Remote Sensing Image Classification Based on Improved Fuzzy c-Means

YU Jie  GUO Peihuang  CHEN Pinxiang  ZHANG Zhongshan  RUAN Wenbin

Abstract  Classification is always the key point in the field of remote sensing. Fuzzy c-Means is a traditional clustering algorithm that has been widely used in fuzzy clustering. However, this algorithm usually has some weaknesses, such as the problems of falling into a local minimum, and it needs much time to accomplish the classification for a large number of data. In order to overcome these shortcomings and increase the classification accuracy, Gustafson-Kessel (GK) and Gath-Geva (GG) algorithms are proposed to improve the traditional FCM algorithm which adopts Euclidean distance norm in this paper. The experimental result shows that these two methods are able to detect clusters of varying shapes, sizes and densities which FCM cannot do. Moreover, they can improve the classification accuracy of remote sensing images.

Keywords  FCM algorithm; GK algorithm; GG algorithm; remote sensing classification

Introduction

As one of the unsupervised classification methods, clustering analysis is a process of partitioning and classifying objects according to their similarity. Also, clustering analysis does not need prior knowledge, and just classify into hard partition and fuzzy partition based on the number of category each pixel belongs to. Hard clustering methods are based on classical set theory, and require that an object either does or does not belong to a cluster. However, fuzzy clustering methods allow objects to belong to several categories simultaneously, with different degrees of membership[1-4]. Because of the uncertainty of remote sensing information and the existence of mixed pixels, it is necessary to study the fuzzy classification. Fuzzy clustering method, as one of an important method of unsupervised classification, is based on the fuzzy set Zadeh created and fuzzy mathematic theory, and classifies objects into a certain category in a certain level of membership and if the membership degree reaches the maximum, then the degree of belongingness is highest[5]. Also, it has been widely used in the field of pattern recognition recently[6].

Fuzzy c-means is a traditional clustering algorithm that has been widely used in fuzzy clustering, yet it does not consider the spatial information between pixels, so it is very sensitive to noise[7, 8]. Furthermore, it has some shortcomings such as the problem of falling into a local minimum. The FCM algorithm computes with the standard Euclidean distance norm, which induces hyper-spherical clusters. Hence, it can only detect clusters with the same shape and orientation[4]. Zhang Lu and Liao Mingsheng proposed an improved fuzzy clustering algorithm considering context information by incorporating the concept of spatial fuzzy membership under Markov Random
Field framework, in this way, accuracy and reliability of clustering can be improved compared with the traditional ones; Reinhard Beichel et al. examined the characteristics of unsupervised fuzzy classification methods and applied it to a section of a Landsat TM scene from the surroundings of Vienna, and obtained a better result. However, there are only a few studies in the field of remote sensing images classification referring to Gustafson-Kessel (GK) and Gath-Geva (GG) algorithm based on improved Fuzzy c-means. Therefore, this paper discusses the application of GK and GG algorithms in the remote sensing images classification.

1 Fuzzy c-means (FCM) algorithm

The FCM algorithm is a kind of clustering algorithm based on partition, and it can classify the sample data automatically. This algorithm optimizes the fuzzy objective function to get the membership degree that each sample belongs to the center of the category. The Fuzzy c-means clustering algorithm divides the n vectors \( x_i (i = 1, 2, \ldots, n) \) into c fuzzy groups, and it is based on the minimization of an objective function called c-means function, which is defined by Dunn as:

\[
J(X;U,V) = \sum_{i=1}^{N} \sum_{k=1}^{c} \mu_{ik}^m \|x_i - v_k\|^2
\]  

(1)

Where \( V = [v_1, v_2, \ldots, v_c] \) is a vector of cluster centers, which have to be determined, \( v_k \) is the cluster center of each category; \( c \) is the number of category; \( U \) is a membership matrix, the \( i \)-th column of \( U \) contains values of the membership function of the \( i \)-th fuzzy subset of \( x_i \). The expression \( \sum_{j=1}^{c} u_{ij} = 1, \forall j = 1, \ldots, n \) constrains the sum of each column to 1, thus, the total membership of each \( x_i \) equals one; \( m \in [1, \infty) \) is the fuzziness weighting exponent.

