Clustering Analysis Method based on Fuzzy C-Means Algorithm of PSO and PPSO with Application in Real Data

JENG-MING YIH, YUAN-HORNG LIN and HSIANG-CHUAN LIU

Abstract**—**The popular fuzzy c-means algorithm (FCM) converges to a local minimum of the objective function. Hence, different initializations may lead to different results. The important issue is how to avoid getting a bad local minimum value to improve the cluster accuracy. The particle swarm optimization (PSO) is a popular and robust strategy for optimization problems. But the main difficulty in applying PSO to real-world applications is that PSO usually need a large number of fitness evaluations before a satisfying result can be obtained. In this paper, the improved new algorithm, "Fuzzy C-Mean based on Picard iteration and PSO (PPSO-FCM)", is proposed. Three real data sets were applied to prove that the performance of the PPSO-FCM algorithm is better than the conventional FCM algorithm and the PSO-FCM algorithm.

Keywords**—**FCM, Picard iteration, PSO algorithms, PPSO-FCM algorithm.

I. Introduction

THE popular fuzzy c-means algorithm (FCM) is developed by using Picard Iteration through the first-order conditions for stationary points of the objective function. It converges to a local minimum of the objective function. Hence, different initializations may lead to different results. The important issue is how to avoid getting a bad local minimum value to improve the cluster accuracy.

The particle swarm optimization (PSO) is a popular and robust strategy for optimization problems. But the main difficulty in applying PSO to real-world applications is that PSO usually need a large number of fitness evaluations before a satisfying result can be obtained. In this paper, the improved new algorithm, "Fuzzy C-Mean based on Picard iteration and PSO (PPSO-FCM)", is proposed. Five real data sets were applied to prove that the performance of the PPSO-FCM algorithm is better than the conventional FCM algorithm and PSO-FCM algorithm.

II. Literature Review

The algorithm of fuzzy C-Means Algorithm are the foundations of this study [1]–[3]. The algorithm will be discussed as follows.

A. Fuzzy c-Mean Algorithm [16]

Fuzzy C-Means Algorithm (FCM) is the most popular objective function based fuzzy clustering algorithm,[14] The objective function used in FCM is given by Equation (1)

$$J_{FCM}^{m}\left(U,A,X\right)=\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}d_{ij}^{2}=\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}\left\|\underline{x}_j-\underline{a}_i\right\|^2 \quad (1)$$

is the membership degree of data object in cluster and it satisfies the following constraint given by Equation (2).

$$\sum_{i=1}^{c}\mu_{ij}=1,\ \forall j=1,2,...,n \quad (2)$$

C is the number of clusters, m is the fuzzy index, m>1(in this paper m=2), which controls the fuzziness of the method. They are both parameters and need to be specified before running the algorithm. is the square Euclidean distance between data object to center . Minimizing objective function Equation (1) with constraint Equation (2) is a non-trivial constraint nonlinear optimization problem with continuous parameters and , Hence there is no obvious analytical solution. Alternatively optimizing one set of parameters while the other set of parameters are considered as fixed, is used here. The updating function for and is obtained as Equation (3) and (4).

The steps of the FCM are listed as follows.

Step 1: Determining the number of cluster; c and m-value (let m=2), given converging error, (such as), randomly choose the initial membership matrix, such that the memberships are not all equal;

Step 2: Find

$$\underline{a}_i^{(k)}=\frac{\sum_{j=1}^{n}\left[\mu_{ij}^{(k-1)}\right]^m\underline{x}_j}{\sum_{j=1}^{n}\left[\mu_{ij}^{(k-1)}\right]^m}\quad i=1,2,...,c \quad (3)$$

$$\mu_{ij}^{(k)}=\left[\sum_{l=1}^{c}\left[\frac{\left(\underline{x}_j-\underline{a}_i^{(k)}\right)'\left(\underline{x}_j-\underline{a}_i^{(k)}\right)}{\left(\underline{x}_j-\underline{a}_l^{(k)}\right)'\left(\underline{x}_j-\underline{a}_l^{(k)}\right)}\right]^{\frac{1}{m-1}}\right]^{-1} \quad (4)$$

Step 3: Increment k;

Until $\max_{1\le i\le c}\left\|\underline{a}_i^{(k)}-\underline{a}_i^{(k-1)}\right\|<\varepsilon \quad (5)$

This paper is partially supported by the National Science Council grant (NSC-97 -2511-S -142-001).The authors would like to thank Dr. C. C. Hung for providing the river-image data.

J. M. Yih is with the Graduate Institute of Educational Measurement and Statistics, and Department of Mathematics Education, National Taichung University 140 Min-Sheng Rd., Taichung City 403, Taiwan (corresponding author to provide phone: 86-913-465279; fax: 86-4-22183500; e-mail: yih@mail.ntcu.edu.tw).

Y. H. Lin., was with Department of Mathematics Education National Taichung University. He is now leader of Department of Mathematics Education National Taichung University. (e-mail: lyh@mail.ntcu.edu.tw).

H. C. Liu is with Department of Bioinformatics, Asia University,Taiwan).

B. PSO-Algorithm

Now, The Particle Swarm Optimization(PSO) provide a quite convenient way to solve the problems of optimization.

(1)The position in the solution space can be described as X={x1,x2,x3,…,xp }, where X is an P-dimension vector.
(2)The objective function is given by Eq. (1) and (8) for PSO-FCM and PSO-FCM-M respectively.
(3)We set the two constant number, which we are called a and b. In this paper, we were given a=b=2,[8].
(4)c and d are random variables in the range from 0 to 1.
(5) Bi(t) is the best solution of the i-th particle at t-th iteration.
(6)G(t) is the best solution of all particles at t-th iteration.
(7)The velocity update of PSO is showed as following: vi(t+1)=vi(t)+a*c*(Bi(t)- xi(t))+ b*d*(G(t)- xi(t)))
(8)The position update of PSO is showed as following:xi(t+1)= xi(t)+ Vi(t).

