http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 6 (4): 541-546, 2007
ISSN 1812-5638

Image Segmentation Using The Enhanced Possibilistic Clustering Method

[1]Zhenping Xie, [1,2]Shitong Wang, [3]DianYou Zhang, [2]F.L. Chung and [2]Hanbin
[1]School of Information, Southern Yangtze University, Wuxi, 214122, China
[2]Department of computing, Hong Kong Polytechnic University, Hong Kong China
[3]Chinese Ship Industry Corporation 723 Research Institute, YangZhou, 225001, China

Abstract: The possibility based clustering method PCM (possibilistic clustering method) was first proposed by Krishnapuram and Keller to overcome FCM for noises and outliers. However, it still has the following weaknesses: 1) the clustering results are dependent on parameter selection and initialization; 2) the outliers cannot be labeled in a reasonable way. In this study, in order to avoid the above weaknesses, a novel modified PCM version, called EPCM (Enhanced PCM), is presented. First, a novel strategy of Flexible Hyperspheric Partition (FHP) is proposed and then, this strategy is used to construct the objective function of EPCM with some novel constraints. The main advantage of EPCM is that it can label the outliers adaptively and accurately, which enhances the clustering performance and increases its potential applications. Our experimental results about artificial datasets and image segmentation confirm the above standpoints.

Key words: Enhanced possibilistic clustering method, outliers, flexible hyperspheric partition, image segmentation

INTRODUCTION

As an important data processing technique, clustering has been widely utilized in a variety of fields (Yu *et al*., 2005; Cheng *et al*., 2002; Cheng and Chau, 2001, 2002) and it is a method for clustering a data set into most similar groups in the same cluster and most dissimilar groups in different clusters. Generally, clustering algorithms may be classified as the following main categories: hierarchical clustering, partition-based clustering, density-based clustering and grid-based clustering (Yu, 2005; Dave and Krishnapuram, 1997, Dave and Sen, 2002). Especially, the partition-based clustering algorithms are most widely and deeply studied. One of the most widely used partition-based clustering algorithms is FCM (Fuzzy C-Means), which assigns fuzzy memberships of data points (objects) to the clusters. In FCM, the fuzzy membership only represents the relative degree of belonging of data point to respective cluster. So the fuzzy memberships of FCM cannot always represent the proper degrees of belonging of data, especially in noise environment (Krishnapuram and Keller, 1993). To overcome this weakness (Krishnapuram and Keller, 1993, 1996) proposed a new clustering algorithm named PCM. Compared with FCM, PCM can effectively eliminate the influence of noise and outliers to clustering. However, the price PCM pays for its freedom to ignore noise points is that PCM is very sensitive to initializations and often results in the coincident cluster problem (Zhang and Yeung, 2004; Barni *et al*., 1996). Moreover, the typicalities, i.e., possibilistic memberships are very sensitive to the choices of the additional parameters needed by the PCM algorithm, which directly determine the clustering accuracy.

Nowadays, several improved PCM algorithms have been proposed to overcome the weakness of the original PCM algorithm (Trimm *et al*., 2001, 2004; Timm and Kruse, 2002; Gustafson and Kessel, 1979; Pal *et al*., 2005; Zhang and Yeung, 2004; yang and Wu, 2006). In Trim *et al*. (2001 and 2004), Timm and Kruse (2002) proposed two possibilistic fuzzy clustering algorithms that can avoid the coincident cluster problem of PCM by adding an inverse function of the distances between cluster centers in PCM's objective function, which acts as a repulsive force and keeps the clusters separate with each others (avoids coincident clusters). In Trim *et al*., (2001 and 2004),Trim and Kruse, 2002, used the same concept to modify the objective function as used in Gustafson and Kessel's clustering algorithm (Grustafon and Kessel, 1979). In Pal *et al*. (2005), integrate the objective functions of both PCM and FCM into a new objective function and present an improved version, called PFCM (possibilistic fuzzy c-means), which can be interpreted as PCM and FCM, respectively in some special cases, where some proper parameters are adopted; PFCM can inherit the merits of both clustering algorithms. In Zhang and Yeung (2004), introduced fuzzy membership of FCM into PCM's objective function and presented an

