The following publication H. Zhang, Q. Wang, W. Shi and M. Hao, "A Novel Adaptive Fuzzy Local Information C -Means Clustering Algorithm for Remotely Sensed Imagery Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 9, pp. 5057-5068, Sept. 2017 is available at https://doi.org/10.1109/TGRS.2017.2702061.

A Novel Adaptive Fuzzy Local Information *C*-Means Clustering Algorithm for Remotely Sensed Imagery Classification

Hua Zhang, Qunming Wang, *Member, IEEE*, Wenzhong Shi, and Ming Hao

***Abstract*— This paper presents a novel Adaptive Fuzzy Local Information *C*-Means (ADFLICM) clustering approach for remotely sensed imagery classification by incorporating the local spatial and gray level information constraints. The ADFLICM approach can enhance the conventional Fuzzy *C*-Means (FCM) algorithm by producing homogeneous segmentation and reducing the edge blurring artifact simultaneously. The major contribution of ADFLICM is use of the new fuzzy local similarity measure based on pixel spatial attraction model, which adaptively determines the weighting factors for neighboring pixel effects without any experimentally set parameters. The weighting factor for each neighborhood is fully adaptive to the image content, and the balance between insensitiveness to noise and reduction of edge blurring artifact to preserve image details is automatically achieved by using the new fuzzy local similarity measure. Four different types of images were used in the experiments to examine the performance of ADFLICM. The experimental results indicate that ADFLICM produces greater accuracy than the other four methods and hence provides an effective clustering algorithm for classification of remotely sensed imagery.**

***Index Terms*—Classification, fuzzy *c*-means clustering, spatial information, local measure similarity, remotely sensed imagery**

I. Introduction

Extracting land cover information from remotely sensed imagery is a common topic in remote sensing and is usually accomplished by classification. When training data is unavailable, unsupervised clustering is widely used for classification of remotely sensed imagery [1]. Many clustering algorithms, such as *K*-means [2], Expectation–Maximum (EM) [3], ISODATA [4], *K*-Nearest-Neighbor (KNN) [5], Markov Random Field (MRF) [6], Fuzzy *C*-Means (FCM) [7], and their variations have been exploited for unsupervised classification. Amongst them, FCM [8], [9] is one of the most commonly used methods. However, due to limited spatial resolution, complexity of ground substances, diversity of disturbance or spectral variation, conventional FCM often produces clustering maps containing salt and pepper noise.

Recently, many researchers have incorporated local spatial information into conventional FCM to enhance the clustering performance [10]-[21]. One of the most commonly used methods is to modify the conventional FCM objective function to include the spatial constraints [10]-[16], [22]-[26]. For example, Pham [23] proposed a Robust Fuzzy *C*-Means algorithm (RFCM) that extended the conventional FCM by including a spatial penalty term. Ahmed *et al.* [10], [22] proposed a FCM_S method by introducing the spatial neighborhood term to the FCM objective function. One shortcoming of FCM_S is time-consuming. To reduce the computational complexity of FCM_S, Chen and Zhang [24] developed two variants, FCM_S1 and FCM_S2 to simplify the computation of neighborhood term. To further accelerate the clustering process, Enhanced FCM (EnFCM) [25] and Fast Generalized FCM (FGFCM) [11] were developed. However, these extended FCM algorithms perform indirectly on the original image, or need a crucial parameter to control the trade-off between the robustness to noise and the effectiveness of preserving the image details, and the selection of these parameters is difficult [10]-[13]. To overcome these problems, Krinidis and Chatzis [13] presented a Fuzzy Local Information *C*-Means (FLICM). However, this method has some weakness in identifying the class boundary pixels and preserving image details [14]. To produce more robust results, Gong *et al* [14] proposed a Reformulated FLICM (RFLICM) which introduces a local coefficient of variation to replace the spatial distance as the local similarity measure. Li *et al.* [16] proposed a Fuzzy *C*-Means with Edge and Local Information (FELICM) based on FLICM to reduce edge degradation. In FLICM and abovementioned enhanced FLICM algorithms, the identification of the center pixel is greatly influenced by its neighboring pixels while the center pixel's own features are not fully considered, failing to take full advantage of the local information encapsulated in the local window. Thus, they may produce over-smooth results for important structures (such as regional borders or edges) and small patches.

To address the aforementioned problems, this paper presents a novel Adaptive Fuzzy Local Information *C*-Means (ADFLICM) clustering approach for remotely sensed imagery classification. In ADFLICM, a novel fuzzy local similarity

Manuscript received April 21, 2016; revised June 10, 2016; November 16, 2016; accepted May 4, 2017. This work was supported by the Fundamental Research Funds for the Central Universities under Grant 2015XKQY09 and the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions. (*Corresponding author: Qunming Wang*.)

H. Zhang and M. Hao are with the School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou, 221116, China. (e-mail: zhhua_79@163.com; haomingcumt@gmail.com).

Q.M. Wang is with the Lancaster Environment Centre, Lancaster University, UK. (e-mail: wqm11111@126.com).

W.Z. Shi is with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. (e-mail: lswzshi@polyu.edu.hk).

measure is defined to replace the fixed parameter α in FCM_S. The new fuzzy local similarity measure S_{ir} possesses several characteristics and advantages: 1) S_{ir} uses a pixel spatial attraction model to describe the relationships between pixels; 2) S_{ir} can be automatically determined by local spatial and gray level relationships between the center pixel and its neighboring pixels in a local window, and it is adaptive to the local image context without any artificial or empirical selection; 3) using S_{ir}, the clustered pixel is influenced by its neighboring pixels and its own features simultaneously, which is useful for retaining edges of regions and small patches when removing noise; 4) S_{ir} makes the proposed algorithm relatively independent of the noise type, making it a promising choice for clustering in the absence of prior knowledge on noise.

The rest of the paper is organized as follows. Section II briefly describes FCM clustering algorithm with spatial constrains and its variants, and the FLICM algorithm. In Section III, the proposed ADFLICM algorithm is introduced explicitly. Section IV illustrates the performance of the proposed algorithm through four experiments. Section V finally concludes the paper.

II. PRELIMINARY THEORY

Supposing an image $X=\{x_1,x_2,\cdots x_i,\cdots,x_N\}$, $x_i\in \mathrm{R}^{\mathrm{n}}$, is a dataset in the $n-$dimensional vector space, N is the number of feature vectors (pixel number in the image), and c is the number of clusters ($2\le c<N$).

A. Fuzzy Clustering with Spatial Constrains (FCM_S)

To enhance the robustness of conventional FCM, Ahmed *et al*. [10] introduced a new term that allows the label of a pixel to be influenced by labels of its neighbors. The neighborhood effect acts as a regularizer and pushes the solution toward piecewise-homogeneous labeling. The objective function of FCM_S is defined as follows:

$$J_m=\sum_{i=1}^{N}\sum_{k=1}^{c}u_{ki}^{m}\|x_i-v_k\|^2+\frac{\alpha}{N_R}\sum_{i=1}^{N}\sum_{k=1}^{c}u_{ki}^{m}\sum_{r\in N_i}\|x_r-v_k\|^2 \quad (1)$$

where x_i is the gray value of the ith pixel, v_k denotes prototype value of kth cluster, u_{ki} represents the degree of fuzzy membership of x_i belonging to the kth cluster, m is the weighing exponent for each fuzzy membership, N_R is its cardinality, N_i is the set of neighborhood pixels in the widow around the ith pixel x_i, and pixel x_r ($r\in N_i$) is the neighborhood pixel that falls into N_i. The parameter α controls the effect from the neighborhood. The details of FCM_S can be found in [10].

A shortcoming of FCM_S is that the computation of the neighborhood term is time-consuming. To reduce the computation burden, Chen and Zhang [24] proposed the FCM_S1 method, in which the neighborhood term in FCM_S is simplified. The modified objective function is written as follows:

$$J_m=\sum_{i=1}^{N}\sum_{k=1}^{c}u_{ki}^{m}\|x_i-v_k\|^2+\alpha\sum_{i=1}^{N}\sum_{k=1}^{c}u_{ki}^{m}\|\overline{x_i}-v_k\|^2 \quad (2)$$

where $\overline{x_i}$ is the mean of neighboring pixels within a local window around x_i. However, FCM_S1 may not be suitable for images with impulse noise [24]. To this end, a variant of FCM_S1, that is, FCM_S2, was proposed by Chen and Zhang [24], in which the median-filtered image is used to replace the mean-filtered image to enhance the robustness to impulse noise like salt and pepper noise.

B. Fuzzy Local Information C-Means Clustering Algorithm

In the objective functions of FCM_S, FCM_S1 and FCM_S2, the parameter α balances the robustness to noise and the effectiveness of preserving the image details, and it has a crucial impact on the final clustering performance. Its selection is difficult when there is no prior knowledge on the noise. In practice, it is generally determined empirically [24]. Moreover, α is fixed for all neighbor windows across the whole image and the local gray level or spatial information may be overlooked. Furthermore, in FCM_S1 and FCM_S2, using the filtered image may lead to the loss of details of the original image.