The FCM algorithm computes with the standard Euclidean distance norm, which is a squared inner-product distance norm.

\[
D_{ik}^2 = \|x_i - v_k\|^2 = (x_i - v_k)^T A (x_i - v_k)
\]  

(2)

Where \( D_{ik} = \|x_i - v_k\|^2 \) is the Euclidean distance between \( i \)-th clustering center and \( k \)-th data point; \( A \) is the norm inducing matrix. It may minimize the expression (1) only if

\[
\mu_{ik} = \frac{1}{\sum_{j=1}^{c} (D_{ik} / D_{jk})^{2(m-1)}} \cdot 1 
\]  

(3)

and

\[
v_i = \sum_{k=1}^{c} \mu_{ik}^m x_k / \sum_{k=1}^{c} \mu_{ik}^m, \quad 1 \leq i \leq c
\]  

(4)

Based on the two conditions above, the FCM algorithm is a simple iterative process.

2 Gustafson-Kessel(GK) algorithm

The GK algorithm was put forward by Gustafson and Kessel in 1979, which extended the standard fuzzy c-means algorithm by employing an adaptive distance norm, in order to detect clusters of different geometrical shapes in one data set. Each cluster has its own norm-inducing matrix \( A \) in the algorithm, which meets the following inner-product norm—squared Mahalanobis distance norm:

\[
D_{ik}^2 = (x_i - v_k)^T A_i (x_i - v_k), 1 \leq i \leq c, 1 \leq k \leq N
\]  

(5)

Where, the matrices \( A_i \) are adopted as optimization variables in the c-means function, thus allowing each cluster to adapt the distance norm to the local topological structure of the data.

The objective function in GK algorithm is defined as:

\[
J(X;U,V,A) = \sum_{i=1}^{N} \sum_{k=1}^{c} \mu_{ik}^m (x_i - v_k)^T A_i (x_i - v_k)
\]  

(6)

In this expression, \( X \) is the data set; \( U \) is also a membership matrix; \( V \) is the center of each category; \( A \) denotes a \( c \)-tuple of the norm-inducing matrices, that is:

\[
A = (A_1, A_2, \ldots, A_c)
\]

To obtain a feasible solution, \( A_i \) must be constrained, usually, using the Lagrange multiplier method, the following expression for \( A_i \) is obtained:

\[
A_i = [\rho_i \det(F_i)]^{1/n} F_i^{-1}
\]  

(7)

where \( \rho_i \) is constant for each cluster, and \( F_i \) is the fuzzy covariance matrix of the \( i \)-th cluster, it is defined by:

\[
F_i = \sum_{k=1}^{c} (\mu_{ik})^m (x_i - v_k)(x_i - v_k)^T / \sum_{k=1}^{c} (\mu_{ik})^m
\]  

(8)

Like the FCM algorithm, the GK algorithm is also an iterative process.
3 Gath-Geva (GG) algorithm

The GG clustering algorithm employs a distance norm based on the fuzzy maximum likelihood estimates, which was proposed by Bezdek and Dunn[4]:

\[
D_k(x_i, v_i) = \sqrt{\frac{\text{det}(F_{w_i})}{\alpha_i} \exp\left(\frac{1}{2} (x_i - v_i^{(0)})^T F_{w_i}^{-1} (x_i - v_i^{(0)})\right)}
\]

(9)

Where \(v_i\) is the center of the \(i\)-th category; \(\alpha_i\) is the prior probability of selecting cluster \(i\), given by:

\[
\alpha_i = \frac{1}{N} \sum_{k=1}^{N} \mu_{ik}
\]

The membership degrees \(\mu_{ik}\) are interpreted as the posterior probabilities of the \(i\)-th cluster for the data point \(x_i\). \(F_{w_i}\) denotes the fuzzy covariance matrix of the \(i\)-th cluster, given by:

\[
F_{w_i} = \sum_{k=1}^{K} (\mu_{ik})^m (x_i - v_i)(x_i - v_i)^T / \sum_{k=1}^{K} (\mu_{ik})^m, 1 \leq i \leq c
\]

(10)

Note that, being different from the GK algorithm, this distance norm involves an exponential term and thus decreases faster than the inner-product norm.

