

# Spectral–Spatial Hyperspectral Image Classification via Non-local Means Filtering Feature Extraction

Bing Tu<sup>1,2,3</sup> · Xiaofei Zhang<sup>1,2,3</sup> · Jinping Wang<sup>1,2,3</sup> · Guoyun Zhang<sup>1,2,3</sup> · Xianfeng Ou<sup>1,2,3</sup>

Received: 11 July 2017 / Revised: 22 December 2017 / Published online: 13 March 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** Hyperspectral image (HSI) classification has been a hot topic of research in recent years. The integration of spectral and spatial context is an effective method for HSI classification. This paper proposes a classification method of HSI based on non-local means (NLM) filtering. Firstly, the classification result of HSI is obtained by adopting the support vector machines. Then, the optimization probability image of spatial structure is obtained by using the spatial context information in the first principal component or the first three principal components of HSI to optimize the initial probability map through the NLM filtering. Finally, the final classification results are calculated based on the maximum probability. Experiment results on three real hyperspectral data demonstrate that the proposed NLM filtering based classification method can improve the classification accuracy significantly. Classification results show the effectiveness and superiority of the proposed methods when compared with other methods.

**Keywords** Hyperspectral image · Spatial context · Non-local means filtering · Maximum probability

---

✉ Bing Tu  
tubing@hnist.edu.cn

✉ Xiaofei Zhang  
xiaofei\_zhang@vip.hnist.edu.cn

<sup>1</sup> School of Information and Communication Engineering, Hunan Institute of Science and Technology, Yueyang, China

<sup>2</sup> Key Laboratory of Optimization and Control for Complex Systems, College of Hunan Province, Yueyang, China

<sup>3</sup> Laboratory of Intelligent-Image Information Processing, Hunan Institute of Science and Technology, Yueyang, China

## 1 Introduction

In recent years, hyperspectral image (HSI) technology has developed rapidly and it has been widely applied to the fields of military monitoring, environmental monitoring, and mineral identification [1–3]. HSI has the feature of spectral and spatial information which determines the ability of distinguishing different category in ground and objects. However, HSI has immense amount of data and spectral bands, and there also exist the fact that different objects with same spectrum and the same spectrum of different objects in hyperspectral images [4]. In traditional classification methods, most of them only considering the single pixel, not the relevance between adjacent pixels in spatial position. This work shows that adjacent pixels in spatial location often belong to the same category in HSI.

Initially, HSI classification is made only according to the spectral information of each pixel in the HSI, e.g., k-nearest neighbor method [5], decision tree classifier [6], Bayesian classifier [7, 8], neural network [9], and support vector machine (SVM) method [10–12]. With the increase of data dimension, the “Hughes” phenomenon is introduced unavoidably because of the high spectral dimension of hyperspectral images [13]. HSI feature extraction is an effective way to reduce the dimension of the original data. The common methods of extracting spectral feature include principal component analysis (PCA) [14] and independent component analysis (ICA) [15]. Aiming at the difference of the data independence standard and the calculation method in PCA and ICA, various independent component analysis methods were proposed such as [16]. In addition, several extensions of feature extraction method of HSI had been studied. For example, Benediktsson et al. [17, 18] proposed a feature extraction method based on the multi scale morphology of HSI, which combines the independent principal components with the morphological method to enhance the classification accuracy of HIS. Demir et al. [19] extracted the intrinsic mode component of the hyperspectral images through the empirical mode decomposition (EMD), and realized the image classification by adopting the summation of principal mode components and generalized SVM classification method. To further improve the classification performance, Erturk et al. [20] proposed empirical mode decomposition-based method by combining the empirical mode decomposition with spectral gradient enhancement technique, which has a better classification performance. Demir et al. [21] proposed a new classification method of morphology HSI based on the histogram, which has shown some promising results in several real hyperspectral data sets.