To overcome the abovementioned limitations, FLICM [13] introduces a novel fuzzy factor G_{ki} as a local similarity measure to remove noise and preserve image details simultaneously.

$$G_{ki}=\sum_{\substack{j\in N_i\\ i\neq j}}\frac{1}{d_{ij}+1}(1-u_{kj})^m\|x_j-v_k\|^2 \quad (3)$$

where d_{ij} is the spatial Euclidean distance between pixels i and j. Introducing G_{ki} to the conventional FCM, the objective function of FLICM is described as:

$$J_m=\sum_{i=1}^{N}\sum_{k=1}^{c}\left[u_{ki}^{m}\|x_i-v_k\|^2+G_{ki}\right] \quad (4)$$

where u_{ki} and v_k are defined as [13].

However, FLICM has limitations in identifying the class boundary pixels and the important structures (such as regional borders or edges) may be over-smoothed [14].

III. ADAPTIVE FUZZY LOCAL INFORMATION C-MEANS CLUSTERING

Motivated by the individual strengths as well as limitations of FCM_S, FCM_S1, FCM_S2 and FLICM, this paper presents a novel Adaptive Fuzzy Local Information *C*-Means (ADFLICM) clustering algorithm. In this section, Section III-A gives an definition of pixel spatial attraction between pixels, and Section III-B describes the proposed local similarity measure. The general framework of ADFLICM is provided in Section III-C.

A. The definition of pixel spatial attraction model

The attraction model was shown to be effective in characterizing the spatial correlation between pixels in the image [27], [28]. In this paper, we generalize the attraction model to incorporate local spatial and gray level information. For two pixels i and j, their attraction with respect to the kth cluster is proportional to their fuzzy memberships u_{ki} and u_{kj}, and inversely proportional to the square of the spatial distance between the two pixels. Accordingly, the pixel spatial attraction $SA_{ij}(k)$ between the two pixels can be described as:

$$SA_{ij}(k) = \frac{u_{ki} \times u_{kj}}{D_{ij}^2} \tag{5}$$

where D_{ij} is a spatial distance between pixels i and j. In this paper, the Chebyshev distance is selected for D_{ij}.

B. A novel local similarity measure S_{ir}

The proposed similarity measure aims to address the following issues: 1) to provide a proper trade-off between the insensitiveness to noise and effectiveness of preserving the details in the image; 2) the trade-off should be determined automatically without any manual parameter selection; 3) the value should change flexibly according to the spatial distances from the center pixel and the gray level differences simultaneously.

Based on the spatial attraction model, we introduce a novel local similarity measure S_{ir} which incorporates both local spatial and gray level information. It is defined as:

$$S_{ir} = \begin{cases} SA_{ir}, & i \neq r \\ 0, & i = r \end{cases} \tag{6}$$

where the ith pixel is the center of local window, the rth pixel ($r \in N_i$) is the neighborhood pixel that falls into N_i. The neighborhood structure of the local window is defined as:

$$N_i = \{r \in N \mid 0 < (a_i - a_r)^2 + (b_i - b_r)^2 \leq Q\} \tag{7}$$

where (a_i, b_i) and (a_r, b_r) denote the coordinates of pixels i and r, respectively, and Q is a constant equal to 2^{L-1} (L is the level of neighborhood). Fig. 1 illustrates the neighborhood structure for different levels. Note that the attractions only exist between the center pixel and its neighboring pixels in the given local window and other pixels outside of the window are assumed too distant to exert any attractions on the center pixels. The distance D_{ir} between pixels i and r can be defined as the Chebyshev distance $D_{ir} = \max(|a_i - a_r|, |b_i - b_r|)$. In addition, other distance metrics such as Euclidean distance and Manhattan distance can also be favorably adopted in our algorithm. It is also worth noting that the shape of local window is not restricted to that in Fig. 1, but also other shapes such as square can also be adopted in our algorithm.

Clearly, the local similarity measure S_{ir} does not involve any experimentally adjusted parameters (except the local window level L) to control the trade-off between the image noise removal and the image details preservation. The weighting factor for the neighborhood effect is automatically determined by the spatial attraction. The introduction of the measure S_{ir} makes the influences on the center pixels from its neighbors change adaptively according to their distance D_{ir}. This is different from the parameter α in FCM_S, FCM_S1 and FCM_S2 that is globally taken as a constant. Moreover, based on S_{ir}, the weighting factor is not only influenced by its neighboring pixels but also the central pixel. In such a way, when the gray level value of the rth neighboring pixel is close to the gray level value of center pixel i, the center pixel should be greatly influenced by this neighboring pixel, and thus, S_{ir} should be large and *vice versa* (*e.g.*, the ith pixel locates on the edge region). This is helpful for reducing the edge blurring artifact.

				5				
		5	5	5	5	5		
	5	4	4	3	4	4	5	
	5	4	2	1	2	4	5	
5	5	3	1	x_i	1	3	5	5
	5	4	2	1	2	4	5	
	5	4	4	3	4	4	5	
		5	5	5	5	5		
				5				

Fig. 1. Definition of neighborhood structure of the local window (a higher level includes pixels labeled as the number of the level and pixels in all lower levels)

C. General framework of ADFLICM

Based on S_{ir}, ADFLICM is proposed for unsupervised remotely sensed imagery classification. It incorporates local spatial and gray level information into the objective function of conventional FCM to enhance the smoothness towards piecewise-homogeneous classification and reduce the edge blurring effect simultaneously, the objective function of ADFLICM is described as:

$$J_m = \sum_{i=1}^{N}\sum_{k=1}^{c} u_{ki}^m \left[\| x_i - v_k \|^2 + \frac{1}{N_R} \sum_{\substack{r \in N_i \\ r \notin i}} (1 - S_{ir}) \| x_r - v_k \|^2 \right] \tag{8}$$

where, N_i is defined in Section III-B. The two necessary conditions of J_m to be at its local minimal extreme, with respect to u_{ki} and v_k are obtained as follows:

$$\boldsymbol{v}_k = \frac{\sum_{i=1}^{N} u_{ki}^m (\boldsymbol{x}_i + \frac{1}{N_R} \sum_{\substack{r \in N_i \\ r \notin i}} (1 - S_{ir}) * x_r)}{(1 + \frac{1}{N_R} \sum_{\substack{r \in N_i \\ r \notin i}} (1 - S_{ir})) \sum_{i=1}^{N} u_{ki}^m} \tag{9}$$

$$u_{ki} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\| x_i - v_k \|^2 + \frac{1}{N_R} \sum_{\substack{r \in N_i \\ r \notin i}} (1 - S_{ir}) * \| x_r - v_k \|^2)}{\| x_i - v_j \|^2 + \frac{1}{N_R} \sum_{\substack{r \in N_i \\ r \notin i}} (1 - S_{ir}) * \| x_r - v_j \|^2} \right)^{1/(m-1)}} \tag{10}$$

The flowchart of proposed method is shown in Fig. 2. The implementation includes the following five steps:

Step 1- Initialization

Set the cluster number c , the weighting exponent m, local window level L and termination criterion ε and the loop counter $b=0$. The standard FCM is implemented to obtain the final fuzzy memberships matrix $U=\{u_{ki}\}_{c\times N}$ as the initial memberships matrix of ADFLICM.

Step 2- Calculating the similarity measure S_{ir}

S_{ir} is calculated according to Equations (5) and (6).

Step 3- Calculating the cluster centers and membership values

Based on the S_{ir} obtained in Step 2, the cluster centers are calculated by Equation (9) and the membership values are calculated by Equation (10).

Step 4- Termination

The iteration will stop when the termination criterion $\max_{k\in[1,c]}\{\|v_k^b - v_k^{(b+1)}\|\} < \varepsilon$ is met; otherwise, $b=b+1$, go back to Step 2 and repeat.

Step 5- assigning the final class to each pixel

After the algorithm converges, the final fuzzy matrix $U=\{u_{ki}\}_{c\times N}$ is produced, and the crisp partition is performed finally by assigning each pixel i to the class c with the greatest membership:

$$C_i = \arg_k\{\max(u_{ki})\},\ k=1,2,3,\cdots,c \tag{11}$$

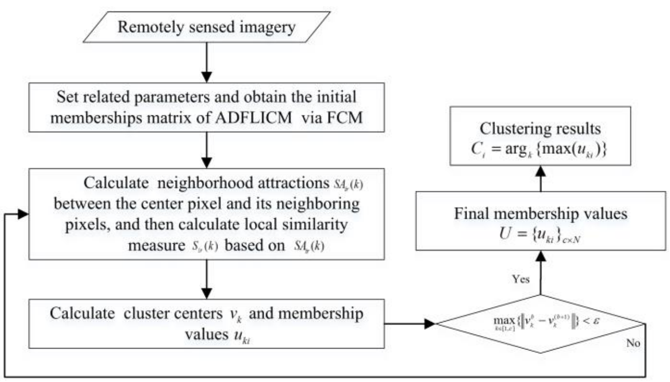
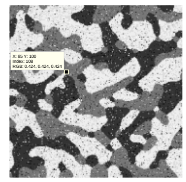


Fig. 2. Flowchart of the proposed ADFLICM clustering approach.