Corresponding Author: Shitong Wang, School of Information, Southern Yangtze University, Wuxi, 214122, China

improved PCM clustering algorithm to overcome the coincident problem of PCM. In Yang and Wu (2006) presented an unsupervised possibilistic clustering algorithm PCA and proved it to be more robust than PCM. Although these improved PCM clustering algorithms can partially overcome the drawbacks of PCM, they often need to pay more attentions to adjust some parameters and are not easy for real applications.

In this study, we propose a novel improved PCM clustering algorithm called EPCM (Enhanced PCM). EPCM has the following two main features:

- It introduces the strategy of flexible hyperspheric partition to avoid estimating the parameter η_i in PCM and its variants, which has an important influence on the clustering results.
- It imposes a novel constraint on the objective function. In the proposed objective functions in (Pal *et al*., 2005; Zhang and Yeung, 2004), there are two partition matrices that play the roles of both fuzzy and possibilistic partitions, respectively, where the aim of introducing fuzzy partition matrix is mainly to avoid the coincident problem. However, in our method, only a partition matrix is needed which can play the same roles of both fuzzy and possibilistic partitions simultaneously.

In summary, the proposed EPCM not only inherits the merits of PCM, but also weakens the coincident problem and parameter sensitivity problem of PCM. Especially, in EPCM, the outliers could be labeled accurately. So it is more suitable for real applications. Our experimental results demonstrate the above advantages of EPCM.

FCM AND PCM ALGORITHM

The most widely used fuzzy clustering algorithm is FCM (Fuzzy C-Means). Its objective function can be described as

$$J_{FCM}(U,V;X)=\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}^{m}d^{2}(x_j,v_i)\quad \text{s.t.}\begin{cases}\sum_{i=1}^{C}u_{ij}=1,\forall j\\ u_{ij}\geq 0,\forall i,j\end{cases}\qquad(1)$$

where the total number of sample points is assumed to be N, $U=[u_{ij}]_{C\times N}$ denotes the fuzzy partition matrix, u_{ij} denotes the fuzzy membership, $V=[v_1\ \ v_2\Lambda v_c]$ denotes C cluster centers, $X=\{x_1, x_2, \Lambda, x_N\}$ denotes the dataset, $d(x_j,\ v_i)$ denotes the distance measure, e.g., the most commonly used Euclidean distance and m is weighted exponent. Optimal partitions U* of X are taken from pairs (U*, V*) that are local minimum of J_{FCM}.

To overcome outliers sensitivity of FCM, Krishnapuram and Keller (1993 and 1996) proposed a new clustering algorithm called PCM (Possibilistic C-Means). The objective function of PCM can be described as

$$J_{PCM}(U,V;X)=\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}^{m}d^{2}(x_j,v_i)+\sum_{i=1}^{C}\sum_{j=1}^{N}(1-u_{ij})^{m}\eta_i\qquad(2)$$
$$\text{s.t. } u_{ij}\geq 0,\ \forall i,j$$

where $U=[u_{ij}]_{C\times N}$ denotes possibilistic partition matrix, U_{ij} denotes the possibilistic membership, $\eta_i(i = 1, 2, \Lambda, C)$ is a scale parameter and it is suggested in [Krishnapuram and Keller, 1993) to be:

$$\eta_i=K\frac{\sum_{j=1}^{N}u_{ij}^{m}d^{2}(x_j,v_i)}{\sum_{j=1}^{N}u_{ij}^{m}}\qquad(3)$$

where K>0 and in general K = 1.