Meanwhile, Gath and Geva reported that the fuzzy maximum likelihood estimates clustering algorithm is able to detect clusters of varying shapes, sizes and densities, and the clusters are not constrained in volume. However, this algorithm is less robust and it needs a good initialization; besides, because of the introduction of exponential distance norm; it converges to a near local optimum.

4 The experiments and analysis

In this paper, the TM image with 298×298 pixels in Wuhan region acquired in July 2005 is used for testing (Fig.1). The objects in the image are classified into 4 categories by manual judgment which based on the feature color, texture and spatial characteristics: vegetation (trees, lawn); residential area (city buildings); water (lake, river) and bare land (roads, bridges, bareness land). Because of a large amount of data, Principal component analysis (PCA) transformation was involved in this experiment to reduce the dimension before the image classification.

4.1 Principal component analysis (PCA)

Due to the correlation among the wave bands for the image, it is unnecessary to analyze all of the bands. Therefore, in order to improve the training speed, PCA method is adopted for reducing the dimensionality of the data set. Principal component analysis, which is based on the projection of correlated high-dimensional data onto a hyper-plane, involves a mathematical procedure that transforms a larger number of correlated variables into a smaller number of uncorrelated variables. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for the best part of remaining variability[4, 9].

The original TM image with seven bands is compressed into three uncorrelated bands after PCA transformation in the experiment. The result of PCA transformation is shown in Fig.2.

4.2 The remote sensing image classification and analysis

GK and GG algorithms are used to classify the experimental image, and then compare the result with what the traditional FCM algorithm does.

4.2.1 The classification with FCM algorithm

According to the introduction in the ‘Fuzzy c-means (FCM) algorithm’ section, FCM algorithm is a simple iterative process. There are three input parameters needed in Eq.(1): \(1 < c < N\), as the number of categories or initializing partition matrix; \(m > 1\), as the fuzziness weighting exponent; and \(\varepsilon > 0\), as the maximum termination tolerance. In the experiment, the number of categories \(c\) is 4 the weighting exponent \(m\) is 2; the tolerance \(\varepsilon = 1e-6\); and the partition matrix \(U\) is initialized randomly. Repeat the calculations from Eq.(1) to Eq.(4) until the difference between partition matrixes is less than the termination tolerance.

The result is shown in Fig.4. The two bridges across the river in the test area are not classified, and some of the vegetation which stands in the residential area is misclassified into residential area. According to the producer accuracy, the vegetation only accounts for 59.7%.

4.2.2 The classification with GK algorithm

In the iterative process of GK algorithm, the num-
number of categories $c$ is 4; the weighting exponent $m$ is 2; the termination tolerance $\varepsilon$ is 1e-6 and the more one parameter $\rho$ is set to one for each category by 1.

Fig. 5 shows the classification result using the GK algorithm. It can be seen that the some of the vegetation in the residential area could be classified well but some could not. However, the producer accuracy of vegetation increases by about 10% than the FCM algorithm. Meanwhile, the bridge is still unable to classify.

4.2.3 Classification with GG algorithm

The GG algorithm is close relative to the parameters, such as the initialization membership matrix. In order to get a better classification result, the resulting partition matrix from the FCM algorithm is utilized. Other parameters of the GG algorithm are the same as the GK algorithm.

The classification result can be seen in Fig. 6, the bridges can be well classified by GG algorithm, and some sands in the river are also shown. Moreover, the vegetation located in the residential area is also classified, whose producer accuracy is up to 73.6%. The whole effect of the classification is better than the former two algorithms.

4.2.4 Analysis of classification accuracy

As mentioned above, both GK and GG algorithms can improve the effect of classification for a type of non-spherical distribution data. Fig. 7 shows the distribution of experimental data set. It can be seen that the data set is non-spherical distribution.