C. FCM-M Algorithm [13],[15]

For improving the above two problems, our previous work [4] proposed the improved algorithm FCM-M which added $-\ln\left|+\Sigma_i^{-1}\right|$ a regulating factor of covariance matrix to each class in objective function, and deleted $\left|M_i\right| = \rho_i$ the constraint of the determinant of covariance matrices in GK Algorithm as the objective function (6).

Using the Lagrange multiplier method, We can minimize the objective function (6). Constraint (7) with respect to the parameters $\underline{a}_i$, μ_{ij}, and Σ_i, we can obtain the solutions as (10), (11), and(13).

We want to avoid the singular problem and to select the better initial membership matrix, the updating functions for $\underline{a}_i$, μ_{ij} and Σ_i are obtained as (8) ~ (3-8). Both of FCM and FCM-M can not exploit all of the memberships with the same value. FCM is a special case of FCM-M, when covariance matrices equal to identity matrices by our previous work [8].

$$J^m_{FCM-M}(U,A,\Sigma,X)=\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^m\left[\left(\underline{x}_j-\underline{a}_i\right)'\Sigma_i^{-1}\left(\underline{x}_j-\underline{a}_i\right)-\ln\left|\Sigma_i^{-1}\right|\right] \tag{6}$$

Constraints: membership,

$$\sum_{i=1}^{c}\mu_{ij}=1,\ \forall j=1,2,\ldots,n \tag{7}$$

$\Sigma=\{\Sigma_1,\Sigma_2,\ldots,\Sigma_c\}$ is the set of covariance of cluster.

Step 1: Determining the number of cluster; c and m-value (let m=2), given converging error, $\varepsilon>0$ (such as $\varepsilon=0.001$).

Method 1: choose the result membership matrix of FCM algorithm as the initial one.

Method 2: let $\underline{a}_i^{(0)}, i=1,2,\ldots,c$ be the result centers of k-mean algorithm, And $d_{ij}=\left\|\underline{x}_j-\underline{a}_i^{(0)}\right\|$ be distances between data object $\underline{x}_j$ to center $\underline{a}_i^{(0)}$.

$$d_i^M=\max_{1\le j\le n}d_{ij}=\max_{1\le j\le n}\left\|\underline{x}_j-\underline{a}_i^{(0)}\right\|,$$
$$d_i^m=\min_{1\le j\le n}d_{ij}=\min_{1\le j\le n}\left\|\underline{x}_j-\underline{a}_i^{(0)}\right\|, \tag{8}$$
$$u_{ij}^{(0)}=\frac{d_i^M-d_{ij}}{d_i^M-d_i^m},\ i=1,2,\ldots,c,\ j=1,2,\ldots,n$$

$$U^{(0)}=\begin{bmatrix}u_{11}^{(0)}\,u_{12}^{(0)}\ldots u_{1n}^{(0)}\\ u_{21}^{(0)}\,u_{22}^{(0)}\ldots u_{2n}^{(0)}\\ \ldots\ \ldots\ \ldots\\ u_{c1}^{(0)}\,u_{c2}^{(0)}\ldots u_{cn}^{(0)}\end{bmatrix}=\begin{bmatrix}u_1^{(0)}(\underline{x}_1)\,u_1^{(0)}(\underline{x}_2)\ldots u_1^{(0)}(\underline{x}_n)\\ u_2^{(0)}(\underline{x}_1)\,u_2^{(0)}(\underline{x}_2)\ldots u_2^{(0)}(\underline{x}_n)\\ \ldots\ \ldots\ \ldots\\ u_c^{(0)}(\underline{x}_1)\,u_c^{(0)}(\underline{x}_2)\ldots u_c^{(0)}(\underline{x}_n)\end{bmatrix} \tag{9}$$

Step 2: Find

$$\Sigma_i^{(k)}=\frac{\sum_{j=1}^{n}\left[\mu_{ij}^{(k-1)}\right]^m\left(\underline{x}_j-\underline{a}_i^{(k)}\right)\left(\underline{x}_j-\underline{a}_i^{(k)}\right)'}{\sum_{j=1}^{n}\left[\mu_{ij}^{(k-1)}\right]^m} \tag{10}$$

$$\Sigma_i^{(k)}=\sum_{s=1}^{p}\lambda_{si}^{(k)}\Gamma_{si}^{(k)}\left(\Gamma_{si}^{(k)}\right)',$$

$$\left[\lambda_{si}^{(-1)}\right]^{(k)}=\begin{cases}\left[\lambda_{si}^{(k)}\right]^{-1} & if\ \lambda_{si}^{(k)}>0\\ 0 & if\ \lambda_{si}^{(k)}=0\end{cases}$$

$$\left[\Sigma_i^{-1}\right]^{(k)}=\sum_{s=1}^{p}\left[\lambda_{si}^{(-1)}\right]^{(k)}\Gamma_{si}^{(k)}\left(\Gamma_{si}^{(k)}\right)' \tag{11}$$

$$\left|\Sigma_i^{-1}\right|^{(k)}=\prod_{1\le s\le p,\lambda_{si}^{(k)}>0}\left[\lambda_{si}^{(-1)}\right]^{(k)} \tag{12}$$

$$\mu_{ij}^{(k)}=\left[\sum_{s=1}^{c}\left[\frac{\mathrm{w}_i'\left[\Sigma_i^{-1}\right]^{(k)}\mathrm{w}_i-\ln\left\|\left[\Sigma_i^{-1}\right]^{(k)}\right\|}{\mathrm{w}_s'\left[\Sigma_s^{-1}\right]^{(k)}\mathrm{w}_s-\ln\left\|\left[\Sigma_s^{-1}\right]^{(k)}\right\|}\right]^{\frac{1}{m-1}}\right]^{-1} \tag{13}$$

$$where\ \mathrm{w}_i=\left(\underline{x}_j-\underline{a}_i^{(k)}\right)$$

Step 3: Increment k; until $\max_{1\le i\le c}\left\|\underline{a}_i^{(k)}-\underline{a}_i^{(k-1)}\right\|<\varepsilon$.