The ADFLICM is robust because of the introduction of the novel local similarity measure. Specifically, as seen from Equation (8), the noise tolerance and detail preservation ability is completely dependent on the local spatial and gray level information from neighbors that is characterized by S_{ir} which is determined automatically. Three basic cases are used to describe the robustness to noise and detail preservation ability of the ADFLICM algorithm. In this section, a synthetic image was used to evaluate the anti-noise performance of the proposed method. The synthetic image with 256×256 pixels included three classes (1 2 3) with three intensity values taken as (55 110 225) was sampled from MRF model using a Gibbs sampler. Fig. 3 (a) and Fig. 4 (a) are the same image corrupted by 'pepper & salt' noise with a level of 3%, and Fig. 5 (a) is the image corrupted by the 'Gaussian' noise (the mean and variance are 0 and 0.01, respectively).

Case 1: The center pixel is not noisy and some of its neighboring pixels are contaminated by noise. A 3×3 window for a pixel marked in Fig. 3 (a) is shown in Fig.3 (b). ADFLICM converged after 15 iterations. As shown in Fig 3 (b), the noisy pixels' gray level values are 8 and 254, which are far different from the center pixel with a value of 108. The local similarity measure S_{ir} balances their membership values and suppresses the influences from noise of the neighbors. Hence, S_{ir} can enhance the ADFLICM's robustness to the noise in this case.



(a)

104	254	126
124	108	8
124	119	114

(b)

0.1113	0.4456	0.3774
0.4845	0.4070	0.3921
0.2202	0.5635	0.1710

(c)

0.3745	0.4291	0.2740
0.1918	0.3988	0.0035
0.2258	0.2840	0.3785

(d)

0.5143	0.1253	0.3485
0.3237	0.1943	0.6045
0.5540	0.1526	0.4505

(e)

0.0103	0.1474	0.0175
0.0078	0.0148	0.2017
0.0228	0.0037	0.0025

(f)

0.9852	0.7051	0.9748
0.9880	0.9791	0.7395
0.9649	0.9953	0.9969

(g)

0.0045	0.1475	0.0077
0.0042	0.0060	0.0588
0.0123	0.0010	0.0006

(h)

143.6344
143.4637
143.9878

(i)

55.7381
115.1681
223.2817

(j)

2	2	2
2	2	2
2	2	2

(k)

2	2	2
2	2	2
2	2	2

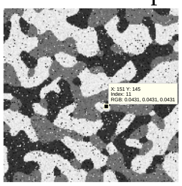
(l)

Fig. 3. Classification results of a synthetic image with noise using ADFLICM (Case 1): (a) is the original image, (b) is a 3×3 window (marked with a rectanguler in the original image, and the center pixel locates at pixel (85, 100)), (c), (d) and (e) show the initial membership degrees of the pixels belonging to the three class, respectively, (f), (g) and (h) show the final membership degrees of the pixels belonging to the three classes after 15 iterations, respectively, (i) shows the intial cluster centers, (j) shows the final cluster centers, (k) shows the classification results using ADFLICM, and (l) shows the reference of the classification results.

Case 2: The center pixel is noisy and the pixels within its local window are not contaminated and homogeneous. ADFLICM in this example converged after 15 iterations. As shown in Fig. 4 (b), while the gray level value of the center pixel is 11, which is far different from those of the neighbors. The local similarity S_{ir} pushes the membership value of the center pixel to the same as that for the non-noisy neighbors.

Case 3: In Cases 1 and 2, pixels in the given local window are homogeneous. In fact, there are also a number of pixels locate on the object boundaries, as illustrated in Case 3. The ADFLICM algorithm converged after 21 iterations in this case. Fig. 5 (b) shows the pixels of two regions. The gray level value of the center pixel is 72, and its neighborhood pixels' values are far different. By incorporation of spatial and gray level information, the neighborhood pixels will greatly influence the center pixel, and improper weighting will result in misclassification of the center pixel. Here, to validate the advantage of ADFLICM, the classification result of FLICM is shown for comparison, see Fig. 6. In FLICM result, the center pixel is misclassified as the second class as G_{ki} cannot properly reflect the damping extent of the neighbors, as discussed in

Section II-B. In ADFLICM, however, the local similarity measure S_{ir} not only incorporates the neighborhood pixels' spatial and gray level information, but also the center pixel's own gray level information. The center pixel's information can alleviate the influence from the neighbors to some extent. As shown in the classification result of ADFLICM, the center pixel is assigned with the correct class, suggesting that the proposed method can keep a balance between the insensitiveness to noise and effectiveness of preserving the details in this case.



(a)

220	223	216
225	11	214
227	215	226

(b)

0.1113	0.4456	0.3774
0.4845	0.4070	0.3921
0.2202	0.5635	0.1710

(c)

0.5143	0.1253	0.3485
0.3237	0.1943	0.6045
0.5540	0.1526	0.4505

(d)

0.3745	0.4291	0.2740
0.1918	0.3988	0.0035
0.2258	0.2840	0.3785

(e)

0.0103	0.1474	0.0175
0.0078	0.0148	0.2017
0.0228	0.0037	0.0025

(f)

0.0045	0.1475	0.0077
0.0042	0.0060	0.0588
0.0123	0.0010	0.0006

(g)

0.9852	0.7051	0.9748
0.9880	0.9791	0.7395
0.9649	0.9953	0.9969

(h)

143.6344
143.4637
143.9878

(i)

55.7381
115.1681
223.2817

(j)

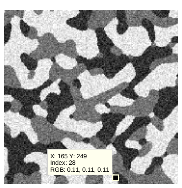
3	3	3
3	3	3
3	3	3

(k)

3	3	3
3	3	3
3	3	3

(l)

Fig. 4. Classification results of a synthetic image with noise based on ADFLICM (Case 2): (a) is the original image, (b) is a 3×3 window (marked with a rectangular in the original image, and the center pixel locates at pixel (151, 145)), (c), (d) and (e) show the initial membership degrees of the pixels belonging to the three class, respectively, (f), (g) and (h) show the final membership degrees of the pixels belonging to the three classes after 15 iterations, respectively, (i) shows the intial cluster centers, (j) shows the final cluster centers, (k) shows the classification results using ADFLICM, and (l) shows the reference.



(a)

233	42	63
245	72	59
195	212	47

(b)

0.1237	0.3641	0.3503
0.4122	0.4539	0.4466
0.3053	0.2685	0.0503

(c)

0.7319	0.3454	0.1702
0.1376	0.0429	0.3846
0.1984	0.4068	0.5324

(d)

0.1444	0.2905	0.4794
0.4502	0.5032	0.1688
0.4963	0.3247	0.4173

(e)

0.0994	0.5642	0.9888
0.0469	0.4424	0.8650
0.0170	0.1044	0.5531

(f)

0.1682	0.3091	0.0100
0.0929	0.3956	0.1192
0.0397	0.1884	0.3373

(g)

0.7324	0.1266	0.0012
0.8602	0.1620	0.0158
0.9433	0.7072	0.1097

(h)

143.4187
144.077
144.3141

(i)

54.6274
115.5244
223.5646

(j)

3	1	1
3	1	1
3	3	1

(k)

3	1	1
3	1	1
3	3	1

(l)

Fig. 5. Classification results of a synthetic image with noise based on ADFLICM (Case 3): (a) is the original image, (b) is a 3×3 window (marked with a rectanguler in the original image, and the center pixel locates at pixel (165, 249)), (c), (d) and (e) show the initial membership degrees of the pixels belonging to the three class, respectively, (f), (g) and (h) show the final membership degrees of the pixels belong to the three classes after 21 iterations, respectively, (i) shows the intial cluster centers, (j) shows the final cluster centers, (k) shows the classification results using ADFLICM, and (l) shows the reference.

0.1844	0.4531	0.8900
0.1369	0.3757	0.6932
0.0844	0.1930	0.4310

(a)

0.2645	0.3594	0.0861
0.2097	0.4016	0.2371
0.1368	0.2899	0.3806

(b)

0.5511	0.1875	0.0240
0.6534	0.2227	0.0697
0.7788	0.5171	0.1884

(c)

143.4187
144.077
144.3141

(d)

57.1634
115.9147
222.9790

(e)

3	1	1
3	2	1
3	3	1

(f)

3	1	1
3	1	1
3	3	1

(g)

Fig. 6. Classification results of the synthetic image in Fig. 5 based on FLICM: (a)-(c) show the final membership degrees of the 3×3 pixels in Fig. 5(b) belonging to the three classes after 25 iterations, respectively, (d) shows the intial cluster centers, (e) shows the final cluster centers, (f) shows the classification results using FLICM, and (g) shows the reference.

Fig. 7 (a)-(c) show the reference image, and classification results of ADFLCIM and FLICM based on the synthetic image with 'pepper & salt' noise. As shown in Fig. 7, the performances of (b) and (c) are very similar. To evaluate the performances quantitatively, the Producer's accuracy, Overall Accuracy and Kappa coefficient were listed in Table I. As seen from Table I, ADFLICM produces a little greater classification accuracy than FLICM. It shows that both AFLICM and FLICM are robust to the noise in the homogenous regions. Fig. 8 (a)-(c) show the reference image, classification result of ADFLCIM and FLICM based on the synthetic image with 'Gaussian' noise. Checking the results in Fig. 8, AFLICM outperforms FLICM, especially for classification of the boundaries. The quantitative assessment in Table II also shows that AFLICM outperforms FLICM.