PCM cancels the column sum constraint of the membership matrix in FCM and has the capability in weakening the influence of noise and outliers in clustering. However, as pointed out by Barni *et al*. (1996), the price PCM pays for its freedom to ignore noise points is that PCM is very sensitive to initializations. Moreover, the typicalities, i.e., possibilistic memberships, can be very sensitive to the choices of the additional parameters η_i. Furthermore, the outliers cannot be labeled reasonably and adaptively in PCM and its variants, which would weaken the worth of real applications. So these aspects of PCM deserve further studying. In this study, we attempt to improve these aspects of PCM and its variants. In the next section, the strategy of flexible hyperspheric partition will be introduced firstly.

THE STRATEGY OF FLEXIBLE HYPERSPHERIC PARTITION FHP

It is obviously that the objective function of PCM can be divided into C independent objective functions from the discussion and for cluster i the corresponding single objective function could be described as:

$$J_{PCM\_i}=\sum_{j=1}^{N}u_{ij}^{m}d^{2}(x_j,v_i)+\sum_{j=1}^{N}(1-u_{ij})^{m}\eta_i\qquad(4)$$

Thus, the optimal membership values of all data points to cluster i can be obtained by minimizing the formula (4) in PCM clustering algorithm. And to the above objective function, the smaller η_i will results in smaller function value and smaller u_{ij}, which inversely result in much smaller η_i if it is updated using (3). So the continuous update for parameters η_i is forbidden in PCM, thus the optimal membership value is strongly dependent on the η_i. Finally, we can draw the conclusion that PCM is strongly dependent on parameters selection, that is to say, η_i should be evaluated properly in the initiation phase. Nevertheless, in many cases, this problem will not happen, because that the clustering results of FCM will be used as initial condition in general and the FCM can get the reasonable clustering results in average sense in most clustering problems. However, for some dataset, the above problem should be considered and settled. On the other hand, it can be easily discovered that the task to label the outliers adaptively is impossible in PCM. To remedy these two problems, the strategy of flexible hyperspheric partition FHP for single cluster cluster i is proposed, the objective function can be stated as:

$$J_{FHP\_i} = \sum_{j=1}^{N} u_{ij}^m (d_{ij}^2 + (1-\alpha_i)\cdot St(d_{ij}^2 - \eta_{Ii})) + \sum_{j=1}^{N} (1-u_{ij})^m (\eta_{Ii} + \alpha_i \cdot St(\eta_{Oi} - d_{ij}^2)) \quad (5)$$

where α_i is a free parameter with range (0, 1), d_{ij} is equal to $d(x_j, v_i)$ and the function St(g) is defined as:

$$St(x) = \begin{cases} x, & x \ge 0 \\ 0, & x < 0 \end{cases}$$, η_{Ii} and η_{oi}

are defined in (6).

$$\eta_{Ii} = \frac{\sum_{j=1}^{N} u_{ij}^m d_{ij}^2}{\sum_{j=1}^{N} u_{ij}^m}, \quad \eta_{Oi} = \frac{\sum_{j=1}^{N} (1-u_{ij})^m d_{ij}^2}{\sum_{j=1}^{N} (1-u_{ij})^m} \quad (6)$$

Now, η_{Ii} and η_{oi} might be updated dynamically in FHP, where η_{Ii} plays the same role as η_i in PCM. Likewise PCM, the alternative updating also may be used to get the optimal u_{ij} for $\forall j$ and v_i in FHP, the detailed description is omitted here. Furthermore, the data points whose memberships to cluster i are smaller than 0.5 may be regarded as the point set that all points shouldn't belong to the cluster i, which is the basis used to label the outliers in EPCM. In the next section, EPCM will be introduced.