D. PSO-FCM Algorithm

Particle Swarm Optimization (PSO) [7] is a quite convenient method for optimizing hard numerical function on metaphor of social behavior of flocks of birds and schools of fish[5].

A swarm consist M individuals, (here,), called particles, which change their position over time. Each particle represents a potential solution to the problem of optimization [6]. In FCM, the problem of optimization is to minimize the value of the objective function. Let the particle k in a D-dimension space (D = nc) be represented as

(i) Let the particle k in a D-dimension space (D = nc) be represented as

$$\begin{aligned}\mu_k &= \left(\overrightarrow{\mu_{k1}}, \overrightarrow{\mu_{k2}}, ..., \overrightarrow{\mu_{kD}}\right) \\ &= (\mu_{k11}, \mu_{k12}, ..., \mu_{k1n}, \\ &\quad \mu_{k21}, \mu_{k22}, ..., \mu_{k2n}, ..., \mu_{kc1}, \mu_{kc2}, ..., \mu_{kcn}) \\ k &= 1, 2, ..., M\end{aligned} \quad (14)$$

(ii) Let the objective function of FCM be the fitness function as follows,

$$\begin{aligned}J_{FCM}^{m}(U, A, X) &= \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m} d_{ij}^{2} \\ &= \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}\left\|\underline{x}_j - \underline{a}_i\right\|^2\end{aligned} \quad (15)$$

where $\underline{a}_i = (\sum_{j=1}^{n}\left[\mu_{ij}\right]^m)^{-1}(\sum_{j=1}^{n}\left[\mu_{ij}\right]^m \underline{x}_j), i = 1, 2, ..., c$ (16)

(iii) Let the particle k in a D-dimension space (D = nc) be represented as

$$\begin{aligned}\mu_k &= \left(\overrightarrow{\mu_{k1}}, \overrightarrow{\mu_{k2}}, ..., \overrightarrow{\mu_{kD}}\right) \\ &= (\mu_{k11}, \mu_{k12}, ..., \mu_{k1n}, \mu_{k21}, \\ &\quad \mu_{k22}, ..., \mu_{k2n}, ..., \mu_{kc1}, \mu_{kc2}, ..., \mu_{kcn}) \\ k &= 1, 2, ..., M\end{aligned} \quad (17)$$

(iv) Let the objective function of FCM be the fitness function as follows,

$$\begin{aligned}J_{FCM}^{m}(U, A, X) &= \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m} d_{ij}^{2} \\ &= \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}\left\|\underline{x}_j - \underline{a}_i\right\|^2\end{aligned} \quad (18)$$

where $\underline{a}_i = (\sum_{j=1}^{n}\left[\mu_{ij}\right]^m)^{-1}(\sum_{j=1}^{n}\left[\mu_{ij}\right]^m \underline{x}_j), i = 1, 2, ..., c$

(19)

(v) The best previous position which possesses the best fitness value of particle k was denoted by $p_k = \left(p_{k1}, p_{k2}, ..., p_{kD)}\right)$, which is also called P_{best}. The index of the best P_{best} among all the particles is denoted by the symbol g. We called that the best fitness value of the position $p_k = \left(p_{k1}, p_{k2}, ..., p_{kD)}\right)$ was also g_{best} .The velocity for the particle k is represented as $v_k = \left(v_{k1}, v_{k2}, ..., v_{kD}\right)$.

(vi) p_{best} and g_{best} location for iteration t according to following two formulas,

$$\begin{aligned}v_k(t+1) &= wv_k(t) + c_1r_1\left(p_{best}(t) - \mu_k(t)\right) \\ &\quad + c_2r_2\left(g_{best}(t) - \mu_k(t)\right)\end{aligned}$$

$\mu_k(t+1) = normalized\left[\mu_k(t) + v_k(t+1)\right]$ satisfying;

$$\mu_{kij}(t+1) = \frac{\mu'_{kij}(t+1)}{\sum_{i=1}^{c}\mu'_{kij}(t+1)}, \quad (20)$$

$$\forall i = 1,2,...,c,\ j = 1,2,...,n,\ k = 1,2,...,m$$

where $\mu'_{kij}(t+1) = \dfrac{U_{kij}(t) - \min_{1\le j\le n} U_{kij}(t)}{\max_{1\le j\le n} U_{kij}(t) - \min_{1\le j\le n} U_{kij}(}$

$$U_{kij}(t) = \mu_{kij}(t) + v_{kij}(t+1) \quad (21)$$

$$k = 1, 2, ..., m,\ i = 1, 2, ..., c,\ j = 1, 2, ..., n$$

Where w is the inertia coefficient which is a constant in the interval [0,1], and can be adjusted in the direction of linear decrease, (In this paper w=0.75); c_1 and c_2 are learning rates which are nonnegative constants(In this paper, $c_1 = c_2 = 2$); r_1 an r_2 are generated randomly in the interval [0,1].

(vii) The termination criterion for iterations is determined according to whether the maximum generation or a designated value of the fitness is reached. In this paper, the given converging error is $\varepsilon = 0.001$.

$$\max_{1\le i\le c}\left\|\underline{a}_i\ (t+1) - \underline{a}_i\ (t)\right\| < \varepsilon = 0.001$$

where $\underline{a}_i\ (t+1) = \dfrac{\sum_{j=1}^{n}\left[\mu_{ij}(t)\right]^m \underline{x}_j}{\sum_{j=1}^{n}\left[\mu_{ij}(t)\right]^m}, i = 1, 2, ..., c$ (22)

E. PPSO-FCM Algorithm[11]

The main difficulty in applying PSO to real-world applications is that PSO usually need a large number of fitness evaluations before a satisfying result can be obtained [8]. For overcoming above problem, In this paper, the improved new algorithm, "Fuzzy C-Mean based on Picard iteration [4] and PSO (PPSO-FCM)", is proposed. This algorithm integrates Picard iteration and PSO for FCM.