TABLE I

COMPARISON OF PRODUCER'S ACCURACY, OVERALL ACCURACY AND KAPPA COEFFICIENT OF ADFLICM AND FLICM BASED ON A SYNTHETIC IMAGE WITH 'pepper & salt' NOISE

Class	Number of Testing samples	ADFLCIM	FLICM
Class 1	19145	99.69%	99.22%
Class 2	18360	99.76%	99.78%
Class 3	28031	99.83%	99.69%
Overall Accuracy		99.77%	99.58%
Kappa Coefficient		0.9965	0.9935

TABLE II

COMPARISON OF PRODUCER'S ACCURACY, OVERALL ACCURACY AND KAPPA COEFFICIENT OF ADFLICM AND FLICM BASED ON A SYNTHETIC IMAGE WITH 'Gaussian' NOISE

Class	Number of Testing samples	ADFLCIM	FLICM
Class 1	19145	99.66%	97.85%
Class 2	18360	99.84%	99.48%
Class 3	28031	99.89%	99.44%
Overall Accuracy		99.81%	98.99%
Kappa Coefficient		0.9970	0.9845

TABLE III

COMPARISON OF CLUSTER VALIDITY VALUES FOR ADFLICM AND FLICM

	V_{PC}	V_{PE}	V_{MPC}	V_{FS}	V_{XB}	V_{K}	V_{T}	V_{PCAES}
ADFLICM	0.907	0.182	0.861	-2.748E+08	0.025	1.678E+03	1.678E+03	2.575
FLICM	0.761	0.456	0.641	-2.196E+08	0.026	1.726E+03	1.726E+03	2.663

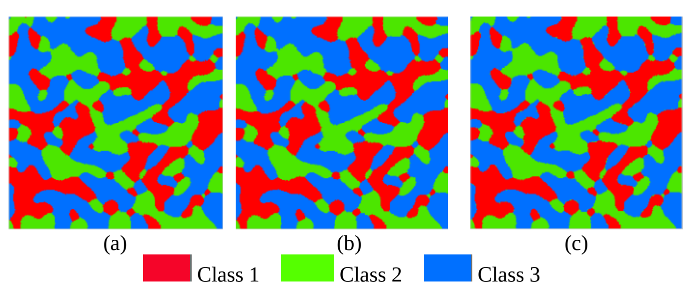


Fig. 7. Classification results for a synthetic image with 'pepper & salt' noise using ADFLICM and FLICM: (a) is the reference image, (b) and (c) are classification results of ADFLICM and FLICM, respectively.

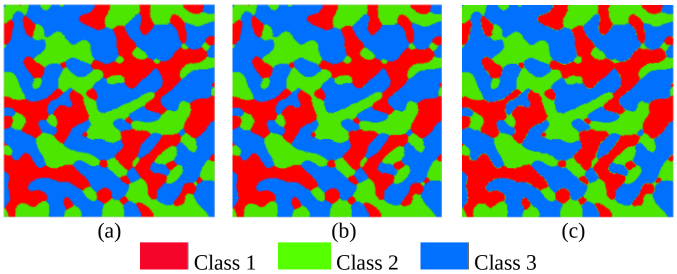


Fig. 8. Classification results on a synthetic image with 'Gaussian' noise using ADFLICM and FLICM: (a) is the reference image, (b) and (c) are corresponding classification results by ADFLICM and FLICM, respectively.

In addition, Partition Coefficient (PC) [29], Partition Entropy (PE) [30], Modification of the Partition Entropy (MPC) [31], Fukuyama and Sugeno (FS) [32], Xie-Beni (XB) [33], Kwon (K) [34], Tang (T) [35] and Partition Coefficient and Exponential Separation (PCAES) [36] indices were used to quantitatively evaluate the cluster validity of the ADFLICM. Table III lists the comparison results. As seen from the table, all indices show that AFLICM outperforms FLICM in terms of cluster validity.

It is worth pointing out that the way to reduce noise in ADFLICM is different from that in other methods. FCM_S1 is relatively suitable for the noisy image corrupted by Gaussian noise by using the mean-type filtering, while FCM_S2 is relatively suitable for the noise image corrupted by impulse noise such as salt and pepper noise by using the median-type filtering. Furthermore, their final clustering results are affected by parameter α, and it is difficult to obtain the optimal one without any prior knowledge on the noise. The FCM_S algorithm uses the original image without any preprocessing steps, but the fixed value of α for all neighboring pixels usually overlooks the local information and α also needs extra work to be determined. FLICM is independent of noise type and free of any parameter choosing, but it may result in over-smoothed borders. The main characteristics of ADFLICM are summarized as following:

1) Using the spatial attraction model, the weighting factors are determined by both center pixels and its neighboring pixels simultaneously;

2) It is less sensitive to noise and is able to reduce the edge blurring artifact when removing isolated pixels;

3) It is free of parameter selection when incorporating the local spatial and gray level information in a given local window;

4) The classification is performed straightforwardly on original image to preserve image details without any preprocessing steps to generate the filtered image.

IV. Experimental Study and Analysis

In this section, the performance of ADFLICM was examined and compared with four fuzzy algorithms (i.e., standard FCM, FCM_S1, FCM_S2 and FLICM) through four experiments. Each algorithm was conducted ten trials and the average classification accuracy and the optimal classification results were provided. All algorithms were implemented with Matlab 2013b. The Producer's Accuracy, Overall Accuracy and Kappa coefficient were used to quantitatively evaluate the classification performance.

A. Experiment 1: TM Image of Xuzhou

In this experiment, ADFLICM was tested using a 30-m resolution multispectral Landsat Thematic Mapper (TM) image (272×165 pixels) acquired on September 14, 2000 (Fig. 9 (a)). The studied area located in Xuzhou City, China. Bands 1, 2, 3, 4, 5 and 7 were used for image classification. The area contains building and bare soil, woodland, water and farmland (Fig. 9 (b)), respectively. The testing samples in the reference image were obtained by referring to the TM image and land use map. Specifically, a 1: 2000 land use map which was produced around the same date as the TM image. They were georeferenced to the same coordinate system as the land use map. Based on the land use map, TM image and fieldwork, some reliable test sample points were selected to generate the reference map. The parameters in five algorithms are: $c=4$, $m=2$, $\varepsilon=1e-5$, $L=2$, and $N_R=8$. The parameter α was set to 4.3 in FCM_S1 and FCM_S2 by repeating test in the interval [0.2, 8].

Fig. 9 (c)-(g) illustrates the classification results derived from the FCM, FCM_S1, FCM_S2, FLICM, and ADFLICM algorithms, respectively. As shown in Fig. 9 (c), due to mixed pixels in TM image and spectral variation, FCM produces a map with salt and pepper noise and shows the weakest performance amongst the five algorithms. With the local spatial information and gray level information are incorporated into the objective function, FCM_S1, FCM_S2, FLICM and ADFLICM produce more homogeneous images. In Fig. 9 (d) and Fig. 9 (e), most of the isolated pixels are removed by FCM_S1 and FCM_S2, but some isolated pixels still remain. As shown in Fig. 9 (f) and Fig. 9 (g), FLICM and ADFLICM are more competent in removing isolated pixels. FLICM removes almost all the isolated pixels and achieves satisfactory result, while some image details are lost. In ADFLICM result, most of the isolated pixels are removed and image details are satisfactorily preserved. For example, in marked area A, many woodland pixels are misclassified as water pixels by FCM, FCM_S1, FCM_S2 and FLICM, and ADFLICM produces more accurate result. In marked area B, many farmland pixels are misclassified into woodland pixels by FLICM, while in ADFLICM result, most of the pixels are correctly classified into farmland. The advantage of ADFLICM can be similarly illustrated by comparison for marked areas C and D. In area C, FLICM obtain more homogeneous classification result than ADFLICM. However, area C covers woodland with some farmland pixels (i.e., heterogeneous). Compared with FLICM, ADFLICM correctly produces more details. This experiment also illustrates that ADFLICM has advantages in classification of both homogeneous and heterogeneous landscapes.

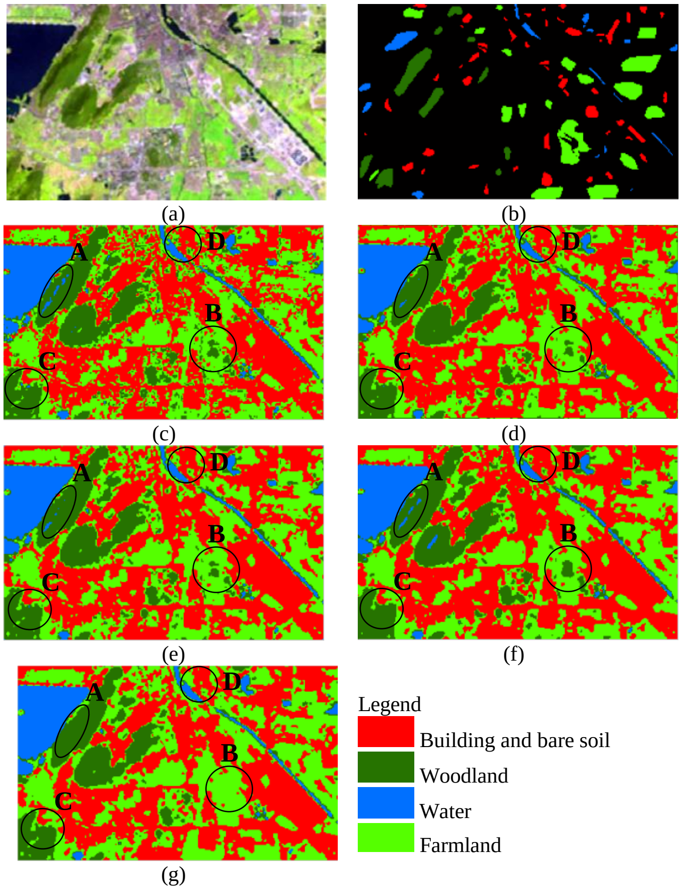


Fig. 9. Dataset and classification results in Experiment 1. (a) is the TM image of Xuzhou (RGB (5, 4, 3)), (b) is the reference image, (c), (d), (e), (f) and (g) are classification results of Fig. 9 (a) produced by FCM, FCM_S1, FCM_S2, FLICM, and ADFLICM, respectively.