ENHANCED POSSIBILISTIC CLUSTERING METHOD

In this section, we present an enhanced possibilistic clustering method (EPCM) by extending the strategy of FHP. EPCM may be viewed as a generalized FHP based clustering algorithm. The objective function of EPCM is written as follows:

$$J_{EPCM} = \sum_{i=1}^{C} \Big(\sum_{j=1}^{N} u_{ij}^m (d_{ij}^2 + (1-\alpha_i)\cdot St(d_{ij}^2 - \eta_{Ii})) + \sum_{j=1}^{N} (1-u_{ij})^m (\eta_{Ii} + \alpha_i \cdot St(\eta_{Oi} - d_{ij}^2)) \Big) \text{ s.t. } \begin{cases} u_{ij} \ge 0 & \forall i,j \\ \sum_{i=1}^{C} u_{ij} \le 1 & \forall j \end{cases} \quad (7)$$

In (7), a new constraint $\sum_{i=1}^{C} u_{ij} \le 1$ is introduced to avoid the coincident problem of PCM. All the current PCM and its variants share a fuzzy partition matrix to overcome the coincident problem and a possibilistic matrix to weaken the influence of outliers and noises. However, by imposing a novel constraint $\sum_{i=1}^{C} u_{ij} \le 1$, our algorithm can effectively solve these two problems simultaneously only with a partition matrix, because this new constraint not only makes the different clusters to exclude with each other but can also weaken the influence of noises and outliers.

Now, let us discuss the clustering implement of EPCM. First, with theorem 1 is introduced as follows:

Theorem 1: Given V, η_{Ii}, η_{oi} and if m = 2 is considered (which is suitable for almost all applications), the optimal u_{ij} by minimizing (7) can be obtained by the following procedure of determining the optimal U_I.

Proof: The proof of this theorem is omitted due to the limitation of the paper space.

Procedure of determining the optimal u_{ij}:

1. Set
   $A_{ij} = d_{ij}^2 + (1-\alpha_i)\cdot St(d_{ij}^2 - \eta_{Ii})$, $B_{ij} = \eta_{Ii} + \alpha_i \cdot St(\eta_{Oi} - d_{ij}^2)$
   $p_{\cdot j} = \sum_{i=1}^{C} \frac{B_{ij}}{A_{ij} + B_{ij}}$
2. If $p_{gj} \le 1$, then $u_{ij} = \frac{B_{ij}}{A_{ij} + B_{ij}}$ and terminate, otherwise, go to 3;
3. Let $L_j = 2(\sum_{i=1}^{C} \frac{1}{A_{ij} + B_{ij}})^{-1}(p_{\cdot j} - 1)$;
4. If $L_j \le 2 \times \min_i \{B_{ij}\}$, then go to 5, otherwise, go to 6;
5. Calculate $u_{ij} = \frac{B_{ij} - L_j/2}{A_{ij} + B_{ij}}$, then terminate this procedure;
6. Calculate the following variables that satisfy the following conditions simultaneously using some greedy algorithm,
   $L_j = 2 \cdot (\sum_{i=1}^{C} \frac{1}{A_{ij} + B_{ij}} - \sum_{i \in (\prod j)} \frac{1}{A_{ij} + B_{ij}})^{-1} \cdot (p_{\cdot j} - \sum_{i \in (\prod j)} \frac{B_{ij}}{A_{ij} + B_{ij}} - 1)$
   $\prod j = \{k \mid u_{kj} = 0\}$
   $u_{ij} = St(\frac{B_{ij} - L_j/2}{A_{ij} + B_{ij}}) \quad i = 1, 2, \cdots, C$
   s.t. $L_j > 0$

Based on the above theorem, we can get the update rules on u_{ij}, v_i (cluster centers), where.

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m} \tag{8}$$

Now, except for the cluster number C, only parameter $\alpha = [\alpha_1, \alpha_2, \Lambda, \alpha_C]^T$ must be assigned by hand in EPCM. Here, we present a clustering validity index function on α to estimate the optimal α. In general, an effective fuzzy clustering strategy is to attempt to assign all elements of partition matrix U to 0 or 1. From the above procedure of determining the optimal u_{ij}, we know that u_{ij} may be taken as the function of α. So we can use the following clustering validity index function on α to estimate the optimal α.

$$J(\alpha) = \sum_{i=1}^{C}\sum_{j=1}^{N} g(u_{ij}(\alpha)),\ g(x) = (1-|2x-1|)^{\gamma} \tag{9}$$

where $g(\cdot)$ is the function of u_{ij} and the exponent γ is set to be 2 from our experimental comparisons. It can be easily found that when u_{ij} approximates 1 or 0, the values of $g(u_{I,\ ij})$ approximates 0. Thus, the minimum of $J(\alpha)$ corresponds to an optimal α.