(i) Randomly choose the first particle in a D-dimension space (D = nc) as follows, such that the memberships are not all equa

$$\mu_1 = (\mu_{111}, \mu_{112}, ..., \mu_{11n}, \mu_{121}, \mu_{122}, ..., \mu_{12n}, ..., \mu_{1c1}, \mu_{1c2}, ..., \mu_{1cn})$$

$$\sum_{i=1}^{c}\mu_{1ij} = 1, \quad j = 1, 2, ..., n \quad (15)$$

(ii) Let the objective function of FCM be the fitness function as follows,

$$J_{FCM}^{m}\left(U,A,X\right)=\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}d_{ij}^{2} = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}\left\|\underline{x}_j-\underline{a}_i\right\|^{2} \quad (16)$$

$$\text{where } \underline{a}_i=\frac{\sum_{j=1}^{n}\left[\mu_{ij}\right]^{m}\underline{x}_j}{\sum_{j=1}^{n}\left[\mu_{ij}\right]^{m}},\ i=1,2,\dots,c \quad (17)$$

(iii) Let the first particle, μ_1 , be the initial value, by using Picard iteration in FCM algorithm we can obtain the local minimum solution, called p_1 .

We can randomly choose the second particle, μ_2 , as follows $\mu_2=normalized\left[p_1+2.5r\left(\mu_1-p_1\right)\right]$ satisfying

$$\sum_{i=1}^{c}\mu_{2ij}=1,\quad 0\le\mu_{2ij}\le 1,\quad i=1,2,\dots,c,\ j=1,2,\dots,n \quad (18)$$

Let the second particle, μ_2 , be the initial value, by using the FCM algorithm we can obtain the local minimum solution, called p_2 . Let the better of p_1 and p_2 be called p_g

We can randomly choose the second particle, μ_3, as follows

$$\mu_3=normalized\left[\mu_3'\right]$$
$$\mu_3'=p_2+2r_1\left(\mu_2-p_2\right)+2r_2\left(\mu_2-p_g\right) \quad (19)$$

(18) is satisfied

$$\sum_{i=1}^{c}\mu_{3ij}=1,\quad 0\le\mu_{3ij}\le 1,\quad i=1,2,\dots,c,\ j=1,2,\dots,n \quad (20)$$

Where r_1 and r_2 are generated randomly in the interval [0,1];Similarly we can randomly choose the particle, μ_m , as follows

$$\mu_3=normalized\left[\mu_3'\right]$$
$$\mu_3'=p_2+2r_1\left(\mu_2-p_2\right)+2r_2\left(\mu_2-p_g\right) \quad (21)$$

satisfying

$$\sum_{i=1}^{c}\mu_{3ij}=1,\quad 0\le\mu_{3ij}\le 1,\quad i=1,2,\dots,c,\ j=1,2,\dots,n \quad (22)$$

Where p_g is the best of $p_1,p_2,\dots,p_{m-1}$, r_1 an r_2 are generated randomly in the interval [0,1];

(iv) The best previous position (which possesses the best fitness value) of particle k is denoted by $p_k=\left(p_{k1},p_{k2},\dots,p_{kD)}\right)$, which is also called p_{best} ,. The index of the best p_{best} among all the particles is denoted by the symbol g. The location $p_g=\left(p_{g1},p_{g2},\dots,p_{gD)}\right)$ is also called g_{best} .

The velocity for the particle k is represented as $v_k=\left(v_{k1},v_{k2},\dots,v_{kD}\right)$.

(v) p_{best} and g_{best} location for iteration t according to following two formulas,

$$v_k(t+1)=wv_k(t)+c_1r_1\left(p_{best}(t)-\mu_{kd}(t)\right)+c_2r_2\left(g_{best}(t)-\mu_k(t)\right)$$
$$\mu_k\left(t+1\right)=normalized\left[\mu_k\left(t\right)+v_k\left(t+1\right)\right]$$

which is satisfied

$$\mu_{kij}(t+1)=\frac{\mu_{kij}'(t+1)}{\sum_{i=1}^{c}\mu_{kij}'(t+1)},\quad \forall i=1,2,\dots,c,\ j=1,2,\dots,n,\ k=1,2,\dots,m \quad (23)$$

where

$$\mu_{kij}'(t+1)=\frac{U_{kij}(t)-\min_{1\le j\le n}U_{kij}(t)}{\max_{1\le j\le n}U_{kij}(t)-\min_{1\le j\le n}U_{kij}(t)},\quad U_{kij}(t)=\mu_{kij}(t)+v_{kij}(t+1),\quad k=1,2,\dots,m,\ i=1,2,\dots,c,\ j=1,2,\dots,n \quad (24)$$

Where w is the inertia coefficient which is a constant in the interval [0,1], and can be adjusted in the direction of linear decrease, (In this paper w=0.75); c_1 and c_2 are learning rates which are nonnegative constants(In this paper, $c_1=c_2=2$); r_1 an r_2 are generated randomly in the interval [0,1];

(vi) The termination criterion for iterations is determined according to whether the maximum generation or a designated value of the fitness is reached. In this paper, the given converging error is $\varepsilon=0.001$.

$$\max_{1\le i\le c}\left\|\underline{a}_i\ (t+1)-\underline{a}_i\ (t)\right\|<\varepsilon=0.001$$

$$\text{where } \underline{a}_i\ (t+1)=\frac{\sum_{j=1}^{n}\left[\mu_{ij}(t)\right]^{m}\underline{x}_j}{\sum_{j=1}^{n}\left[\mu_{ij}(t)\right]^{m}},\quad i=1,2,\dots,c \quad (25)$$

III. Data Resource

We have Three real data sets were applied to prove that the performance of the PPSO-FCM algorithm is better than the conventional FCM algorithm and the PSO-FCM algorithm. One real data is Iris Data. Another is Image Data.