In order to improve the classification performance further, spectral–spatial feature extraction methods have been studied extensively in recent years. [22] proposed the spectral–spatial classification algorithm of HSI based on SVM and the markov random field (MRF). Based on this observation, Zhang et al. [23] proposed a adaptive MRF classification method, which can adjust the weight of MRF energy function and spectrum parameter adaptively according to the region of the pixel. In the processing of hyperspectral image acquisition, there are many factors causing interference to introduce noise. Therefore, Li et al. [24, 25] proposed the multivariate logistic regression to the study of posterior probability distribution,

which can further improve the classification accuracy through combining the posterior probability with the Markov random field statistical model. In order to fully extract the spectral feature information of hyperspectral remote sensing images, Kang et al. [26] proposed a kind of HSI classification method based on the edge-preserving filtering.

In addition, the sparse representation based methods have also been applied in HSI. For sparse representation, pixels from the same class usually share the same spectral characteristics and thus can be represented by liner combination of others, i.e., a test pixel can be represented by a product of structured dictionary and sparse coefficients. Chen et al. [27] presented a HSI classification method using dictionary-based sparse representation, which can achieve a good classification performance. Considering that the joint sparse representation may include pixels from different classes through a fixed square window, Fang et al. [28] applied multiscale adaptive sparse representation (MASR) in the HSI spectral–spatial classification, which can adjust the size and shape of regions adaptively for each test pixel. Moreover, a large number of research results show that the spectral–spatial classification based on SR can effectively improve HSI classification accuracy by combining spatial and spectral information [29, 30].

The filtering method is also an effective way to enhance the contour feature. Compared with bilateral filtering [26], the NLM filtering can achieve a better performance. In recent years, the NLM filtering was applied in the image denoising widely, i.e., MR image denoising [31, 32], ultrasonic image denoising [33, 34], and computed tomography (CT) image denoising [35]. In addition, the processing speed of the NLM can also increase significantly by using the fast NLM algorithm [36–38]. And thus, the NLM method is consider to apply into the HSI classification [39].

In this paper, the spectral–spatial classification method based on the probability of the NLM optimization is proposed, and the main steps are shown as follows: (1) adopting the support vector machines to obtain the classification result of HSI. (2) using the spatial context information in the main component of HSI to optimize the initial probability through the non-local means filtering, and then, obtain the optimization probability image of spatial structural. (3) calculating the final classification results based on maximum probability. Experiments based on three real data set show that the NLM filtering can smooth the probabilities while ensure the consistency between the smoothed probabilities and real object boundaries. Moreover, the NLM filtering method can improve the classification accuracy of SVM significantly, and thus, will be quite useful in real application.

The rest of this paper is organized as follows. Section 2 introduces the NLM and structural similarity non-local means (SSIM–NLM) filtering algorithm. Section 3 describes the spectral–spatial classification method. Section 4 is experimental results and discussions. Finally, the conclusion is given in Sect. 5.

## 2 Method of filtering

### 2.1 Non-local means filtering

The NLM algorithm has been recently proposed to deal with the image denoising problem [40]. The basic idea of NLM filtering is to construct the weight of the average by measuring the image similarity, which makes full use of the redundant information and preserves the image features furthest in denoising processing [41]. Here, the NLM filtering is used to calculate the weighted average of all pixels as the estimated value of pixel  $i$ , and the output of NLM filtering is calculated as

$$NL_i = \sum_{j \in I} W_{ij}(I) p_j \quad (1)$$

where  $I$  is the input image,  $p_j$  represents the probability of  $j$ , and  $j$  is neighboring pixel surrounding the pixel being compared  $i$ . Among them, the weight  $W_{ij}$  depends on the similarity between the pixel  $i$  and pixel  $j$ , which meets the conditions of  $0 \leq W_{ij} \leq 1$  and  $\sum_j W_{ij} = 1$ . Therefore, the probability of the NLM filtering is calculated as