TABLE IV
COMPARISON OF PRODUCER'S ACCURACY, OVERALL ACCURACY AND KAPPA COEFFICIENT OF FIVE CLASSIFICATION METHODS IN EXPERIMENT 1

Class	Number of Testing samples	FCM	FCM_S1	FCM_S2	FLICM	ADFLICM
Building and bare soil	1275	93.02%	93.49%	94.59%	**96.24%**	92.47%
Woodland	1186	93.76%	98.82%	97.64%	92.50%	**99.92%**
Water	627	**98.72%**	95.53%	98.25%	98.41%	98.09%
Farmland	2647	80.36%	83.42%	84.81%	83.26%	**92.14%**
Overall Accuracy		87.95%	90.17%	91.11%	89.71%	**94.47%**
Kappa Coefficient		0.8285	0.8592	0.8725	0.8534	**0.9196**

The quantitative results are listed in Table IV. As seen from Table IV, FCM_S1, FCM_S2, FLICM and ADFLICM yield greater classification accuracies than FCM. Amongst all methods, ADFLICM obtains the greatest accuracy. Taking the Overall Accuracy as an example, ADFLICM produces a value of 94.47%, with gains of 6.52%, 4.30%, 3.36% and 4.76% over FCM, FCM_S1, FCM_S2 and FLICM, respectively.

B. Experiment 2: ZY-3 Image of Xuzhou

In this experiment, a 6-m resolution multispectral ZY-3 image (400×400 pixels) was used for validation. It is located in Xuzhou City, China, and was acquired on August 11, 2012 (Fig. 10 (a)). The four multispectral bands of the ZY-3 image were used for image classification. The reference image contains building and bare soil, greenhouse, water and vegetation (Fig. 10 (b)). Based on the panchromatic imagery with 2-meter resolution, the multispectral imagery and the fieldwork, the testing samples were obtained, as shown in Fig. 10 (b). The parameters used in this experiment are: $c=4$, $m=2$, $\varepsilon=1e-5$, $L=2$, $N_R=8$ and $\alpha=3.9$ (in FCM_S1 and FCM_S2).

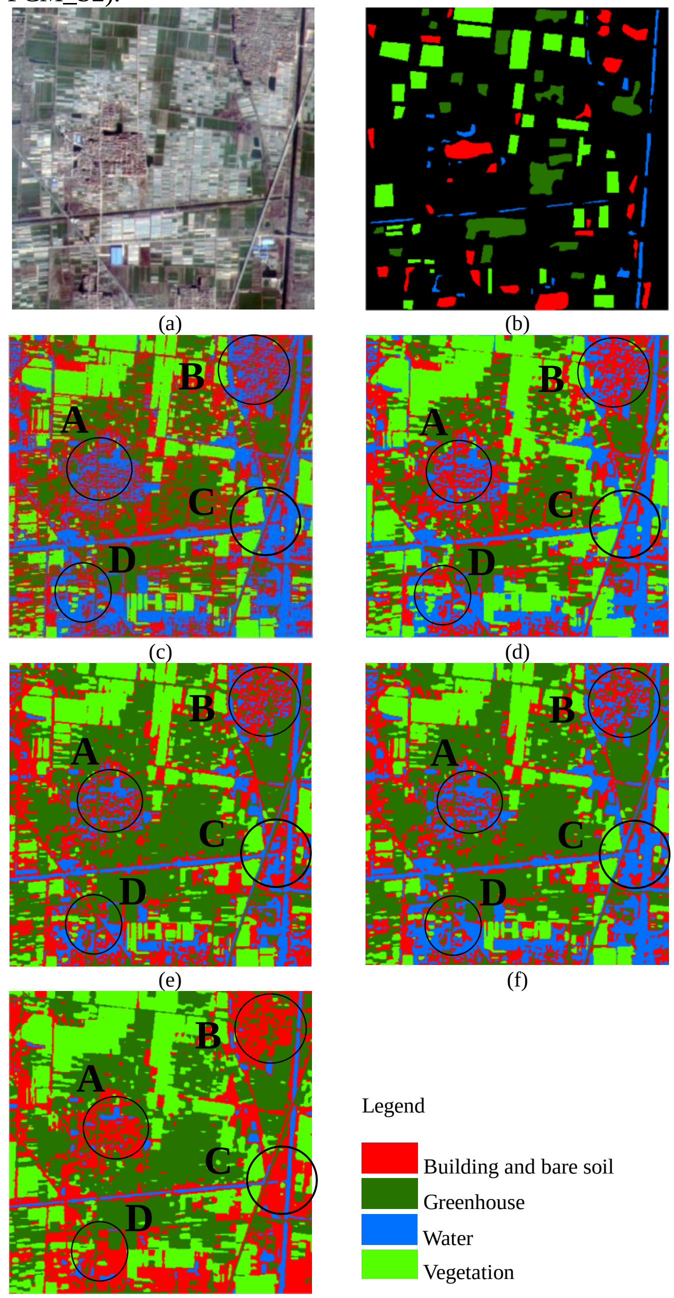


Fig. 10. Dataset and classification results in Experiment 2. (a) ZY-3 image of Xuzhou (RGB (3, 2, 1)). (b) Reference image. (c), (d), (e), (f) and (g) are the classification results of Fig. 10 (a) produced by FCM, FCM_S1, FCM_S2, FLICM, and ADFLICM, respectively..

Fig. 10 (c)-(g) shows the classification results of the ZY-3 image of Xuzhou by the FCM, FCM_S1, FCM_S2, FLICM, and ADFLICM algorithms, respectively. Visually, as shown in Fig. 10 (c), there are a lot of salt and pepper noise in the FCM result. As for FCM_S1, FCM_S2 and FLICM, they enhance FCM to some extent, but show weaker performance than ADFLICM. This can be illustrated by referring to marked areas A-D, where many building and bare soil pixels are misclassified as water pixels by FCM, FCM_S1, FCM_S2 and FLICM, but

ADFLICM show more accurate result. The main reason may be that very different spatial constraints were incorporated into the conventional FCM objective function. Specifically, in FCM_S1 and FCM_S2, the selection of parameter α is difficult and the improper α will constrain improper spatial information on the objective function, which will result in very inaccurate predictions. In FLICM, the classification of the center pixel is greatly influenced by its neighboring pixels but the center pixel's own feature is not fully considered. Thus, they may produce over-smooth results for important structures and small patches. In ADFLICM, however, by introducing the local similarity measure S_{ir} , the weighting factor is not only influenced by its neighboring pixels but also the central pixel, which can provide more proper trade-off between the center pixel and its neighboring pixels.

Table V lists the quantitative results for the five methods. Again, ADFLICM produces the greatest Producer's Accuracy for all classes (except the water class) as well as the greatest Overall Accuracy and Kappa coefficient values.

TABLE V
COMPARISON OF PRODUCER'S ACCURACY, OVERALL ACCURACY AND KAPPA COEFFICIENT OF FIVE CLASSIFICATION METHODS IN EXPERIMENT 2

Class	Number of Testing samples	FCM	FCM_S1	FCM_S2	FLICM	ADFLICM
Building and bare soil	6289	43.01%	49.80%	52.74%	43.01%	**77.77%**
Greenhouse	8029	90.66%	93.70%	99.35%	99.48%	**99.81%**
Water	2697	99.89%	99.93%	99.85%	100.00%	98.85%
Vegetation	12048	93.19%	98.46%	92.73%	92.54%	**98.56%**
Overall Accuracy		82.26%	86.75%	86.57%	84.43%	**94.43%**
Kappa Coefficient		0.7525	0.8133	0.8102	0.7816	**0.9199**

C. Experiment 3: QuickBird Image of Xuzhou

The QuickBird image (400×400 pixels) containing three 0.61-m fused multispectral bands (red, green, and blue) was used in this experiment. It covers an urban area in Xuzhou, China, and was acquired in August 2005 (Fig. 11 (a)). The land cover categories are road, bare soil, water, vegetation1 and vegetation2. The testing samples in the reference map Fig. 11 (b) were obtained from the multispectral imagery, combined with some fieldwork. The parameters used in this experiment are: $c=5$, $m=2$, $\varepsilon=1e-5$, $L=2$, $N_R=8$, and $\alpha=4.6$ (in FCM_S1 and FCM_S2).

The classification results of the five methods are shown in Fig. 11 (c)-(g). In the FCM result, there are a lot of noise in the bare soil region that are misclassified as road. FCM_S1 and FCM_S2 have very close performance and both show fewer noise than FCM. However, more bare soil pixels are misclassified as road, see marked areas A and B. With respect to FLICM and ADFLICM, they are obviously superior to the other three methods. Moreover, the inter-comparison between FLICM and ADFLICM reveals that the latter are more advantageous in persevering spatial details (such as thin features of the vegetation 1 class in area D).

