

An automatic fuzzy c -means algorithm for image segmentation

Yan-ling Li · Yi Shen

Published online: 4 June 2009
© Springer-Verlag 2009

Abstract Fuzzy c -means (FCM) algorithm is one of the most popular methods for image segmentation. However, the standard FCM algorithm must be estimated by expert users to determine the cluster number. So, we propose an automatic fuzzy clustering algorithm (AFCM) for automatically grouping the pixels of an image into different homogeneous regions when the number of clusters is not known beforehand. In order to get better segmentation quality, this paper presents an algorithm based on AFCM algorithm, called automatic modified fuzzy c -means cluster segmentation algorithm (AMFCM). AMFCM algorithm incorporates spatial information into the membership function for clustering. The spatial function is the weighted summation of the membership function in the neighborhood of each pixel under consideration. Experimental results show that AMFCM algorithm not only can spontaneously estimate the appropriate number of clusters but also can get better segmentation quality.

Keywords Image segmentation · Fuzzy c -means · Fuzzy clustering · K -means · Spatial information

1 Introduction

Image segmentation is one of the most difficult and challenging problems in the image processing. It denotes

a process by which an image is partitioned into non-overlapping regions such that each region is homogeneous and the union of two adjacent is heterogeneous. The extent of homogeneity of the segmented region can be measured using some image property [e.g. Pixel intensity (Dunn 1974)]. Most of segmentation algorithms aim at the concrete problem; there is not a universal segmentation method for all images.

Clustering can be defined as the optimal partitioning of a given set of n data points into c subgroups, such that data points belonging to the same group are as similar to each other as possible, whereas data points from two different groups have the maximum difference. Image segmentation can be treated as a clustering problem where the features describing a pixel correspond to a pattern, and each image region corresponds to a cluster. Therefore, many clustering algorithms have widely been used to solve the segmentation problem [e.g. K -means (Tou and Gonzalez 1974), FCM, ISODATA (Ball and Hall 1967) and Snob (Wallace and Boulton 1968)]. In the last decades, fuzzy segmentation methods, especially the fuzzy c -means algorithm (FCM) (Bezdek 1981), have been widely used in the image segmentation (Udupa and Samarasekera 1996; Yamany et al. 1999) and such a success chiefly attributes to introduction of fuzziness for the belongingness of each image pixel, which makes the clustering methods able to retain more information from the original image than the crisp or hard segmentation methods (Pham and Prince 1999). However, most of the unsupervised fuzzy clustering algorithms assume a prior knowledge of the number of classes, c , while in many practical situations, the appropriate number of classes is unknown or impossible to determine even approximately. Find an optimal number of clusters in a large dataset is usually a challenging task.

Y. Li (✉) · Y. Shen
Department of Control Science and Engineering,
Institute of System Engineering, Huazhong University
of Science and Technology, 430074 Wuhan, China
e-mail: ly175@163.com

Y. Shen
e-mail: yishen64@163.com

Because the fuzzy partitions obtained using the FCM algorithm depend on the choice of c . It is necessary to validate each of the fuzzy c -partitions once they are found (Pal and Bezdek 1995). This validation is carried out by a cluster validity index, which evaluates each of the fuzzy c -partitions and determines the optimal partition or the optimal number clusters (c) from them. Many validation criteria have been proposed for evaluating fuzzy partitions (Kim et al. 2004), such as Bezdek's partition coefficient (V_{pc}) and partition entropy (V_{pe}), Xie–Beni's (V_{xb}), Fukuyama–Sugeno's (V_{FS}), Kwon's (V_K), Boudraa's (V_{CWB}) and Dae–WonKim's (V_{OS}). Researchers use nests ISODATA and GA to get the optimal partition according to various validation criteria. However, it is necessary to pre-define the maximum number of clusters (c) when we use the exhaustive attack method to get the optimal partition number. Obviously, the exhaustive attack method is feasible when the data point is less. But the exhaustive attack method is impossible when the data point is big. Moreover, we have not a generic theory and implement method to determine the maximum number of clusters (c) at present. In Kim et al. (2004), a new cluster validity index is proposed that determines the optimal partition and optimal number of clusters for fuzzy partitions obtained from the fuzzy c -means algorithm. The proposed validity index exploits an overlap measure and a separation measure between clusters. In Hall and Kanade (2005), the authors use a fuzzy cluster validity metric proposed by Xie and Beni as the criterion for evaluating a partition produced by swarm based clustering. In Shelokar et al. (2004), Hall and Kanade (2007) and Huang et al. (2006), colony algorithm is used to get the initial cluster centers.