$$W_{ij} = \frac{1}{Z(i)} e^{-\frac{d(i,j)}{h^2}} \quad (2)$$

$$d(i,j) = \left\| v(N_i) - v(N_j) \right\|_{2,a}^2 \quad (3)$$

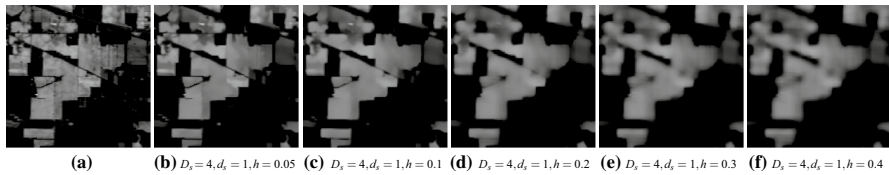
$$Z(i) = \sum_j e^{-\frac{d(i,j)}{h^2}} \quad (4)$$

where  $\left\| v(N_i) - v(N_j) \right\|_{2,a}^2$  represents the Gaussian weighted Euclidean distance [42],

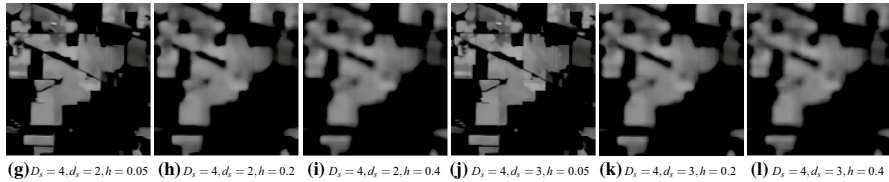
and  $a > 0$  is the standard deviation of the Gaussian kernel. The similarity between  $i$  and  $j$  is determined by the similarity of the gray value vector  $v(N_i)$  and  $v(N_j)$ .  $N_i$  represents the fixed square area whose center is pixel  $i$ , and  $h$  is the smoothing coefficient, which controls the filtering intensity of images.

In the realization of NLM filtering, two windows in fixed size are set, including a searching window ( $D * D, D = 2 * D_s + 1$ ) and a neighborhood window ( $d * d, d = 2 * d_s + 1$ ). The neighborhood window slides in the search window, determining the weight of pixels according to the similarity of the Eq. (2) by adopting the method based on integral figure to realize the fast NLM filtering.

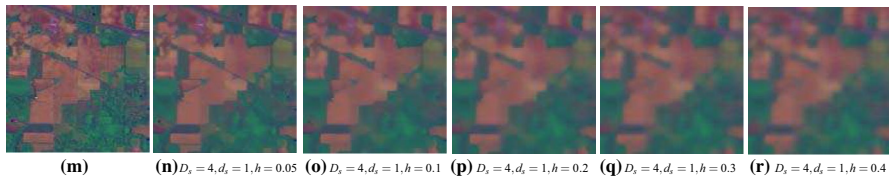
Figures 1, 2, 3 and 4 compare the effect of NLM method on the image filtering with different parameters. Figures 1a and 3m represent the filtering results of the



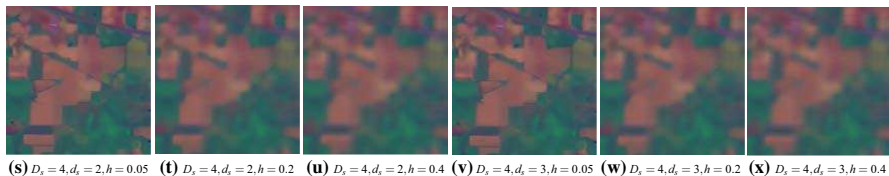
**Fig. 1** **a** Gray input image. **b–f** Gray image filtered result obtained NLMs with different parameter settings



**Fig. 2** **g–l** Gray image filtered result obtained NLMs with different parameter settings



**Fig. 3** **m** Color input image. **n–r** Color images filtered result obtained NLMs with different parameter settings (Color figure online)