Furthermore, in order to effectively avoid the coincident cluster problem of PCM, we impose (10) and (11) as the constraints of (9).

$$\begin{aligned}&\sqrt{\eta_{Ii} + (\eta_{Oi} - \eta_{Ii}) \times \alpha_i / 2} \le \\ &\min\{\| v_i - v_k \| - \sqrt{\eta_{Ik}}, k \ne i\}\end{aligned} \quad \forall i \tag{10}$$

$$0.05 \le \alpha_i \le 0.6 \quad \forall i \tag{11}$$

where the upper and lower boundaries of α in (11) can be slightly adjusted according to the practical requirements. The current upper and lower boundaries are based on our experiments.

In terms of the above analysis, we now give the complete description of EPCM as follows:

Algorithm EPCM

1. Use FCM clustering algorithm to obtain an initial fuzzy partition U;
2. Compute v_i using (8) and update η_{Ii}, η_{oi} using (6);
3. Minimize the index function (9) with the constraints (10) and (11) to obtain the optimal parameter $\alpha = [\alpha_1, \alpha_2, \Lambda, \alpha_C]^T$;
4. Repeat the following steps:
   a) Use the above procedure of determining the optimal u_{ij};
   b) Update the v_i $\forall_i$ using (8);
   c) Update the η_{Ii}, η_{oi} using (6);

   Until some termination conditions are satisfied.

In terms of the above discussion and the design strategy of EPCM, the relationship between data points and clusters can be proposed as follows:

$$x_j \in \begin{cases} \text{k cluster} & u_{kj} \ge 0.5 \\ \text{outlier cluster} & u_{kj} < 0.5 \end{cases},\ k = \arg\max_i u_{ij} \tag{12}$$

EXPERIMENTS STUDIES

In this section, we will examine the clustering performance of EPCM, where two experiments are arranged. The first is used to examine the ability of EPCM in labeling the outliers adaptively and accurately. Experimental results in the second experiment, several color image segmentation results demonstrate its potential in real applications.

Experiment 1: In this experiment, the datasets X400 and X550 introduced in (Pal *et al*., 2005) are used to examine the adaptive ability of EPCM in labeling the outliers. X400 consists of two clusters which are satisfied with two dimension standard normal distributions, whose centers are $[5.0\ 6.0]^T$ and $[5.0\ 5.6]^T$ and the number of data points of every cluster is 200. X550 is produced by X400 added into 150 noises that are random and uniformly distributed on $[0, 15] \times [0, 11]$.

The clustering results obtained by EPCM on two datasets are displayed in Fig. 1. It appears that EPCM adaptively label the outliers for two cases, where the X400 has no noises but X550 has. Furthermore, it can be found that the outliers labeled by EPCM on X550 are real noises compared to X400, i.e., the labeled results of EPCM are accurate on X550. In summary, EPCM can label the outliers adaptively and accurately on dataset X550, which will be still valid for some other datasets from our other experiments such as the followed image segmentation experiment.

Experiment 2: In this experiment, we will apply EPCM to the color peach image segmentations. Image segmentation is a fundamental and important research topic in image processing field. Clustering is one of the important image segmentation methods. The first task of color image segmentation is to choose a feature space to represent the pixels of an image. In our experiment, the HSV feature space is adopted.