In this article, an automatic fuzzy clustering algorithm (AFCM) is proposed which can determine the number of clusters automatically. In order to get better segmentation quality, we presents an algorithm based on AFCM algorithm, called automatic modified fuzzy c -means cluster segmentation algorithm (AMFCM). AMFCM algorithm incorporates spatial information into the membership function for clustering. The spatial function is the weighted summation of the membership function in the neighborhood of each pixel under consideration. Experimental results show that AMFCM algorithm is an effective method and outperforms the standard FCM algorithm and AFCM algorithm.

The rest of the paper is organized as follows. In Sect. 2, we formally describe the problem of fuzzy clustering, briefly outline the standard FCM algorithm, and provide a fuzzy clustering validity measures. The proposed algorithm is illustrated in Sect. 3. Experimental results along with statistical tests of significance are presented in Sect. 4. In Sect. 5, we conclude this paper.

2 Fuzzy c -means clustering algorithm

Let $X = \{x_1, x_2, \dots, x_n\}$ be a set of n patterns or data points, each having d features. These patterns can also be represented by a profile data matrix $X_{n \times d}$ having n d -dimensional row vectors. The i th row vector \vec{X}_i characterizes the i th object from the set X and each element X_{ij} in \vec{X}_i corresponds to the j th real value feature $j = 1, 2, \dots, d$ of the i th pattern $i = 1, 2, \dots, n$. Given such a $X_{n \times d}$, a clustering algorithm tries to find a partition $X = \{X_1, X_2, \dots, X_c\}$ such that the similarity of the patterns in the same cluster X_i is maximum and patterns from different clusters differ as far as possible. The partitions should maintain the following properties:

- (1) $X_1 \cup X_2 \cup \dots \cup X_c = X$
- (2) $X_i \cap X_j = \phi, 1 \leq i \neq j \leq c$
- (3) $X_i \neq \phi \forall i \in \{1, 2, \dots, c\}$

In the case of fuzzy clustering, a pattern may belong to all the classes with a certain fuzzy membership grade for each class. In this case, we need to evolve an appropriate partition matrix $U = [u_{ij}]_{c \times n}$, where $u_{ij} \in [0, 1]$, such that u_{ij} denotes the grade of membership of the j th element to the i th cluster.

2.1 The standard FCM algorithm

The standard FCM algorithm is an iterative optimization that minimizes the cost function defined as follows:

$$J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 \quad (1)$$

subject to

$$\begin{aligned} \sum_{i=1}^c u_{ij} &= 1, \quad 1 \leq j \leq n \\ 0 < \sum_{j=1}^n u_{ij} < n, \quad 1 \leq i \leq c \\ \sum_{i=1}^c \sum_{j=1}^n u_{ij} &= n \end{aligned} \quad (2)$$

where $m \in (1, \infty)$ controls the fuzziness of the resulting partition, and $m = 2$ is used in this study, u_{ij} represents the membership of pixel x_j in the i th clustering, $d_{ij} = \|x_j - v_i\|$ represents distance between the pixel x_j and the cluster center v_i . In the image clustering, the most commonly used feature is the gray-level value, or intensity of image pixel. Thus the FCM cost function is minimized when high membership values assigned to pixels whose intensities are close to the centroid of their clusters, and low membership values are assigned when the point is far from the centroid. The membership function represents the probability that a

pixel belongs to a specific cluster. In the FCM algorithm, the probability is dependent solely on the distance between the pixel and each individual cluster center in the feature domain. The membership function and cluster centers are updated by the following:

$$u_{ij} = \left[\sum_{k=1}^c \left(\frac{d_{ij}^2}{d_{kj}^2} \right)^{\frac{1}{m-1}} \right]^{-1} \tag{3}$$

$$v_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \tag{4}$$

Starting with an initial guess for each cluster center, the FCM converges to a solution for v_i representing the local minimum or a saddle point of the cost function. Convergence can be detected by comparing the changes in the membership function or the cluster center at two successive iteration steps. After the convergence, defuzzification is applied to assign each pixel to a specific cluster for which the membership is maximal. The standard FCM algorithm degenerate to hard c -means algorithm when $u_{ij} \in \{0, 1\}$ and $m = 1$. In this case, the cost function defined as follows:

$$J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij} d_{ij}^2 \tag{5}$$

subject to

$$u_{ij} = \begin{cases} 1 & d_{ik} = \min_{1 \leq r \leq n} \{d_{ir}\} \\ 0 & \text{others} \end{cases} \tag{6}$$