**Fig. 4** **s–x** Color images filtered result obtained NLMs with different parameter settings (Color figure online)

gray and color input images . Figure 1b–f, combined with the color image parameter, set the  $D_s = 4$ , the  $d_s = 1$ , and  $h = 0.05, h = 0.1, h = 0.2, h = 0.3, h = 0.4$ , respectively. These two sets of images demonstrate that when the search window  $D_s$  and neighborhood window  $d_s$  are fixed, the filter output will become increasingly blurred with the increase of smoothing coefficient  $h$ . It is shown that the filtered output image is much smoother with proper parameter when compared with the initial input image.

## 2.2 Structural similarity non-local means filtering

The NLM filtering can achieve a good denoising performance in the smooth area but with the limitation in the complex structure information area. Taking other factors such as brightness, contrast and structure information into account, Wang zhou et al. proposed the structural similarity index (SSIM) based on NLM filtering, which shows much better performance with respect to the weighted coefficient, and thus, can be used to evaluate the image similarity and meet the sensitive to structure characteristics of objects.

$$SSIM(i, j) = \frac{(2\mu_i\mu_j + C_1)(2\sigma_{ij} + C_2)}{(\mu_i^2 + \mu_j^2 + C_1)(\sigma_i^2 + \sigma_j^2 + C_2)} \quad (5)$$

where  $\mu_i$  and  $\sigma_i^2$  represent the average and the variance of gray scale in the  $i$ 's neighborhoods, respectively.  $\sigma_{ij}$  represents the gray value covariance of the neighborhoods  $N_i$  and  $N_j$ .  $C_1$  and  $C_2$  are chosen as a minimum constants in case of the denominator is zero. According to the Gaussian distance, we can know that the closer the  $d(i, j)$  is to zero, the better the similarity of the two small pieces is. In other words, the higher the SSIM value, the better the similarity. In order to be consistent with the Gaussian weighted Euclidean distance, the parameters for similar structure is defined as follows:

$$S(i, j) = \frac{1 - SSIM(i, j)}{2} \quad (6)$$

where  $S(i, j) \in [0 \sim 1]$ . The higher the  $S(i, j)$  is, the lower the similarity of the two small pieces would be, and when the two small pieces are in complete agreement, the parameter value is 0, which shows the same monotonicity with the Gaussian weighted Euclidean distance. Here, the improved  $d(i, j)$  based on structural similarity parameters can be calculated as

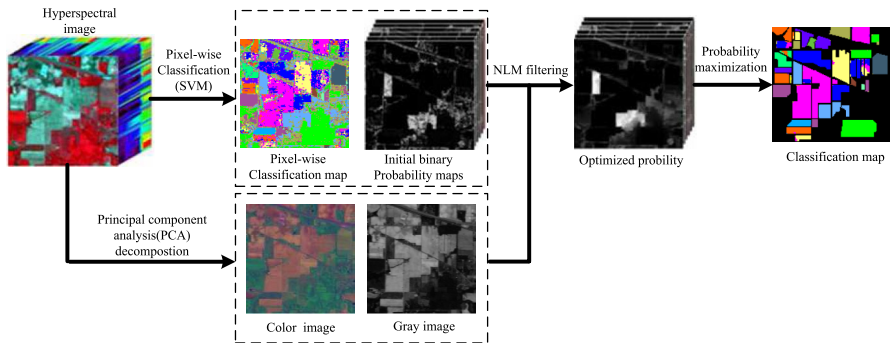
$$d(i, j) = \frac{S(i, j)}{E[S(i, j)]} \left\| v(N_i) - v(N_j) \right\|_{2,a}^2 \quad (7)$$

Therefore, the probability of the SSIM-based non-local filtering method is shown as follows:

$$W_{i,j} = \frac{1}{Z(i)} e^{-\frac{\frac{S(i,j)}{E[S(i,j)]} \left\| v(N_i) - v(N_j) \right\|_{2,a}^2}{h^2}} \quad (8)$$

