into FCM. One of the characters of the algorithms is more simple than the conventional ones. Moreover, we have verified the effectiveness through numerical examples. The algorithms can achieve natural classification results.

Our proposed algorithms is expected to handle data more flexibly than the conventional FCM from the viewpoint that we can give any data supervised membership grades.

In the forthcoming paper, we have to verify the algorithms through more real data. Moreover, we also have to estimate those using classification functions. In addition, we expect that our concept that u_{ki} is substituted for $|u_{ki} - \bar{u}_{ki}|$ is very simple so that we can apply the concept to other objective functions.

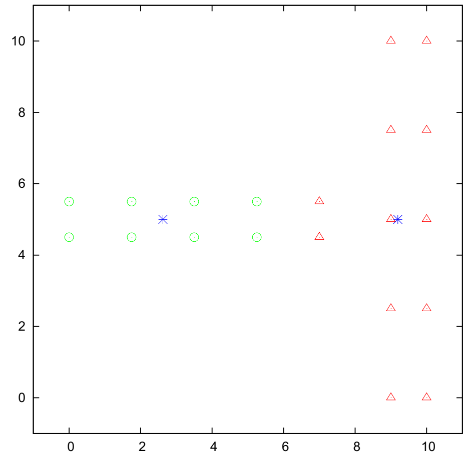

Fig. 5. Result by sSFCM in Case 2 ($m = 1$)

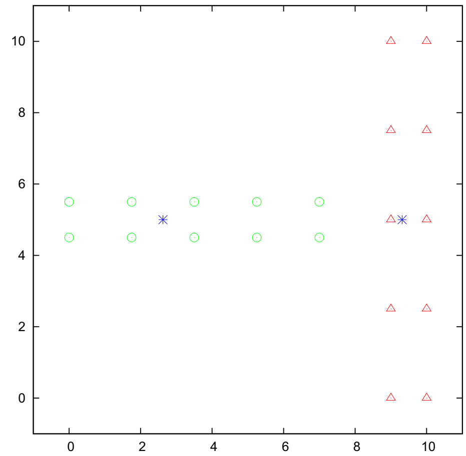

Fig. 6. Result by sSFCM in Case 3 ($m = 1$)

ACKNOWLEDGMENT

This study has partly been supported by the Grant-in-Aid for Scientific Research (C) (No.21500212) and (B) (No.19300074), Japan Society for the Promotion of Science.

REFERENCES

[1] L. A. Zadeh : “Fuzzy sets”, Information and Control, Vol.8, pp.338-353 (1965).

[2] M.R. Anderberg : ‘Cluster Analysis for Applications’, Academic Press, New York (1973).

[3] J.C. Bezdek : ‘Pattern Recognition with Fuzzy Objective Function Algorithms’, Plenum Press, New York (1981).

[4] F. Höppner , F. Klawonn, R. Kruse, T. Runkler : ‘Fuzzy Cluster Analysis’, Wiley, Chichester (1999).

[5] K. Wagstaff, C. Cardie, S. Rogers, S. Schroedl : “Constrained K-means Clustering with Background Knowledge”, In Proceedings of ICML, pp.577–584, (2001).

[6] S. Basu, A. Banerjee, R. Mooney : “Semi-supervised Clustering by Seeding” , In ICML-2002, pp.19–26, (2002).

[7] A. Bouchachia, W. Pedrycz : “A Semi-supervised Clustering Algorithm for Data Exploration”, Proceedings of Fuzzy Sets and Systems, IFSA 2003, LNAI2715, pp.328–337 (2003).

[8] N. Grira, M. Crucianu, N. Boujemaa : “Unsupervised and Semi-supervised Clustering:a Brief Survey”, Report of the MUSCLE European Network of Excellence (2005).

[9] Mitsuaki Yamasaki, Sadaaki Miyamoto, In-Jae Lee : “Semi-Supervised Clustering Using Target Clusters”, 24th Fuzzy System Symposium, pp.133–138 (2008) (in Japanese).

TABLE III
MEMBERSHIP GRADE (u_{k1}, u_{k2}) OF EACH DATA BY SSFCM $(m = 1)$

k	$\bar{u}_{ki} = 0.0$	$\bar{u}_{ki} = 0.3$	$\bar{u}_{ki} = 0.6$
1	$(1.0, 0.0)$	$(1.0, 0.0)$	$(1.0, 0.0)$
2	$(1.0, 0.0)$	$(1.0, 0.0)$	$(1.0, 0.0)$
3	$(1.0, 0.0)$	$(1.0, 0.0)$	$(1.0, 0.0)$
4	$(1.0, 0.0)$	$(1.0, 0.0)$	$(1.0, 0.0)$
5	$(1.0, 0.0)$	$(1.0, 0.0)$	$(1.0, 0.0)$
6	$(1.0, 0.0)$	$(1.0, 0.0)$	$(1.0, 0.0)$
7	$(1.0, 0.0)$	$(1.0, 0.0)$	$(1.0, 0.0)$
8	$(1.0, 0.0)$	$(1.0, 0.0)$	$(1.0, 0.0)$
9	$(0.0, 1.0)$	$(0.3, 0.7)$	$(0.6, 0.4)$
10	$(0.0, 1.0)$	$(0.3, 0.7)$	$(0.6, 0.4)$
11	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
12	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
13	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
14	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
15	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
16	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
17	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
18	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
19	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$
20	$(0.0, 1.0)$	$(0.0, 1.0)$	$(0.0, 1.0)$