2.2 Cluster validity functions

In order to obtain a quantitative comparison, two types of cluster validity functions, fuzzy partition and feature structure, are often used to evaluate the performance of clustering in different clustering methods. The representative functions for the fuzzy partition are partition coefficient V_{pc} (Bezdek 1974) and partition entropy V_{pe} (Bezdek 1975). They are defined as follows:

$$V_{pc} = \frac{1}{n} \sum_{i=1}^c \sum_{j=1}^n u_{ij}^2 \tag{7}$$

and

$$V_{pe} = \frac{-1}{n} \sum_{i=1}^c \sum_{j=1}^n (u_{ij} \log u_{ij}) \tag{8}$$

The idea of these validity functions is that the partition with less fuzziness means better performance. As a result, the best clustering is achieved when the value V_{pc} is maximal or V_{pe} is minimal.

Other validity functions based on the feature structure, for example, Xie–Beni’s V_{xb} (Xie and Beni 1991), defined as

$$V_{xb} = \frac{\sum_{i=1}^c \sum_{j=1}^n u_{ij}^2 \|x_j - v_i\|^2}{n \left(\min_{i \neq k} \{ \|v_k - v_i\|^2 \} \right)} \tag{9}$$

A good clustering result generates samples that are compacted within one cluster and samples that are separated between different clusters. Minimizing V_{xb} is expected to lead to a good clustering.

3 Automatic modified FCM algorithm

3.1 Automatic FCM algorithm

The standard FCM algorithm is often determined by the number of clusters manually and randomly initialized in the membership value for fuzzy clustering algorithm. In this paper, we propose a method to improve the standard FCM algorithm and which can automatically determine the optimal cluster number called automated fuzzy c -means algorithm. The proposed algorithm can be summarized as follows: firstly, we use hard c -means algorithm to find the initial two centroids which are closed to the real cluster centers. Secondly, FCM is applied to calculate the membership matrix that will be iterated. We calculate the pixel proportion of each cluster (pop_i) by summarizing pixels in a cluster and diving by the number of image pixels. The optimal clusters will be reached when the pop_i of at most two clusters have values less than δ . The parameter δ is represented as the smallest percentage for the existing cluster. A cluster whose area is $<10\%$ will be discarded. Because it is too small an area to become a meaningful area. So, the value of δ is set as 0.1.

The proposed algorithm is denoted AFCM algorithm and can be summarized in the following step:

- Step 1 Set the initialize two cluster centers by using a random generator from the original dataset. Set $m = 2$, decide ε , where ε is a small positive constant.
- Step 2 Calculate initial centroids by hard c -means algorithm by using (5).
- Step 3 Repeat
 - (1) Update centroid $V^{(l)}$ by using (4).
 - (2) Compute the membership matrix $U^{(l)}$ by using Eq. 3. Until $\|U^{(l)} - U^{(l+1)}\| < \varepsilon$.
- Step 4 Check the automatic optimal cluster numbers:
 - (1) Calculate pop_i value: $pop_i = |c_i|/n$, where $|c_i|$ is the size of cluster i ; $i = 1, 2, \dots, c$. n is the total number of image pixels.

- (2) Stop process when the optimal cluster is realized, otherwise go to step 5.

Step 5 Increase the number of cluster $c = c + 1$, set initial value of new centroid v_c , go to step 3.

We can get the optimal membership matrix U when the algorithm is finished. Then, we can get the final segmentation image.