## 3 Spectral-spatial classification with NLM Filtering

Figure 5 is the hyperspectral image classification framework based on NLM filtering. The main steps are shown as follows. Firstly, the initial binary probability maps of HSI



**Fig. 5** Schematic of the proposed non-local means based spectral-spatial classification method

are obtained by adopting the SVM. Then, the NLM filtering-based method are conducted on initial binary probability maps with gray or color guidance image to obtain the optimization probability maps. Finally, the final classification result can be calculated according to the maximum probability. This paper proposes four classification methods, i.e., the classification method based on the NLM filtering and the gray guided image (NLM\_G), the method method based on the NLM filtering and the color guided image (NLM\_C), the method based on the SSIM–NLM filtering and the gray guided image (SNLM\_G), and the method based on the SSIM–NLM filtering and the color guided image (SNLM\_C). All experiments adopt the default parameter settings. The major steps of the proposed method can be concluded in Algorithm 1.

### 3.1 Initial binary probability maps

As shown in Figure 5, the proposed spectral-spatial classification method based on NLM filtering algorithm obtains the initial probability classification results by adopting a pixel level classifier, such as SVM, which can be expressed in the form of probability map:  $p = (p_1, \dots, p_n)$ , and  $c = (c_1, \dots, c_i)$ . Here,  $p_i \in [0, 1]$  shows the probability of pixels  $i$  belonging to the part of class  $n$ , and the initial binary probability  $p_i$  is shown as follows:

$$p_i = \begin{cases} 1 & \text{if } c_i = n \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

where  $L \equiv \{1, \dots, n\}$  represents the set of labels, and  $c = (c_1, \dots, c_i)$  denotes the classification map of labels.

### 3.2 Probability graph filtering and classification

Without considering the spatial information of HSI, all the probability values of the per-pixel classification based on SVM are valued at either 0 or 1, therefore, there

are lots of misclassification in the initial probability maps and the classification results of the edges are frequently inconsistent with the real target edges. To solve this problem, the NLM filtering is conducted on the initial probability maps, and the probability can be expressed as weighted average of the initial probability in the local window [13], which can be calculated by

$$\widehat{p}_i = \sum_j W_{i,j}(I) p_j \quad (10)$$

where  $i$  and  $j$  represent as the center of the pixel and its neighborhood pixel, respectively. The filtering weights  $W$  is chosen such that the filter preserves the spatial structure information of a guided image  $I$ , and  $W_{i,j}$  represents the optimized probability. In this paper, the NLM filtering and SSIM–NLM filtering are adopted as the specified filtering. The weights  $p_i$  obtained by the NLM filtering and SSIM–NLM filtering are expressed as follows:

$$W_{i,j}(I) = \frac{1}{Z(i)} e^{-\frac{\|v(N_i) - v(N_j)\|_{2,a}^2}{h^2}} I_j \quad (11)$$

$$W_{i,j}(I) = \frac{1}{Z(i)} e^{-\frac{\frac{S(i,j)}{E[S(i,j)]} \|v(N_i) - v(N_j)\|_{2,a}^2}{h^2}} I_j \quad (12)$$

In addition, the guided image also has an important effect on the filtering results. In this paper, the gray image and the color image obtained by PCA are adopted as the guided image.

1. The gray guided image. As shown in Fig. 5, the dimension reduction is performed on the HSI, and the first principal component is chosen as the guided image of NLM filtering, which provides the optimal representation of the mean square value of the HSI and keeps the significant information as much as possible.
2. The color guided image. Instead of using the single principal component, the color image is composed of the first three principal components of HSI, which has the largest mean square build three-channel pseudo color image. Compared with the gray guided image, the color guided image has a great impact on spatial characteristics of HSI by using the first three principal components.