A. *Experiment of Iris Data*

The performances of three clustering methods, FCM, PSO-FCM, and PPSO-FCM, are compared in the experiments. The Iris data with sample size 150 is used as first example. The features of the Iris data contain Length of Sepal, Width of Sepal, Length of Petal, and Width of Petal. The samples were assigned the original 3 clusters based on the clustering analysis. The results were shown in Table 1.

Table 1 The characteristics of 3 clusters.

Cluster	Sample size	Concepts	Average distance of center
1	50	Setosa	0.481705
2	50	Versicolor	0.706870
3	50	Virginica	0.819339

Each cluster has 50 sample points. The classification accuracies were shown in Table 2.

Table 2 Classification accuracies of testing samples.

Algorithm	Accuracies (%)
FCM	89.33
PSO-FCM	90.00
PPSO-FCM	90.67

From Table 2, we find that the PPSO-FCM has the best result, up to 90.67%.

B. *Experiment of Iris Data with Random*

The Iris Data [12] with sample size 150 are used as second example. The items of the Iris Data contain Length of Sepal, Width of Sepal, Length of Petal, and Width of Petal. The samples were assigned the original 3 clusters based on the clustering analysis. The results were shown in Table 3.

Table 3 The characteristics of 3 clusters.

Cluster	Samples size	Concepts	Average distance of center
1	50	Setosa	0.48170523644
2	50	Versicolor	0.70687020404
3	50	Virginica	0.81933940766

Each 10 sample points were randomly drawn from Cluster 2 and Cluster 3, respectively, and 3 from Cluster 1. The 23 sample points of raw data of Iris were listed in Appendix.The classification accuracies of testing samples were shown in Table 4.

Table 4 Classification accuracies of testing samples.

Algorithm	Accuracies (%)
FCM	21
PSO-FCM	48
PSO-FCM-M	61

From the data of Table 4, we found that the PSO-FCM-M Algorithm could obtain the best result, up to 61%. The singular problem and detecting the local extreme value problem are improved by the Eigenvalues method and the algorithm of Particle Swarm Optimization. Finally, two numerical examples showed that the new fuzzy clustering algorithm (PSO-FCM-M) gave more accurate clustering results than that of FCM algorithm.

C. *Experiment of Image Data*

A real data set of Image with 262144 sample points from the original Image was selected[9]. All of the sample points were assigned to 3 clusters for clustering analysis. The results were shown in Table5-6 and Figure 1-3.

Table5 The Characteristics of 3 Clusters

Cluster	Samples size	Colors	Average distance of center
1	188118	Red	83.3725
2	3382	Green	104.7881
3	70644	Blue	126.5261

The classification accuracies of testing samples were shown in Table 6.

Table6 Classification Accuracies of Testing samples

Algorithm	Accuracies (%)
FCM	30.65
PSO-FCM	54.67
PPSO-FCM	75.73

From the data of Table 4, we found that the FPSO-FCM could obtain the best result, up to 75.73%.

the improved new algorithm, "Fuzzy C-Mean based on Picard iteration and PSO (PPSO-FCM)", is proposed. Above three real data sets were applied to prove that the performance of the PPSO-FCM algorithm is better than the conventional FCM algorithm and PSO-FCM algorithm.

As shown in Fig. 1, We use the Algorithm of FCM, we could obtain the result, 30.65%.

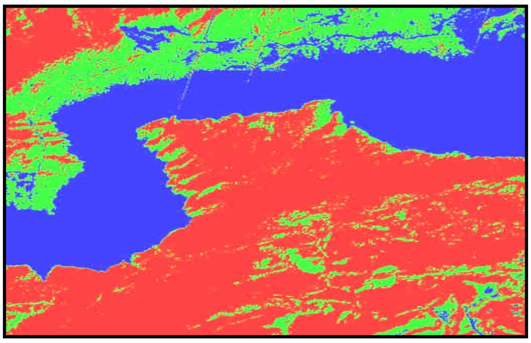

Fig. 1. FCM, obtain the result, 30.65%

As shown in Fig. 2, We use the Algorithm of PSO-FCM, we could obtain the result, 54.67% ◦

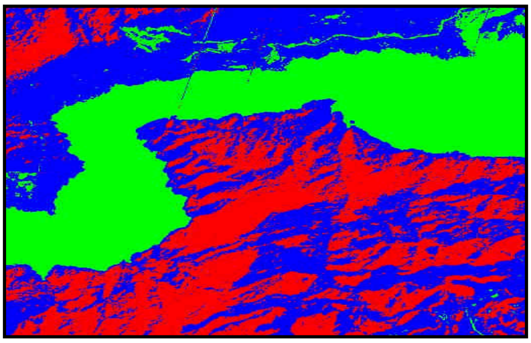

Fig 2. PSO-FCM, obtain the result, 54.67%

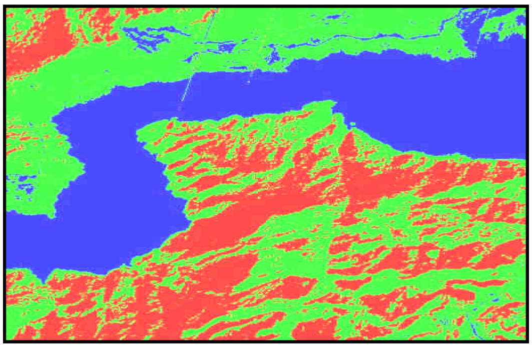

Fig. 3. PPSO-FCM, obtain the result, 75.73%

As shown in Fig.1-3, We use the Algorithm of PPSO-FCM, we could obtain the best result, Two real data sets were applied to prove that the performance of the PPSO-FCM algorithm is better than the conventional FCM algorithm and the PSO-FCM algorithm.