3.2 Automatic modified FCM algorithm

3.2.1 The spatial feature of pixel

One of the important characteristics of an image is that neighboring pixels are highly correlated. In other words, these neighboring pixels possess similar feature values, as we can know that they should be the close membership value according to the amended membership function and the probability that they belong to the same cluster is great. This spatial relationship is important in clustering, but it is not utilized in a standard FCM algorithm. All samples are used as dispersive points when using the standard FCM algorithm to cluster. So the standard FCM algorithm is sensitive to noise. To exploit the spatial information, a spatial function is defined as

$$h_{ij} = \frac{\sum_{t \in \Omega_j} u_{it} \beta_t}{\sum_{k=1}^c \sum_{t \in \Omega_j} u_{kt} \beta_t} \quad (10)$$

where Ω_j represents a square window centered on pixel x_j in the spatial domain. A 3×3 window was used in this paper. Just like the membership function, the spatial function h_{ij} represents the probability that pixel x_j belongs to i th clustering. The spatial function is the largest if all of its neighborhood pixels belong to i th clustering, and is the smallest if none of its neighborhood pixels belong to i th clustering. In the spatial function, c is the number of desired clusters, β_t is the contribution factor of the neighbor x_t , and the weight u_{it} is the membership of the pattern x_t to the i th clustering. The factor β_t represents the contribution of the neighbor x_t 's membership to the overall spatial dissimilarity. Generally, the nearer the neighbor is, the stronger the interaction becomes, and the bigger the contribution will be. Consequently, for each neighbor $x_t (t \in \Omega_j)$, β_t is a non-increasing function of the distance between sites j and t , and is defined as follow:

$$\beta_t = \frac{1}{1 + \exp(\theta \|j - t\|)} \quad (11)$$

where the coefficient θ determines how fast the interaction between two sites diminished with the increasing of the distance of them. A small θ leads to a similar contribution of different neighbors; whereas a big θ makes the spatial

dissimilarity highly depend on those adjacent neighbors. In this paper, θ is empirically set to 0.6.

3.2.2 Automatic modified FCM algorithm

With the spatial feature of pixel, the modified membership of the current pixel is defined as follows:

$$u_{ij}^* = \frac{u_{ij} h_{ij}}{\sum_{k=1}^c u_{kj} h_{kj}} \quad (12)$$

the new cluster center is represented as follows:

$$v_i^* = \frac{\sum_{j=1}^n (u_{ij}^*)^m x_j}{\sum_{j=1}^n (u_{ij}^*)^m} \quad (13)$$

then the proposed automatic modified FCM algorithm can be summarized in the following step:

- Step 1 Get the number of clusters by AFCM algorithm.
 Step 2 Calculate the spatial function h_{ij} and β_t by using (10) and (11).
 Step 3 Update centroid $V^{(l)}$ by using (13).
 Step 4 Update the membership matrix $U^{(l)}$ by using Eq. 12.
 Step 5 If $\|U^{(l)} - U^{(l+1)}\| < \varepsilon$ then stop; otherwise, jump to step 2.

4 Experiment results and analysis

In this section, we test the three algorithms which are FCM, AFCM and AMFCM on three images. One is the cameraman image with 256×256 pixels, another is the simulative brain image with 256×256 pixels, and the last one is the lena image is size of 512×512 . They are shown, respectively, in Figs. 1a, 2a and 3a. For all cases, unless otherwise stated, the fuzzy weighting exponent $m = 2$, $\delta = 0.1$ and $\varepsilon = 1e - 5$. All the algorithms are coded in matlab7.4. For each example, after we get the appropriate cluster number from the AFCM algorithm, we use the cluster number to process FCM method. Figures 1b, 2b and 3b show the segmented results of the FCM algorithm for the three images. Figures 1c, 2c and 3c show the segmented results of the AFCM algorithm for the three images. The segmented results of the AMFCM algorithm for the three images are, respectively, showed in Figs. 1d, 2d and 3d. Table 1 tabulates the experimental results of the three algorithms for the three images.

Table 1 shows the validity functions used to evaluate the performance of the three algorithms for three images. As we can see, for V_{pc} (V_{pe}), the AFCM algorithm is smallest (greatest). AFCM algorithm can automatically estimate the appropriate number of clusters, but the segmentation quality of AFCM is worst in the three algorithms. Visually,