After obtaining the optimized probability images, the category of the pixels  $i$  in HSI can be selected by the maximum probability, which is shown as follows:

$$\widehat{c}_i = \arg \max_n \widehat{p}_i \quad (13)$$



**Algorithm 1** The Spectral-spatial classification based on NLM filtering

1: **INPUT:** Hyperspectral image,  $x = (x_1, \dots, x_i) \in \mathbb{R}^{d \times I}$  denote a hyperspectral image of  $d$ -dimensional feature vectors, and  $L \equiv \{1, \dots, n\}$  be the set of labels,  $c = (c_1, \dots, c_i)$  represents the classification map of labels. Given a training set  $T_\tau \equiv \{(x_1, c_1), \dots, (x_\tau, c_\tau)\} \in (\mathbb{R}^d \times L)^\tau$ , where  $\tau$  is the total number of training samples.  $S \in (1, 2, 3, 4)$  represents the different optimized probability  $\hat{p}_i$  obtained by different filtering based methods, i.e., the NLM\_G, the NLM\_C, the SNLM\_G and the SNLM\_C methods. Two fixed window size parameters  $D_s$  and  $d_s$ , and the smoothing coefficient  $h$ .

2: **OUTPUT:**  $\hat{c} = (\hat{c}_1, \dots, \hat{c}_i)$  is the spectral-spatial classification result and each pixel with a result  $\hat{c}_i \in L$ .

3: **BEGIN**

4: Input training set  $T_\tau$ , and the initial classification result  $c$  obtained by SVM.

5: According formula (8), calculate the initial probability  $p_i$ , and  $c_i = \{1, \dots, n\}$ .

6: Obtain the gray guide image  $I_g$  and the color guide image  $I_c$  by adopting PCA.

7: **if S=1 then**

$$\hat{p}_i = \sum_{j \in I} W_{i,j}(I) p_j$$

$$\text{where } W_{i,j}(I) = \frac{1}{Z(i)} e^{-\frac{d(i,j)}{h^2}} I_{g_j} \quad \text{and} \quad d(i,j) = \|v(N_i) - v(N_j)\|_{2,a}^2$$

**else if S=2 then**

$$\hat{p}_i = \sum_{j \in I} W_{i,j}(I) p_j$$

$$\text{where } W_{i,j}(I) = \frac{1}{Z(i)} e^{-\frac{d(i,j)}{h^2}} I_{c_j} \quad \text{and} \quad d(i,j) = \|v(N_i) - v(N_j)\|_{2,a}^2$$

**else if S=3 then**

$$\hat{p}_i = \sum_{j \in I} W_{i,j}(I) p_j$$

$$\text{where } W_{i,j}(I) = \frac{1}{Z(i)} e^{-\frac{\frac{S(i,j)}{E[S(i,j)]} \|v(N_i) - v(N_j)\|_{2,a}^2}{h^2}} I_{g_j} \quad \text{and} \quad d(i,j) = \|v(N_i) - v(N_j)\|_{2,a}^2$$

**else if S=4 then**

$$\hat{p}_i = \sum_{j \in I} W_{i,j}(I) p_j$$

$$\text{where } W_{i,j}(I) = \frac{1}{Z(i)} e^{-\frac{\frac{S(i,j)}{E[S(i,j)]} \|v(N_i) - v(N_j)\|_{2,a}^2}{h^2}} I_{c_j} \quad \text{and} \quad d(i,j) = \|v(N_i) - v(N_j)\|_{2,a}^2$$

**end if**

$$\hat{c}_i = \arg \max_n \hat{p}_i$$

8: **END**

## 4 Experiments

### 4.1 Experimental setup

#### 4.1.1 Data sets

The proposed method is performed on three hyperspectral data sets, i.e., the Indian Pines image, the University of Pavia image and the Salinas image.