Table VI displays the quantitative results. As shown in the table, FCM, FCM_S1 and FCM_S2 have close classification accuracies (the Overall Accuracies are all below 80%). FLICM and ADFLICM produce Overall Accuracies of 87.5% and 91.4%. The accuracy gains of ADFLICM over FCM, FCM_S1, FCM_S2 and FLICM are 12.64%, 14.59%, 12.56% and 3.93%, respectively.

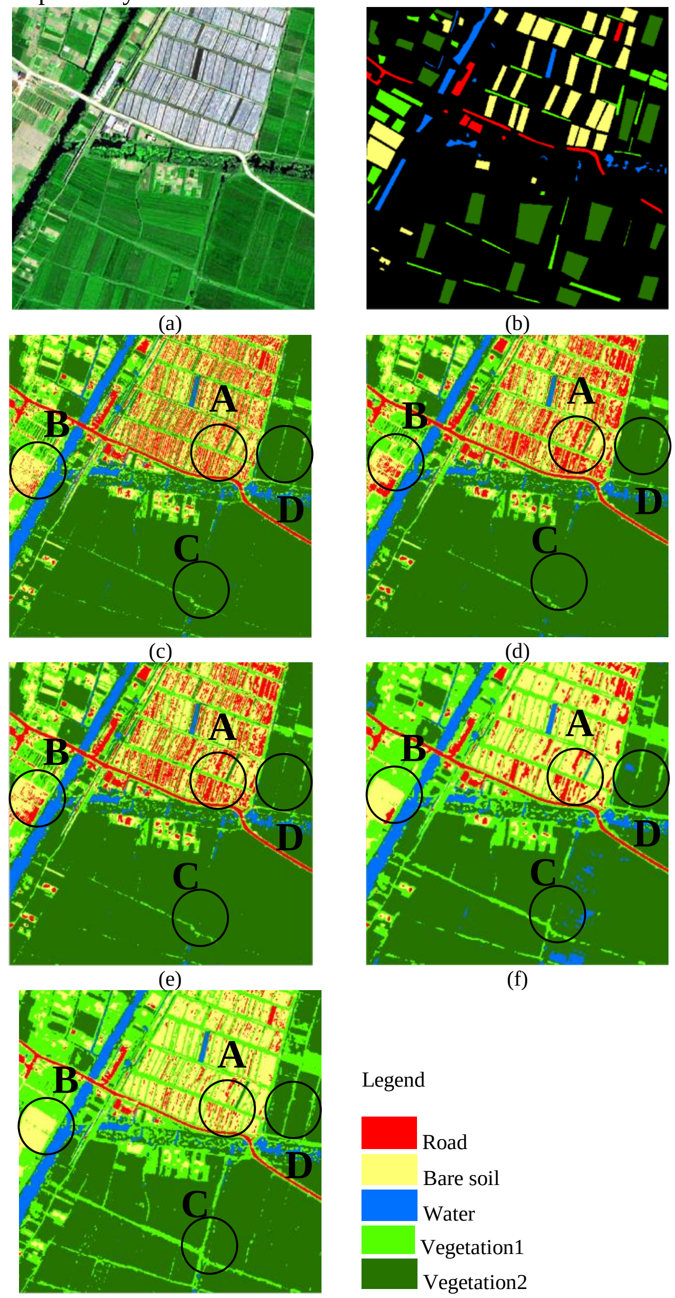


Fig. 11. Dataset and classification results in Experiment 3. (a) *QuickBird* image of Xuzhou (RGB (R, G, B)). (b) Reference image. (c), (d), (e), (f) and (g) are the classification results of Fig. 11 (a) produced by FCM, FCM_S1, FCM_S2, FLICM, and ADFLICM, respectively.

TABLE VI
COMPARISON OF PRODUCER'S ACCURACY, OVERALL ACCURACY AND KAPPA COEFFICIENT OF FIVE CLASSIFICATION METHODS IN EXPERIMENT 3

Class	Number of Testing samples	FCM	FCM_S1	FCM_S2	FLICM	ADFLICM
Road	3405	89.54%	92.01%	**92.54%**	91.51%	83.55%
Bare soil	17831	61.67%	52.61%	58.72%	77.28%	**84.75%**
Water	4322	94.33%	95.53%	95.35%	**98.24%**	96.55%
Vegetation1	8013	53.39%	58.87%	58.47%	86.57%	**91.99%**
Vegetation2	19288	99.78%	**99.83%**	99.87%	94.22%	97.63%
Overall Accuracy		78.79%	76.84%	78.87%	87.50%	91.43%
Kappa Coefficient		0.7117	0.6901	0.7151	0.8309	0.8822

D. Experiment 4: ROSIS Image of University of Pavia

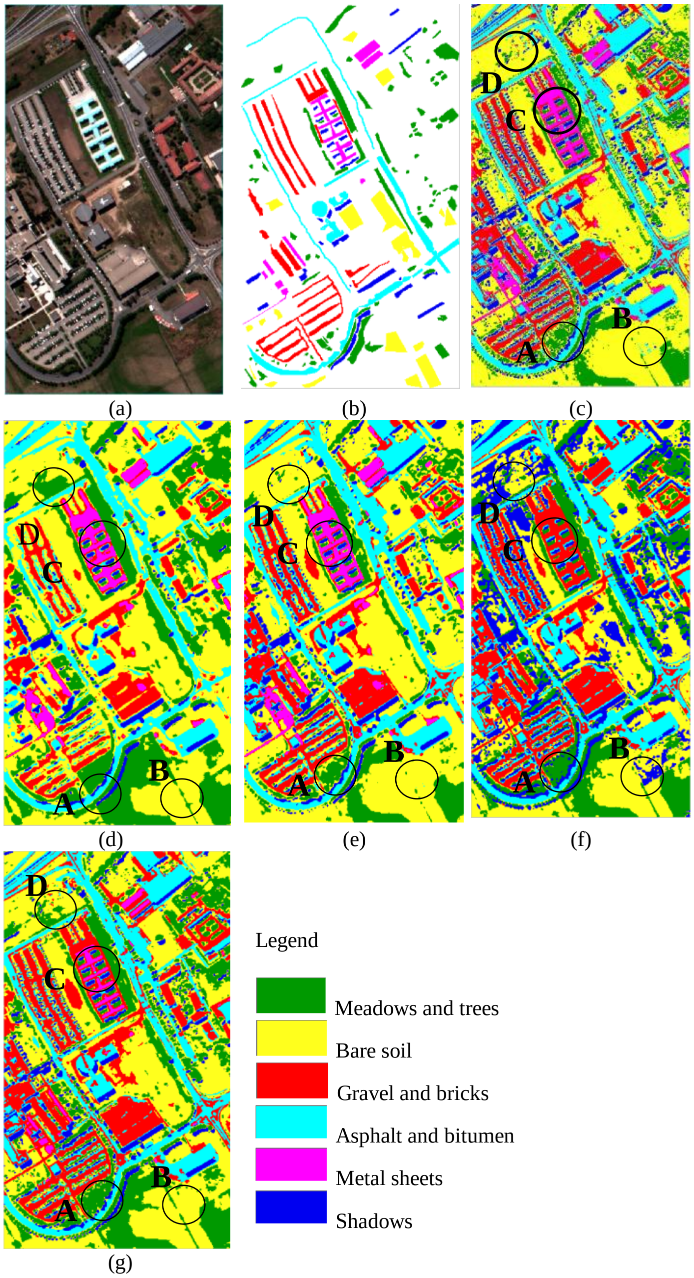


Fig. 12. Dataset and classification results in Experiment 4. (a) *ROSIS* image of Pavia University (RGB (50, 27, 17)). (b) Reference image. (c), (d), (e), (f) and (g) are the classification results of Fig. 12 (a) produced by FCM, FCM_S1, FCM_S2, FLICM, and ADFLICM, respectively..

In this experiment, a hyperspectral dataset was acquired by the ROSIS optical sensor over the urban area of the University of Pavia, Italy. The number of spectral bands is 103 and the spatial size is 610×340. The spatial resolution is 1.3-m. Six classes were identified in the area: meadows and trees, bare soil, gravel and bricks, asphalt and bitumen, metal sheets and shadows. The testing samples were obtained by manual drawing from the multispectral imagery. The parameters used in this experiment are: $c=6$, $m=2$, $\varepsilon=1e-5$, $L=2$, $N_R=8$, and $\alpha=5.3$. To reduce the computational complexity, the principal component analysis-based feature reduction was carried out, and the first six components were used.

The classification results of the five methods are presented in Fig. 12 (c)-(g). Again, ADFLICM is visually more accurate than the other four methods. This can be illustrated by referring to marked areas A-D. With respect to the quantitative evaluation, it should be noted that FLICM and ADFLICM yield lower Producer's Accuracy for ‘Metal sheets’ class with 0% and 64.74%, respectively. The reason may be that improper local spatial information was incorporated. In FLICM and AFLICM, the classification of the center pixel is greatly influenced by its neighboring pixels. If the weighting factor for the neighborhood effect is not properly determined, the final classification will be greatly affected, leading to misclassification. It should be stressed that, however, the less accurate classification occurred for the Metal sheets class, according to the whole classification performance, compared with FCM, FCM_S1, FCM_S2 and FLICM, ADFLICM is visually and quantitatively more accurate.