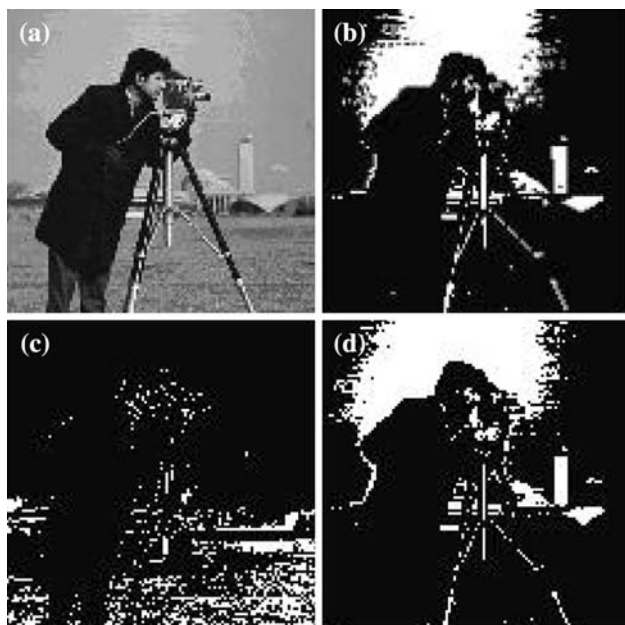


Fig. 1 Comparison of segmented results on cameraman image with 256×256 pixels. **a** primary image, **b** segmented result of FCM algorithm, **c** segmented result of AFCM algorithm, **d** segmented result of AMFCM algorithm

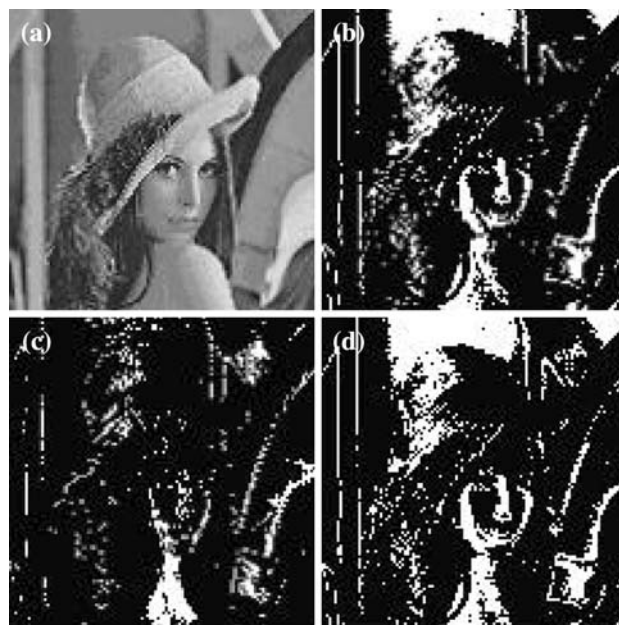


Fig. 3 Comparison of segmented results on lena image with 512×512 pixels. **a** primary image, **b** segmented result of FCM algorithm, **c** segmented result of AFCM algorithm, **d** segmented result of AMFCM algorithm

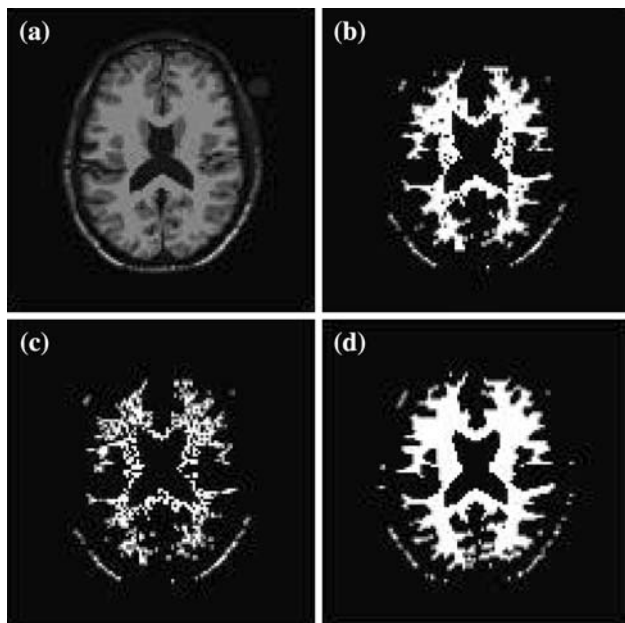


Fig. 2 Comparison of segmented results on simulative brain image with 256×256 pixels. **a** primary image, **b** segmented result of FCM algorithm, **c** segmented result of AFCM algorithm, **d** segmented result of AMFCM algorithm

the quality of Figs. 1d, 2d and 3d are best which indicates that AMFCM algorithm performs best. For V_{pc} (V_{pe}), the AMFCM algorithm is greater (smaller) than conventional FCM and AFCM algorithm. In most cases, the validity