Figure 2 show the original images and their segmentation results using different clustering algorithms including EPCM, FCM, PCM, IPCM (Improved PCM) (Zhang and Yeung, 2004) and PFCM respectively, where cluster number C is set to 3 and outliers are denoted by black points in the segmentation results. Here, every color image is mainly composed of three parts: peach, background and shadow. Besides of these, a lot of pixels indicate other colors. To get a good segmentation result, clustering algorithms should have robustness and high

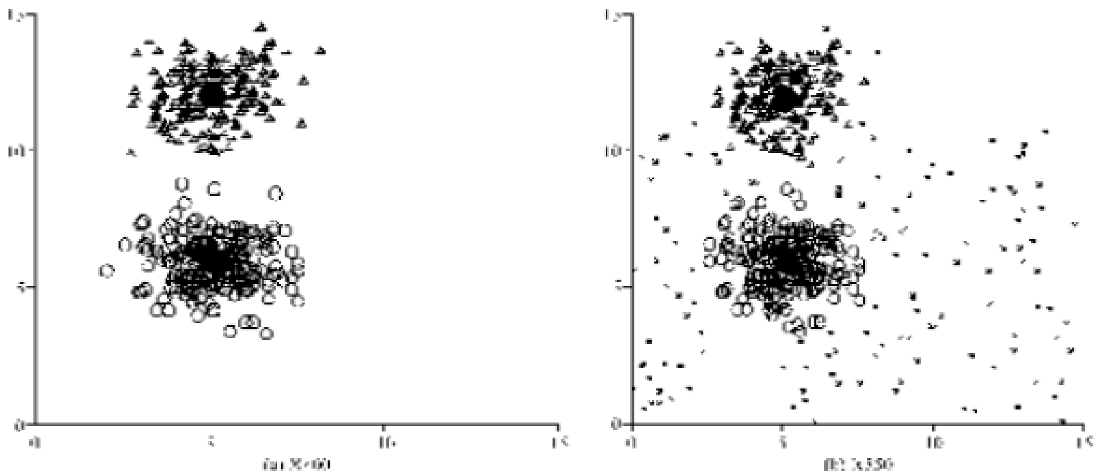

Fig. 1: The clustering results obtained by EPCM on datasets X400 and X550

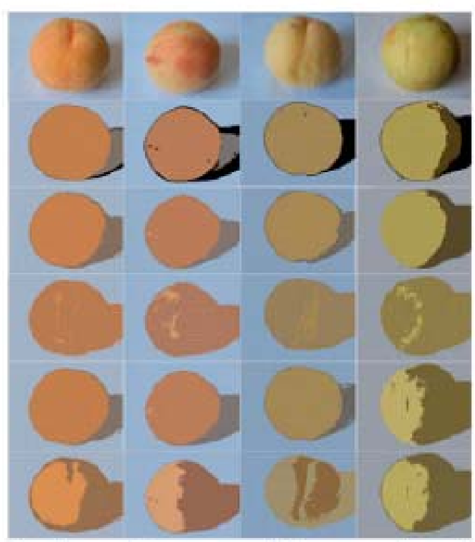

Fig. 2: The segmentation results obtained by several clustering algorithms for the peach images The images in the first row are the original images, and the 2nd~7th rows are the corresponding segmentation results obtained by EPCM, FCM, PCM, IPCM and PFCM respectively

clustering performance (that is to say, the clustering algorithm can label the outliers adaptively and accurately) because of the existence of outliers in original images. It is obvious from Fig. 2 that the results obtained by EPCM have the good effect and superior to all others, which is mainly because that EPCM not only can get the robust clustering centers but also has the ability to label the outliers adaptively and accurately, while the PCM and its variants cannot reach the latter. From this experiment, the good clustering performance is also clarified.

CONCLUSIONS

In this study, to overcome the weaknesses of algorithm PCM, we propose a new enhanced possibilistic clustering algorithm EPCM. The distinctive features can be concluded as follows.