TABLE VII
COMPARISON OF PRODUCER'S ACCURACY, OVERALL ACCURACY AND KAPPA COEFFICIENT OF FIVE CLASSIFICATION METHODS IN EXPERIMENT 4

Class	Number of Testing samples	FCM	FCM_S1	FCM_S2	FLICM	ADFLICM
Meadows and trees	10561	65.93%	94.94%	73.70%	85.36%	91.62%
Bare soil	7484	94.45%	91.86%	95.14%	85.31%	92.61%
Gravel and bricks	8311	81.76%	67.844%	84.50%	92.24%	95.88%
Asphalt and bitumen	8325	87.72%	73.43%	91.92%	92.00%	88.67%
Metal sheets	3253	90.01%	92.62%	93.21%	0.00%	64.74%
Shadows	3374	93.42%	92.26%	94.43%	93.21%	92.68%
Overall Accuracy		82.85%	84.19%	86.66%	81.67%	89.95%
Kappa Coefficient		0.7896	0.8051	0.8363	0.7722	0.8752

E. Analysis of Neighborhood Level in ADFLICM

In the proposed ADFLICM method, the neighborhood level L affects the classification result. Using two datasets, we investigated the neighborhood level L from 1 to 5, and the results are shown in Figs. 13 and 14, and Tables VIII and IX. Fig. 13 (a)-(e) shows the classification results for the Xuzhou TM image using L =1, 2, 3, 4 and 5, respectively. The results are all affected by the noise to different extents. When L increases, more homogeneous result is obtained and more isolated pixels are removed, but some image details are lost. Focusing on Fig. 13 (b), when L =2, most of the isolated pixels are removed and image details are satisfactorily preserved. The quantitative results for the two images are listed in Tables VIII and IX. As seen from Table VIII, the greatest classification accuracies is achieved when L =2, where the computational cost is acceptable. Tables IX gives a similar result, though the result with L =3 is little better than that with L =2, but the cost is much more than that with L =2, Thus, in this paper, we select L =2 in the ADFLICM considering both classification accuracy and computational cost.

TABLE VIII
OVERALL ACCURACY, KAPPA COEFFICIENT AND COMPUTATIONAL COST OF ADFLICM IN RELATION TO DIFFERENT L FOR THE XUZHOU TM IMAGE

Neighborhood Level	L =1	L =2	L =3	L =4	L =5
Overall Accuracy	92.54%	94.47%	92.74%	91.18%	89.20%
Kappa Coefficient	0.8916	0.9196	0.8930	0.8739	0.8468
Computational Cost (seconds)	8.04	18.77	37.00	59.31	71.00

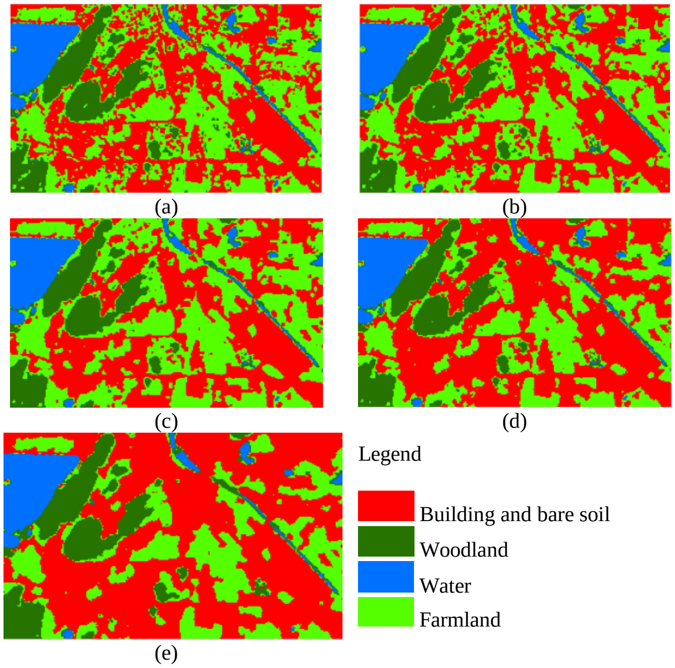


Fig. 13. Classification results for the Xuzhou TM image. (a), (b), (c), (d) and (e) are the classification results produced by ADFLICM when L =1, 2, 3, 4 and 5, respectively.

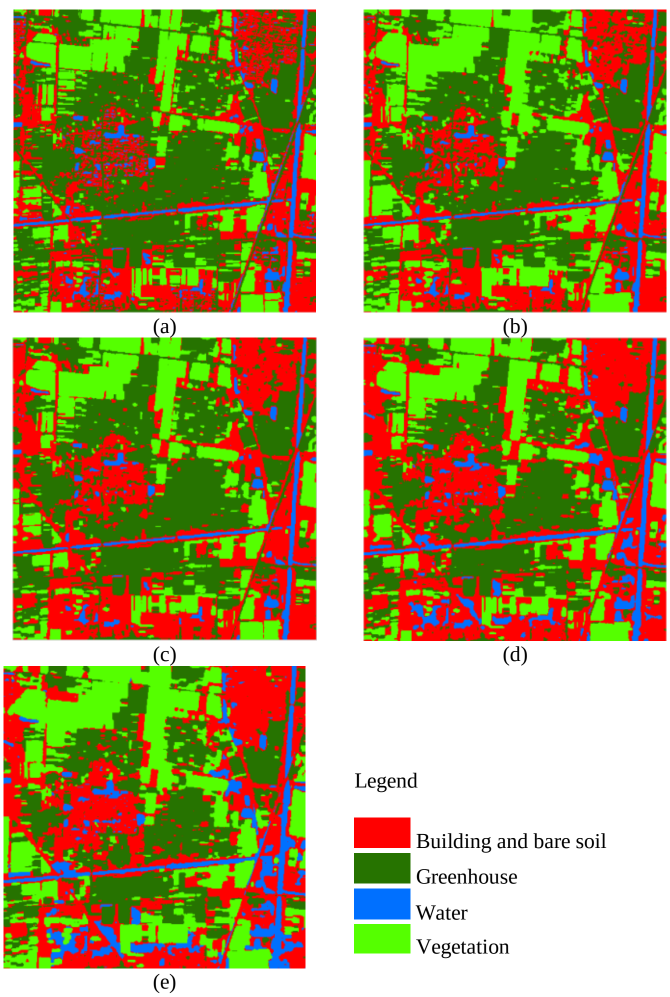


Fig. 14. Classification results for the Xuzhou ZY-3 image. (a), (b), (c), (d) and (e) are the classification results produced by ADFLICM when L =1, 2, 3, 4 and 5, respectively.

TABLE IX
OVERALL ACCURACY, KAPPA COEFFICIENT AND COMPUTATIONAL COST OF ADFLICM IN RELATION TO DIFFERENT L FOR THE XUZHOU ZY-3 IMAGE

Neighborhood Level	L =1	L =2	L =3	L =4	L =5
Overall Accuracy	89.95%	94.43%	94.86%	93.68%	92.49%
Kappa Coefficient	0.8563	0.9199	0.9225	0.9242	0.9156
Computational Cost (seconds)	65	100	142	254	348

F. Computational Complexity Analysis

All four images were used to test the computational complexity of the proposed ADFLICM algorithm and other FCM-based algorithms. All algorithms were tested on an Intel Xeon ® CPU X5675 at 3.06-GHz, and every algorithm repeats ten times on each test image. Table X displays the average computational cost for the five methods. It is clearly that ADFLCIM takes the most time for every image, as weightings and effects on the center pixel from its neighboring pixels need to be calculated in each iteration, but this is the cost of increasing the classification accuracy.

TABLE X
COMPUTATIONAL COMPLEXITY OF ADFLICM AND OTHER FCM RELATED METHODS (SECONDS)

No.	Number of class	Image size	FCM	FCM_S1	FCM_S2	FLICM	ADFLICM
1	4	165×272×6	1.68	4.75	5.01	18.43	18.77
2	4	400×400×4	3.26	11.07	10.75	60.02	84.77
3	5	512×512×3	7.01	21.86	20.35	118.75	126.50
4	6	610×340×6	19.45	29.20	29.03	190.38	211.97

V. CONCLUSION

A novel Adaptive Fuzzy Local Information *C*-Means (ADFLICM) clustering algorithm for remotely sensed imagery classification is proposed in this paper. The proposed algorithm is able to overcome the drawbacks of the well-known FCM by incorporating local spatial and gray level information. The ADFLICM is effective in removing noise pixels and reducing the edge blurring artifact simultaneously. This advantage is based on the definition of a new local similarity measure, which can provide proper trade-off between the center pixel and its neighboring pixels. Experiments on four separate datasets were conducted to demonstrate the effectiveness of ADFLICM. Compared with existing FCM, FCM_S1, FCM_S2 and FLICM, ADFLICM is more accurate by visual and quantitative evaluation. Therefore, ADFLICM is an effective unsupervised classifier for remotely sensed imagery.

In further studies, additional research will be conducted on the selection of the initial cluster center, determining the cluster number and reducing the computational cost.