D. Experiment of Mathematics Teaching Data

A real data set of students with sample size 493 from elementary schools was selected. The main factors of the data were calculated by using factor analysis. According to the main factors, the samples were assigned to 4 clusters based on the clustering analysis. The results were shown in Table 5.

Table 5 the characteristics of Math Teaching data.

Cluster	Samples size	concepts	Average distance of center
1	115	division	1.2576760
2	128	ordering	1.2968550
3	168	multiplication	1.1244569
4	82	place value	1.7861002

Each 15 sample points were randomly drawn from Cluster 1, cluster 2, and cluster 3, respectively, and 5 from cluster 4.

Table 6 Accuracies for Math teaching data

Algorithm	Accuracies
FCM	0.5800
PSO-FCM	0.7800
PPSO-FCM	0.8000

From the data of Table 6, to compare the values of the accuracies and Rand index of above three algorithms, we can find that PPSO-FCM algorithm has the best performance, and PSO-FCM algorithm is better than the traditional FCM algorithm.

E. Experiment of wine data

A wine data set was downloaded from website, ftp://ftp.ics.uci.edu/pub/machine-learning-databases.

The sample included 178 instances, 3 classes of wine, and 13 features for each instance. The performances of three clustering methods, FCM, PSO-FCM, and PPSO-FCM, are compared in the experiments. The Wine data with sample size 178 is used as first example [10]. The features of the wine data contain Alcohol, Malic Acid, Ash, Alcalinity of ash, Magnesium, Total phenols, Flavanoids, Nonflavanoid phenols, Proanthocyanins, Color intensity, Hue, OD280/OD315 of diluted wines and Proline. The samples were assigned the original 3 clusters based on the clustering analysis. The results were shown in Table 7.

Table 7 The characteristics of 3 clusters

cluster	sample size	type of wines	average distance of center
1	50	first	427.9094
2	50	Second	282.5046
3	50	Third	253.0249

The classification accuracies were shown in Table 8.

Table 8 Classification accuracies of testing samples

Algorithms	Accuracies (%)
FCM	49.44
PSO-FCM	57.87
PPSO-FCM	68.54

From Table 8, we find that the PPSO-FCM algorithm has the best result, up to 68.54%.

IV. CONCLUSION

A conclusion section is not required. Although a conclusion An improved new fuzzy clustering algorithm is developed to obtain better quality of fuzzy clustering result. The objective function includes the regulating terms about the covariance matrices. The update equations for the memberships and the cluster centers and the covariance matrices are directly derived from the Lagrange's method. The fuzzy c-mean algorithm is different from the GK and GG algorithms. The singular problem and detecting the local extreme value problem are improved by the Eigenvalues method and the algorithm of Particle Swarm Optimization. Finally, two numerical examples showed that the new fuzzy clustering algorithm (PSO-FCM-M) gave more accurate clustering results than that of FCM algorithm. The popular fuzzy c-means algorithm (FCM) is developed by using Picard Iteration through the first-order conditions for stationary points of the objective function. It converges to a local minimum of the objective function. Hence, different initializations may lead to different results. The important issue is how to avoid getting a bad local minimum value to improve the cluster accuracy.

The particle swarm optimization (PSO) is a popular and robust strategy for optimization problems. But the main difficulty in applying PSO to real-world applications is that PSO usually need a large number of fitness evaluations before a satisfying result can be obtained. In this paper, the improved new algorithm, "Fuzzy C-Mean based on Picard iteration and PSO (PPSO-FCM)", is proposed. Five real data sets were applied to prove that the performance of the PPSO-FCM algorithm is better than the conventional FCM algorithm and PSO-FCM algorithm.

A new fuzzy clustering algorithm is developed to obtain better quality of fuzzy clustering result. The algorithm of Particle Swarm Optimization. Finally, Four numerical examples showed that the new fuzzy clustering algorithm (PPSO-FCM) gave more accurate clustering results than that of FCM algorithm. The improved new algorithm, "Fuzzy C-Mean based on Picard iteration and PSO (PPSO-FCM)", is proposed. Three real data sets were applied to prove that the performance of the PPSO-FCM algorithm is better than the conventional FCM algorithm and the PSO-FCM algorithm.

APPENDIX

We Show **T**he 23 sample points of raw data of Iris 、The Confusion Matrix of FCM 、PSO- FCM and PPSO-FCM and the original Images as follows.

Table 9 raw data of Iris

Cluster	Length of Sepal	Width of Sepal	Length of Petal	Width of Petal
1	5.00	3.40	1.50	0.20
1	4.30	3.00	1.10	0.10
1	5.70	4.40	1.50	0.40
2	7.00	3.20	4.70	1.40
2	6.90	3.10	4.90	1.50
2	6.30	3.30	4.70	1.60
2	6.30	2.50	4.90	1.50
2	6.10	2.80	4.70	1.20
2	6.80	2.80	4.80	1.40
2	6.70	3.00	5.00	1.70
2	6.70	3.10	4.70	1.50
2	5.90	3.20	4.80	1.80
2	6.00	2.70	5.10	1.60
3	4.90	2.50	4.50	1.70
3	6.00	2.20	5.00	1.50
3	6.90	3.20	5.70	2.30
3	5.60	2.80	4.90	2.00
3	6.70	3.30	5.70	2.10
3	7.20	3.20	6.00	1.80
3	6.20	2.80	4.80	1.80
3	6.10	3.00	4.90	1.80
3	6.00	3.00	4.80	1.80
3	6.30	2.80	5.10	1.50

Table 10.The Confusion Matrix of FCM Algorithm for Iris Data .

	New Cluster A'	New Cluster B'	New Cluster C'
Initial Cluster A	47	3	0
Initial Cluster B	13	37	0
Initial Cluster C	0	0	50

Table 11.The Confusion Matrix of PSO-FCM Algorithm for Iris Data.