The Indian Pines image was recorded by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor in North-western Indiana, USA. The image is composed of  $145 \times 145$  pixels, including 220 bands, and the spatial resolution is

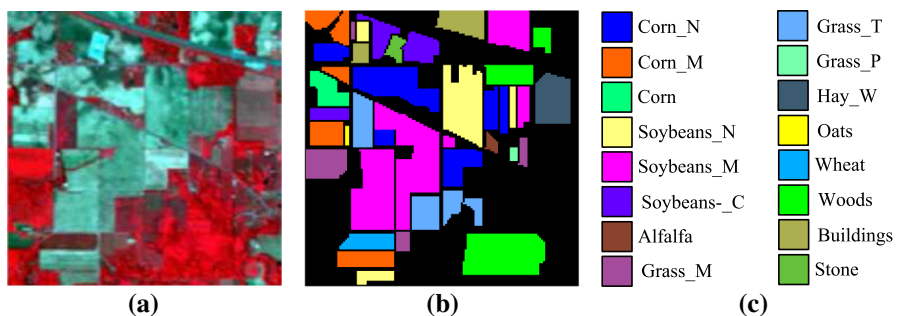
20 m per pixel and a spectral coverage ranging from 0.4 to 2.5  $\mu\text{m}$ . After removing the 20 water absorption bands (absorption bands no. 104–108, 150–163, 220), the rest of 200 bands are used for experiments in this paper. Figure 6 shows the false-color composite of the Indian Pines image and the corresponding reference data.

The University of Pavia image was recorded by ROSIS-03 remote sensing equipment, which is composed of  $610 \times 340$  pixels. The image includes 115 bands with the spatial resolution of 1.3 m per pixel and the wavelength ranging from 0.43 to 0.86  $\mu\text{m}$ . After 12 noise bands removed, the rest of 103 bands are used for experiments. Figure 7 shows the false-color composite of the University of Pavia image and the corresponding reference data.

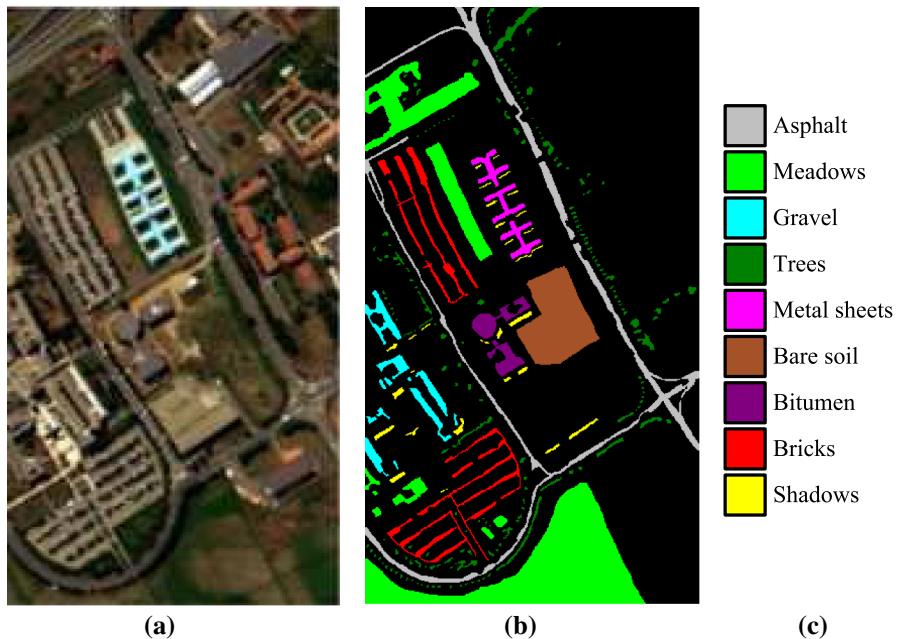
The Salinas image data was recorded by AVIRIS remote sensing sensor at Salinas valley, California, USA. The image has 224 bands of size  $512 \times 217$  pixels and the spatial resolution is 3.7 m. Twenty water absorption bands (no. 108–112, 154–167, 224) were discarded before classification. The false-color composite of the Salinas image and the corresponding reference data which contain 16 classes are presented in Fig. 8.