ACKNOWLEDGMENT

The authors would like to thank the Editor and anonymous reviewers whose insightful suggestions have significantly improved this paper. The authors would also like to thank Prof. P. Gamba from the University of Pavia for providing the ROSIS image of University of Pavia and kind discussions on Experiment 4.

REFERENCES

[1] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens*., vol. 28, no. 5, pp. 823-870, Mar. 2007.

[2] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and K-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.

[3] P. R. Kersten, J. S. Lee, and T. L. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 519–527, Mar. 2005.

[4] N. B. Venkateswarlu and P.S.V.S.K. Raju," Fast isodata clustering algorithms," *Pattern Recognit.*, vol. 25, no, 3, pp. 335-342, Mar. 1992.

[5] E. Blanzieri and F. Melgani, "Nearest neighbor classification of remote sensing images with the maximal margin principle," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1804–1811, Jun. 2008.

[6] K. Xu *et al.*, "Unsupervised satellite image classification using Markov field topic model," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 1, pp. 130–134, Jan. 2013.

[7] J. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithm and applications in image segmentation," *Graph. Models Image Process.*, vol. 58, no. 3, pp. 246–261, May. 1996.

[8] J. C. Dunn, "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters," *Journal of Cybernetics*, vol. 3, pp. 32-57. 1973.

[9] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms. New York, NY, USA: Plenum, 1981.

[10] M. N. Ahmed et al., "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Trans. Med. Imag*., vol. 21, pp. 193–199, 2002. Mar. 2002.

[11] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognit*., vol. 40, no. 3, pp. 825–838, Mar. 2007.

[12] Z. M. Wang, "Adaptive spatial information-theoretic clustering for image segmentation," *Pattern Recognit*., vol. 42, no. 9, pp. 2029–2044, Sep. 2009.

[13] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," *IEEE Trans. Image Process*., vol. 19, no. 5, pp. 1328–1337, May 2010.

[14] M. G. Gong, Z. Q. Zhou and J, J Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Trans. Image Process*., vol. 21, no. 4, pp. 2141–2151, May. 2012.

[15] M. G. Gong et al., "Fuzzy c-means clustering with local information and kernel metric for image segmentation," *IEEE Trans. Image Process*., vol. 22, no. 2, pp. 573–584, Feb. 2013.

[16] N. Li et al., "A Spatial Clustering Method With Edge Weighting for Image Segmentation," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1124–1128, Sep. 2013

[17] Q. Wang and W Shi, "Unsupervised classification based on fuzzy c-means with uncertainty analysis," *Remote Sens Lett.*, vol. 4, no. 11, pp. 1087–1096, Nov. 2013

[18] X. C. Yu et al., "Land cover classification of remote sensing imagery based on interval-valued data fuzzy c-means algorithm," *Sci China Earth Sci.*, vol. 57, no. 6, pp. 1306-1313, 2013.

[19] J. Ji and K. L. Wang, "A robust nonlocal fuzzy clustering algorithm with between-cluster separation measure for SAR image segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 12, pp. 4929-4936, Dec. 2014

[20] Y. Zhong et al., "An adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1235-1248, Apr. 2014.

[21] I. Dópido et al., "A quantitative and comparative assessment of unmixing-based feature extraction techniques for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 421-435, Apr. 2012.

[22] M.N. Ahmed, S.M. Yamany, A.A. Farag, T. Moriaty, Bias field estimation and adaptive segmentation of MRI data using a modified fuzzy c-means algorithm, in: Proceedings of the 13th International Conference on Computer Assisted Radiology and Surgery, CARS'99, Paris, 1999, p. 1004;

[23] D. L. Pham, "Spatial Models for Fuzzy Clustering," *Comput Vis Image Underst*, vol. 84, no. 2, pp. 285-297, Nov. 2001.

[24] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Trans. Syst., Man, Cybern.*, vol. 34, pp. 1907–1916, Aug, 2004

[25] L. Szilagyi *et al.*, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in *Proc. 25th Annu. Int. Conf. IEEE EMBS*, 2003, pp. 17–21.

[26] Wang *et al.* (2013) X. N. Wang, X. B. Lin and Z. Yuan, "An edge sensing fuzzy local information c-means clustering algorithm for image segmentation" in Lecture Notes in Computer Science, vol. 8589, Intelligent Computing Methodologies, D.S. Huang et al, Ed. 2014, pp. 230–240.

[27] G. Y. Sun, "A novel approach for edge detection based on the theory of universal gravity," *Pattern Recognit.*, vol. 40, no. 10, pp. 2766–2775, Oct. 2007.

[28] K. C. Mertens et al., "A sub-pixel mapping algorithm based on sub-pixel/pixel spatial attraction models," *Int. J. Remote Sens*., vol. 27, no. 15, pp. 3293-3310, Aug. 2006.

[29] J. C. Bezdek, "Numerical taxonomy with fuzzy sets," *J. Math. Biol.*, vol. 1, no. 1, pp. 57-71, May. 1974.

[30] J. C. Bezdek, "Cluster validity with fuzzy sets, " *J. Cybernet.*, vol. 3, no. 3, pp. 58-73, 1974.

[31] R. N. Dave, "Validating fuzzy partitions obtained through *c*-shells clustering," *Pattern Recongn Lett.*, vol. 17, no. 6, pp. 613-623, May. 1996.

[32] Y. Fukuyama and M. Sugeno, "A new method of choosing the number of clusters for the fuzzy *c*-means method,". in: Proceedings of Fifth Fuzzy Systems Symposium, pp. 247-250, Jan. 1989.

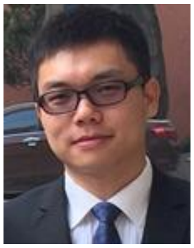
[33] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 13, no. 8, pp. 841-847, Aug. 1991.

[34] S. H. Kwon, "Cluster validity index for fuzzy clustering," *Electron. Lett.*, vol. 34, no. 22, pp. 2176-2177, Oct. 1998.

[35] Y. G. Tang, F.C. Sun and Z.Q. Sun, "Improved validation index for fuzzy clustering," in: American Control Conf., June 8-10, 2005, Portland, OR, USA.

[36] K. L. Wu and M. S. Yang, "A cluster validity index for fuzzy clustering," *Pattern Recongn Lett*., vol. 26, no. 9, pp. 1275-1291, Jul. 2005.

Hua Zhang is an associate Professor in GIS and remote sensing, Key Laboratory for Land Environment and Disaster Monitoring of SBSM, China University of Mining and Technology. He obtained his doctoral degree from China University of Mining and Technology, Xuzhou, China in 2012. He has authored or coauthored over 10 peer-reviewed articles in international journals such as *IEEE Transactions on Geoscience and Remote Sensing*, and *ISPRS Journal of Photogrammetry and Remote Sensing*. His current research interests include multi/hyperspectral and high-resolution remotely sensed images processing, uncertainty in classification, pattern recognition, and remote sensing applications. He was a recipient of the Excellent Doctoral Dissertation Award and the Excellent Graduates in Jiangsu Province, China, in 2012.

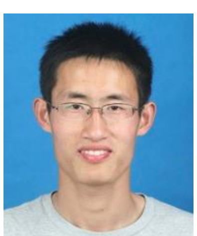


Qunming Wang (M'15) received the B.S. degree and M.S. degree from Harbin Engineering University, China, in 2010 and 2012, and the Ph.D. degree from The Hong Kong Polytechnic University, Hong Kong, in 2015.

He is now a Senior Research Associate in Lancaster Environment Centre, Lancaster University, U.K. From June to December 2013, he was a Visiting Ph.D. Student with Geography and Environment, University of Southampton, Southampton, U.K. He has authored or coauthored around 40 peer-reviewed articles in international journals such as *Remote Sensing of Environment*, *IEEE Transactions on Geoscience and Remote Sensing*, and *ISPRS Journal of Photogrammetry and Remote Sensing*. His current research interests focus on remote sensing image analysis and geostatistics.

Dr. Wang serves as a reviewer for 20 international journals, including most of the journals in remote sensing. He was awarded the hypercompetitive Hong Kong Ph.D. Fellowship to support his three-year Ph.D. study and his PhD thesis was awarded as the outstanding thesis in the Faculty. He was a recipient of the Excellent Master Dissertation Award and the Excellent Graduates in Heilongjiang Province, China, in 2012. He is an Associate Editor for *Computers and Geosciences*.

Wenzhong Shi is a Professor in GIS and remote sensing, Department of Land Surveying and Geo-Informatics, the Hong Kong Polytechnic University. He obtained his doctoral degree from University of Osnabrück in Vechta, Germany in 1994. His current research interests include GIS and remote sensing, uncertainty and spatial data quality control, image processing for high resolution satellite images. He has published some 400 research articles (including over 100 SCI papers) and 10 books. He received the State Natural Science Award from the State Council of China in 2007 and The Wang Zhizhuo Award from International Society for Photogrammetry and Remote Sensing in 2012.



Ming Hao is a postdoctoral researcher at the School of Environmental Science and Spatial Informatics, China University of Mining and Technology. He received his Doctoral degree in China University of Mining and Technology in 2015, China. He has authored or coauthored over 10 peer-reviewed articles in international journals such as *Remote Sensing, International Journal of Remote Sensing and IEEE Geoscience and Remote Sensing Letters.* His research interests include image processing, information extraction from remote sensing images and remote sensing image fusion.