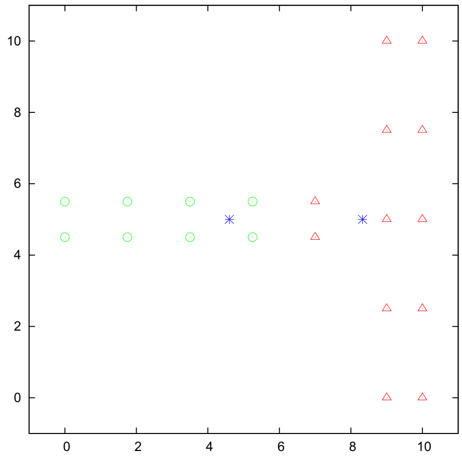

Fig. 7. Result by eSFCM in Case 1 $(\lambda = 1)$

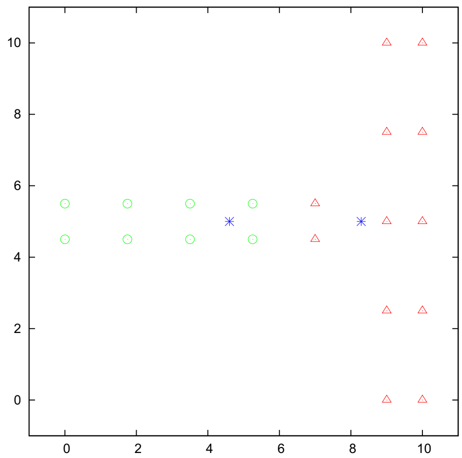

Fig. 8. Result by eSFCM in Case 2 $(\lambda = 1)$

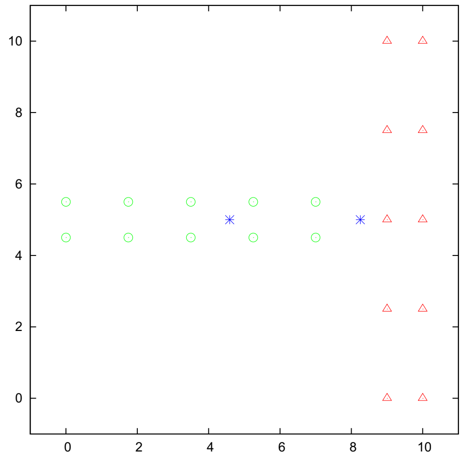

Fig. 9. Result by eSFCM in Case 3 $(\lambda = 1)$

TABLE IV
MEMBERSHIP GRADE (u_{k1}, u_{k2}) OF EACH DATA BY ESFCM $(\lambda = 1)$

k	$\bar{u}_{ki} = 0.0$	$\bar{u}_{ki} = 0.3$	$\bar{u}_{ki} = 0.6$
1	$(0.76, 0.24)$	$(0.76, 0.24)$	$(0.76, 0.24)$
2	$(0.76, 0.24)$	$(0.76, 0.24)$	$(0.76, 0.24)$
3	$(0.84, 0.16)$	$(0.84, 0.16)$	$(0.84, 0.16)$
4	$(0.84, 0.16)$	$(0.84, 0.16)$	$(0.84, 0.16)$
5	$(0.94, 0.06)$	$(0.94, 0.06)$	$(0.94, 0.06)$
6	$(0.94, 0.06)$	$(0.94, 0.06)$	$(0.94, 0.06)$
7	$(0.94, 0.06)$	$(0.93, 0.07)$	$(0.93, 0.07)$
8	$(0.94, 0.06)$	$(0.93, 0.07)$	$(0.93, 0.07)$
9	$(0.25, 0.75)$	$(0.47, 0.53)$	$(0.69, 0.31)$
10	$(0.25, 0.75)$	$(0.47, 0.53)$	$(0.69, 0.31)$
11	$(0.36, 0.64)$	$(0.37, 0.63)$	$(0.37, 0.63)$
12	$(0.21, 0.79)$	$(0.21, 0.79)$	$(0.21, 0.79)$
13	$(0.02, 0.98)$	$(0.03, 0.97)$	$(0.03, 0.97)$
14	$(0.21, 0.79)$	$(0.21, 0.79)$	$(0.21, 0.79)$
15	$(0.36, 0.64)$	$(0.37, 0.63)$	$(0.37, 0.63)$
16	$(0.34, 0.66)$	$(0.34, 0.66)$	$(0.34, 0.66)$
17	$(0.20, 0.80)$	$(0.21, 0.79)$	$(0.21, 0.79)$
18	$(0.09, 0.91)$	$(0.09, 0.91)$	$(0.09, 0.91)$
19	$(0.20, 0.80)$	$(0.21, 0.79)$	$(0.21, 0.79)$
20	$(0.34, 0.66)$	$(0.34, 0.66)$	$(0.34, 0.66)$

TABLE V
CLUSTER CENTER (v_1, v_2)

v_i	v_1	v_2
sSFCM ($m = 2$, Case 1)	$(2.72, 5.00)$	$(9.08, 5.00)$
sSFCM ($m = 2$, Case 2)	$(2.71, 5.00)$	$(9.21, 5.00)$
sSFCM ($m = 2$, Case 3)	$(2.70, 5.00)$	$(9.34, 5.00)$
sSFCM ($m = 1$, Case 1)	$(2.63, 5.00)$	$(9.08, 5.00)$
sSFCM ($m = 1$, Case 2)	$(2.63, 5.00)$	$(9.20, 5.00)$
sSFCM ($m = 1$, Case 3)	$(2.63, 5.00)$	$(9.31, 5.00)$
eSFCM ($\lambda = 1$, Case 1)	$(4.60, 5.00)$	$(8.33, 5.00)$
eSFCM ($\lambda = 1$, Case 2)	$(4.60, 5.00)$	$(8.29, 5.00)$
eSFCM ($\lambda = 1$, Case 3)	$(4.59, 5.00)$	$(8.25, 5.00)$