	New Cluster A'	New Cluster B'	New Cluster C'
Initial Cluster A	47	3	0
Initial Cluster B	12	38	0
Initial Cluster C	0	0	50

Table12.The Confusion Matrix of PPSO-FCM Algorithm for Iris Data .

	New Cluster A’	New Cluster B’	New Cluster C’
Initial Cluster A	47	3	0
Initial Cluster B	11	39	0
Initial Cluster C	0	0	50

Table13. Confusion Matrix of FCM Algorithm for Math Teaching Data .

	New ClusterA’	New Cluster B’	New Cluster C’	New Cluster D’
Initial Cluster A	6	1	0	8
Initial Cluster B	0	15	0	0
Initial Cluster C	1	9	5	0
Initial Cluster D	0	2	0	3

Table14 Confusion Matrix of PSO-FCM Algorithm for Math Teaching Data

	New Cluster A’	New Cluster B’	New Cluster C’	New Cluster D’
Initial Cluster A	8	0	0	7
Initial Cluster B	0	15	0	0
Initial Cluster C	0	0	12	3
Initial Cluster D	0	2	0	3

Table15 Confusion Matrix of PPSO-FCM Algorithm for Math Teaching Data

	New Cluster A’	New Cluster B’	New Cluster C’	New Cluster D’
Initial Cluster A	9	2	0	4
Initial Cluster B	0	15	0	0
Initial Cluster C	0	0	12	3
Initial Cluster D	0	0	1	4

Fig. 4. The original river-image

Fig. 5. The river-image classified by 3 color

ACKNOWLEDGMENT

This paper is partially supported by the National Science Council grant (NSC-97 -2511-S -142-001).The authors would like to thank Dr. C. C. Hung for providing the river-image data.

REFERENCES

[1] J. C. Dunn, A Ffuzzy Relative of the ISODATA Process and Its Use in Detecting Compact, Well-separated Clusters, Journal of Cybernetics, Vol. 3, No. 3, 1973, pp. 32-57.

[2] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, New York, Plenum Press, 1981.

[3] H. C. Liu, J. M. Yih, D. B. Wu, and S. W. Liu, Fuzzy C-means Algorithm based on Adaptive Mahalanobis Distances, Information Sciences 2007 Proceedings of the 10th Joint Conferences, Salt Lake City, Utah, 2007, pp. 1398-1404.

[4] Y. Yao, A. N. Muhammed, and H. Zhou, Some Remarks on the Convergence of Picard Iteration to a Fixed Point for a Continuous Mapping, Applied Mathematics E-notes, Vol. 7, 2007.

[5] J. Kennedy, and R. C. Eberhart, Particle Swarm Optimization, Proc. of the IEEE International Conference on Neural Networks, Vol. IV, 1995, pp.1942-1948.

[6] R. C. Eberhart, and J. Kennedy, A New Optimizer Using Particle Swarm Theory, Proc. of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 1995, pp.39-43.