In order to analyze the classification accuracy, 300 samples are randomly selected for the valuation. The same data set is used for a comparison between the improved methods (GK and GG) and traditional methods (ISODATA and FCM). The classification accuracy is shown in Table 1 and Table 2. From the error
matrix, the total classification accuracy and Kappa coefficient using GK and GG algorithm have increased compared with the traditional FCM algorithm. Moreover, the accuracy of GG algorithm is up to 83.7%.

Table 1  Error matrix

| Methods | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 |
|---------|-----------|-----------|-----------|-----------|
| ISODATA |           |           |           |           |
| Class 1 | 24        | 27        | 21        | 0         |
| Class 2 | 1         | 79        | 25        | 0         |
| Class 3 | 0         | 16        | 50        | 1         |
| Class 4 | 1         | 3         | 8         | 44        |
| Class 1 | 43        | 21        | 5         | 3         |
| Class 2 | 2         | 92        | 11        | 0         |
| Class 3 | 2         | 14        | 50        | 1         |
| Class 4 | 2         | 2         | 1         | 51        |
| Class 1 | 50        | 10        | 9         | 3         |
| Class 2 | 1         | 96        | 7         | 1         |
| Class 3 | 6         | 11        | 49        | 1         |
| Class 4 | 2         | 2         | 1         | 51        |
| Class 1 | 53        | 9         | 9         | 1         |
| Class 2 | 6         | 92        | 7         | 0         |
| Class 3 | 4         | 8         | 54        | 1         |
| Class 4 | 2         | 1         | 1         | 52        |

Table 2  The comparisons of accuracy

| Method | ISODATA | FCM | GK | GG |
|--------|---------|-----|----|----|
| Accuracy% | 65.7     | 78.7 | 82.0 | 83.7 |
| Kappa    | 0.5284   | 0.7062 | 0.7534 | 0.7772 |

5 Conclusions

Aiming at the limitation that traditional FCM algorithm cannot detect categories of different geometrical shapes in one data set, two algorithms (GK and GG) are discussed for improving remote sensing image classification. The experimental result shows that the classification accuracy in the fuzzy partition is higher than traditional hard partition method ISODATA. Furthermore, in the fuzzy partition, the GK and GG algorithm discussed in the paper are more effective than the FCM algorithm for classification of remote sensing images, and the GG algorithm is the best among the three.

References

[1] Wang Cuiling, Wang Yimao (2007) Classification methods of remote sensing image based on FCM clustering [OL]. http://www.paper.edu.cn

[2] Cheng Jian, Lu Kui (2006) The program designing of GK fuzzy clustering based on Matlab[J]. Journal of Anhui University of Science and Technology (Natural Science), (1): 38-40 (in Chinese)

[3] Luo Jiancheng, Yee Leung, Zhou Chenghu (1999) Scale space based hierarchical clustering method and its application to remotely sensed data classification[J]. Acta Geodaetica et Cartographica Sinica, 28(4)

[4] Balazs B, Janos A, Balazs F (2005) Fuzzy clustering and data analysis toolbox for use with Matlab[OL]. http://www.mathworks.com

[5] Zhang Lu, Liao Mingsheng (2006) Contextual fuzzy clustering of remote sensing imagery [J]. Journal of Remote Sensing, 10(1): 58-65

[6] Ge Jianhua, Xiao Haibo, Zhang Min (2007) The FCM data clustering analysis by Matlab[J]. Fujian computer, 4:89-90(in Chinese)

[7] Yang Yong, Zheng Chongxun, Lin Pan, et al (2005) A new algorithm for image segmentation based on modified fuzzy c-Means Clustering [J]. Journal of Optoelectronics Laser, 16(9): 1118-1122 (in Chinese)

[8] Pan Wei, Fu Jia (2007) A automatic classification weighted fuzzy c-means image segmentation algorithm [J]. Periodical of Ocean University of China, 37(3): 485-488 (in Chinese)

[9] Zhao Yingshi (2006) The principles and methods of remote sensing application [M]. Beijing: Science Publishing Company (in Chinese)

[10] Beichel R, Bolter R, Pinz A (1999) Fuzzy clustering of a landsat TM Scene[C]. Geoscience and Remote Sensing Symposium, Hamburg