Table 1 The experimental results of the three algorithms for three images

Image	Algorithm	V_{pc}	V_{pe}
256×256	FCM	0.7517	0.5235
Cameraman	AFCM	0.6788	0.6007
Image	AMFCM	0.9194	0.1424
256×256	FCM	0.8930	0.2234
Simulative brain	AFCM	0.5905	0.6357
Image	AMFCM	0.9650	NaN
512×512	FCM	0.7048	0.6504
Lena	AFCM	0.6768	0.6994
Image	AMFCM	0.8815	0.2064

functions based on the fuzzy partition were best for the AMFCM algorithm which indicates that AMFCM algorithm is an effective method and outperforms the standard FCM algorithm and AFCM algorithm.

5 Conclusions

FCM algorithm is one of the most popular methods for image segmentation. However, the cluster number must be known in advance when we use the standard FCM algorithm. This paper has presented a new AFCM algorithm. An important feature of AFCM algorithm is that it is able to find the optimal number of clusters automatically.

However, the segmentation quality of AFCM is poor. So, we proposed AMFCM algorithm based on AFCM algorithm which not only can get better segmentation quality but also can find the optimal number of clusters automatically. Experimental results show that the proposed AMFCM algorithm outperforms the standard FCM algorithm and AFCM algorithm. Not only does the proposed AMFCM algorithm can automatically estimate the appropriate number of clusters, it also has better segmentation quality than FCM's.

Acknowledgments The authors would like to thank the anonymous referees for their helpful comments and suggestions to improve the presentation of the paper. This research is supported by the National Natural Science Foundation of China (No. 60874031), the Fund for International Cooperation and Exchange of the National Natural Science Foundation of China (No. 60740430664) and the Special Research Fund for the Doctoral Program of Higher Education of China (20070487052).

References

- Ball G, Hall D (1967) A clustering technique for summarizing multivariate data. *Behav Sci* 12:153–155
- Bezdek JC (1974) Cluster validity with fuzzy sets. *J Cybern* 3:58–73
- Bezdek JC (1975) Mathematical models for systematic and taxonomy. In: *Proceedings of eight international conference on numerical taxonomy*, San Francisco, pp 143–166
- Bezdek JC (1981) *Pattern recognition with fuzzy objective function algorithms*. Plenum Press, New York
- Dunn JC (1974) A fuzzy relative of the ISODATA process and its use in detecting compact, well-separated clusters. *J Cybern* 3:32–57
- Hall LO, Kanade PM (2005) Swarm based fuzzy clustering with partition validity. *Proc. 14th IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE'05)*. IEEE Press, Piscataway, NJ, 991–995
- Hall LO, Kanade PM (2007) Fuzzy ants and clustering. *IEEE Trans Syst Man Cybern A Syst Hum* 37(5):758–769
- Huang GR, Wang XF, Cao XB (2006) Ant colony optimization algorithm based on directional pheromone diffusion. *Chin J Electron* 15(3):447–450
- Kim DW, Lee KH, Lee D (2004) On cluster validity index for estimation of optimal number of fuzzy clusters. *Pattern Recognit* 37:2009–2024
- Pal NR, Bezdek JC (1995) On cluster validity for the fuzzy c-means model. *IEEE Trans Fuzzy Syst* 3(3):370–379
- Pham DL, Prince JL (1999) An adaptive fuzzy c-means algorithm for image segmentation in the presence of intensity inhomogeneities. *Pattern Recognit Lett* 20:57–68
- Samarasekera S (1996) Fuzzy connectedness and object definition: theory, algorithm and applications in image segmentation. *Graph Models Image Process* 58(3):246–261
- Shelokar PS, Jayaraman VK, Kulkarni BD (2004) An ant colony approach for clustering. *Anal Chim Acta* 509(2):187–195
- Tou JT, Gonzalez RC (1974) *Pattern recognition principles*. Addison-Wesley, London
- Wallace CS, Boulton DM (1968) An information measure for classification. *Comput J* 11(2):185–194
- Xie XL, Beni GA (1991) Validity measure for fuzzy clustering. *IEEE Trans Pattern Anal Mach Intell* 3:841–846
- Yamany SM, Farag AA, Hsu S (1999) A fuzzy hyperspectral classifier for automatic target recognition (ATR) systems. *Pattern Recognit Lett* 20:1431–1438