- Due to the introduction of the strategy of flexible hyperspheric partition, EPCM can avoid the use of some parameters that are needed in PCM. Thus, EPCM actually avoid these parameters' influence on the clustering results.
- In particular, EPCM not only can effectively neglect outliers in the dataset but also has ability to label them adaptively and accurately that cannot be realized by the many current possibilistic clustering algorithms, which results in better clustering performance and feasibility in real applications such as image segmentations illustrated in this study.

Although EPCM reveals better performances than PCM and its variants in the above, it has some problems to be further studied. For example, just like PCM, it also cannot obtain very satisfactory clustering results for the non-ball distributed datasets. To deal with these problems, the kernel method is usually introduced. Based on the similar idea in the paper, we can extend the EPCM to solve these problems in near future.

ACKNOWLEDGMENTS

This research is partially supported by Hong Kong PolyU grant, Jiangsu Natural Science Foundation (BK2003017), New century Outstanding Young Scholar Grant of Ministry of Education of China (NCET-04-0496), 2005 Key R&D Project of Ministry of Education of China, The Key Lab. of National Pattern Recognition of Automation Institute of CAS of China.

REFERENCES

Barni, M., V. Cappellini and A. Mecocci, 1996. Comments on A possibilistic approach to clustering. IEEE Trans. Fuzzy Sys., 4: 393-396.

Cheng, C.T. and K.W. Chau, 2001. Fuzzy iteration methodology for reservoir flood control operation. J. Am. Water Resour. Assoc., 37: 1381-1388.

Cheng, C.T. and K.W. Chau, 2002. Three-person multi-objective conflict decision in reservoir flood control. Eur. J. Operat. Res. 142: 62-631.

Cheng, C.T., C.P. Ou and K.W. Chau, 2002. Combining a fuzzy optimal model with a genetic algorithm to solve multiobjective rainfall-runoff model calibration. J. Hydrol., 268: 72-86.

Dave, R.N. and R. Krishnapuram, 1997. Robust clustering methods. A unified view. IEEE Trans. Fuzzy Sys., 5: 270-293.

Dave, R.N. and S. Sen, 2002. Robust fuzzy clustering of relational data. IEEE Trans. Fuzzy Sys., 10: 713-727.

Gustafson, E.E. and W.C. Kessel, 1979. Fuzzy clustering with a fuzzy covariance matrix. In: Proc. IEEE Conf. Decision and Control, San Diego, CA, pp: 761-76

Jian Yu, 2005. General c-means clustering model. IEEE Trans. on Pattern Analysis and Machine Intelligence, 27: 1197-1211.

Krishnapuram, R. and J.M. Keller, 1993. A possibilistic approach to clustering. IEEE Trans. Fuzzy Sys., 1: 98-110.

Krishnapuram, R. and J.M. Keller, 1996. The possibilistic c-Means algorithm: Insights and recommendations. IEEE Trans. Fuzzy Sys., 4: 385-393.

Pal, N.R., K. Pal, J.M. Keller and J.C. Bezdek, 2005. A possibilistic fuzzy c-means clustering algorithm. IEEE Trans. Fuzzy Sys., 13: 517-530.

Timm, H., C. Borgelt, C. Doring and R. Kruse, 2001. Fuzzy cluster analysis with cluster repulsion. Presented at the Euro. Symp. Intelligent Technologies (EUNITE), Tenerife, Spain.

Timm, H. and R. Kruse, 2002. A modification to improve possibilistic fuzzy cluster analysis. Presented at the IEEE Int. Conf. Fuzzy Systems, FUZZ-IEEE' 2002, Honolulu, HI.

Timm, H., C. Borgelt, C. Doring and R. Kruse, 2004. An extension to possibilistic fuzzy cluster analysis. Fuzzy Sets Syst., 147: 3-16

Yang, M.S. and K.L. Wu, 2006. Unsupervised possibilistic clustering. Pattern Recognition, 39: 5-21.

Zhang, J.S. and Y.W. Yeung, 2004. Improved possibilistic C-means clustering algorithms. IEEE Trans. Fuzzy Sys., 12: 209-217.