[7] R. C. Eberhart, and Y. Shi, Evolving Artificial Neural Networks, Proc. of the 1998 International Conference on Neural Networks and Brain, Beijing, China, 1998, pp.5-13.
[8] J. F Chang, S. C. Chu, J. F. Roddick and J. S. Pan, A Parallel Particle Swarm Optimization Algorithm with Communication Strategy. Journal of Information, Vol. 31, 2006, pp. 261-279.
[9] C. C Hung, and G. Germany, K-means and Iterative Selection Algorithms in Image Segmentation, Proceedings of IEEE Southeast Conference, Session 1: Software Development, Jamaica, West Indies, 2003.
[10] C. C Hung, and G. Germany, "K-means and Iterative Selection Algorithms in Image Segmentation," in proceedings of IEEE Southeast Conference, Session 1: Software Development, Jamaica, West Indies, April 4-6, 2003.
[11] Jeng-Ming Yih, Yuan-Horng Lin, Hsiang-Chuan Liu (2008). Clustering Analysis Method Besed on Fuzzy C-Means Algorithm of PSO and PPSO with Application in Image Data. Proceedings of the 8th WSEAS International Conference on APPLIED COMPUTER SCIENCE (ACS'08), pp. 54-59, ISSN: 1790-5109/ ISBN: 978-960-474-028-4. [Venice, Italy, November 21-23, 2008].
[12] R. A. Fisher. Annals of Eugenics. 7, 179 (1936).
[13] Jeng-Ming Yih, Yuan-Horng Lin, Hsiang-Chuan Liu (2008). FCM Algorithm Besed on Unsupervised Mahalanobis Distances with Better Initial Values and Separable Criterion. Proceedings of the 8th WSEAS International Conference on APPLIED COMPUTER SCIENCE (ACS'08), pp. 326-331, ISSN: 1790-5109/ ISBN: 978-960-474-028-4. [Venice, Italy, November 21-23, 2008].
[14] Seyed Amir Hadi Minoofam, Azam Bastanfard, (2008) A Novel Algorithm for Generating Muhammad Pattern Based on Cellular Automata Pro ceedings of the 13th WSEAS International Conference on APPLIED MATHEMATICS (MATH'08),pp339-344, ISSN: 1790-2769, ISBN: 978-960-474-034-5.
[15] 4. Hsiang-Chuan Liu, Jeng-Ming Yih, Der-Bang Wu, Chin-Chun Chen (2007). Fuzzy C-Mean Algorithm Based on Mahalanobis Distance and New Separable Criterion. Hong Kong, China,, August 21, 2007.
[16] Hsiang-Chuan Liu, Jeng-Ming Yih, Shin-Wu Liu (2007). Fuzzy C-Mean Algorithm Based On Mahalanobis Distances and Better Initial Values. Information Sciences 2007 Proceedings of the 10th Joint Conference& The 12th International Conference on Fuzzy Theory & Technology. (FTT 2007), pp. 1398-1404. EI-paper [Salt Lake City, Utah, U. S. A., July 18-24, 2007].
[17] Hsiang-Chuan Liu, Jeng-Ming Yih, Tian-Wei Sheu, Shin-Wu Liu (2007). A new Fuzzy Possibility Clustering Algorithms Based on Unsupervised Mahalanobis Distances. *Proceedings of International Conference on Machine Learning and Cybernetics 2007 (ICMLC 2007), Vol.7. No.7*, pp. 3939-3944, ISBN 1-4244-0973X. **EI-paper [**Hong Kong, China, August 19-22, 2007 **]**.
[18] Hsiang-Chuan Liu, Der-Bang wu, Jeng-Ming Yih, Shin-Wu Liu (2008). Fuzzy Possibility C-Mean Based on Complete Mahalanobis Distance and Separable Criterion. *Eighth Intelligent Systems Design and Applications (ISDA 2008), pp. 89-94,* ISBN:978-0-7695-3382-7. [Kaohsiung, Taiwan, November 26-28, 2008].
[19] Hsiang-Chuan Liu, Jeng-Ming Yih, Der-Bang Wu, Shin-Wu Liu (2008). Fuzzy C-Mean Algorithm Based on "Completez" Mahalanobis Distances. *Proceedings of 2008 International Conference on Machine Learning and Cybernetics.* pp. 3569-3574, ISBN: 978-1-4244-2095-7. **EI-paper** [Grand Park Hotel, Kunming, China, July 12-15, 2008].
[20] Hsiang-Chuan Liu, Der-Bang Wu, Jeng-Ming Yih, Liu, S. W. (2007). Fuzzy Possibility C-Mean Based on Mahalanobis Distance and Separable Criterion. *Proceedings of The 7th WSEAS International Conference on APPLIED COMPUTER SCIENCE (ACS'07)*, pp. 570-518, ISSN: 1790-5117/ ISBN: 978-960-6766-18-3. **ISI-paper** [Venice, Italy, November 21-23, 2007].
[21] Hsiang-Chuan Liu, Jeng-Ming Yih, Wen-Chih Lin, Tung-Sheng Liu (2008). Fuzzy C-Means Algorithm Based on PSO and Mahalanobis Distances. *2008 International Symposium on Intelligent Information (ISII2008),* **EI-paper** [Tokai University, Kumamoto, Japan, December 12-13, 2008].
[22] Hsiang-Chuan Liu, Jeng-Ming Yih, Der-Bang wu, Shin-Wu Liu (2008). Fuzzy C-Mean Clustering Algorithms Based on Picard Iteration and Particle Swarm Optimization. *Second International Symposium on Intelligentv Information Technology Application.* pp., ISBN: 978-0-7695-3497-8.**EI-paper** [Shanghai,China, December 21-22, 2008].
[23] Jeng-Ming Yih, Yuan-Horng Lin, Hung, W. L.(2007). Fuzzy Approach Method for Concept Structure Analysis Based on FLMP and ISM with Application in Cognition Diagnosis of Linear Algebra. *Information Sciences 2007 Proceedings of The 10th Joint Conference & The 12th International Conference on Fuzzy Theory & Technology (FTT 2007)*, pp. 1356-1362, **EI-paper [** Salt Lake City, Utah, U. S. A., July 18-24, 2007].
[24] Jeng-Ming Yih, Yuan-Horng Lin (2007). A Fuzzy Basis on Knowledge Structure Analysis for Cognition Diagnosis and Application on Geometry Concepts for Pupils. *The Third International Conference on Intelligence Information Hiding and Multimedia Signal Processing. (IIH-MSP 2007) Vol.2*, pp. 183-186, ISBN: 0-7695-2994-1. **EI-paper** [Kaohsiung City, Taiwan, November 26-28, 2007].
[25] Jeng-Ming Yih, Yuan-Horng Lin (2007). An Integration of Fuzzy Theory and ISM for Concept Structure Analysis with Application of Learning MATLAB. *The Third International Conference on Intelligence Information Hiding and Multimedia Signal Processing. (IIH-MSP 2007) Vol.2*, pp. 187-190, ISBN: 0-7695-2994-1. **EI-paper** [Kaohsiung City, Taiwan, November 26-28, 2007].
[26] Yuan-Horng Lin, Jeng-Ming Yih (2008). Fuzzy Logic Approach on Cognition Diagnosis with Application on Number Concept for Pupils. *Proceedings of 2008 International Conference on Machine Learning and Cybernetics.* pp. 3575-3580, ISBN: 978-1-4244-2095-7. **EI-paper** [Grand Park Hotel, Kunming, China, July 12-15, 2008]
[27] Yuan-Horng Lin, Jeng-Ming Yih, Min-Yen Chen (2008). Polytomous IRS with Application in Concepts Diagnosis and Clustering on Fraction Subtraction for Pupils. *Second International Symposium on Intelligentv Information Technology Application.* pp. 463-467, ISBN: 978-0-7695-3497-8.**EI-paper**[Shanghai, China, December 21-22, 2008]
[28] Yuan-Horng Lin, Jeng-Ming Yih (2008). An Integration of Concept Structure Analysis and S-P Chart with Application in Equality Axiom Concepts Diagnosis. *Second International Symposium on Intelligentv Information Technology Application.* pp. 468-472, ISBN: 978-0-7695-3497-8.**EI-paper**[Shanghai, China, December

21-22, 2008].

J. M. Yih is with the Graduate Institute of Educational Measurement and Statistics, and Department of Mathematics Education, National Taichung University 140 Min-Sheng Rd., Taichung City 403, Taiwan and the Dean of General Affairs of National Taichung University

In the future, We will consider Fuzzy Possibility C-Mean and new Separable Criterion([17]-[20]),select better initial values([16],[21]-[22] and Fuzzy Approach Method for Concept Structure Analysis[23]-[28]