#### 4.1.2 Quality indexes

This paper adopts three common evaluation indexes of HSI classification accuracy: i.e., overall classification accuracy (OA), average classification accuracy (AA), and Kappa coefficient. OA refers to the percentage of correctly classified pixels. AA is the mean of the percentage of correctly classified pixels for each class. The kappa coefficient gives the percentage of correctly classified pixels corrected by the number of agreements that would be expected purely by chance.



**Fig. 6** **a** The false-color composite of the Indian Pines image. **b, c** Reference data for the Indian Pines image (Color figure online)



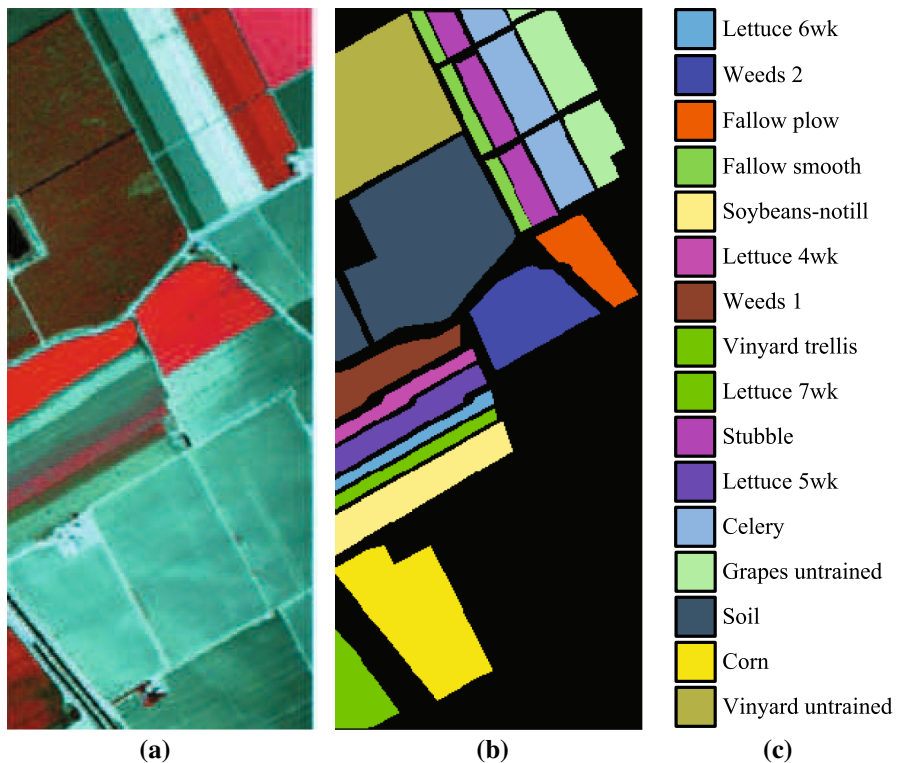
**Fig. 7** **a** Tree-band color composite of the University of Pavia image. **b, c** Ground truth data of the University of Pavia image (Color figure online)

## 4.2 Classification results

### 4.2.1 Parameter analysis

This section analyzes the effect of algorithm parameters based on the NLM filtering spectral-spatial classification algorithm. The NLM algorithm contains two parameters based on fixed window size  $D_s$ ,  $d_s$  and a smoothing coefficient  $h$ . The neighborhood window slides in the search window, and the value of pixels is determined according to the similarity between the neighborhood. Figures 9, 10 and 11, 12 show the influence of the final object recognition accuracy of different parameter settings in the Indian Pines with hyperspectral data.

As shown in Table 2, 10% reference data pixels from the ground are chosen as the training set randomly. In analyzing the effect of  $d_s$  on parameters, the effect of different  $h$  to the classification accuracy of the algorithm is compared with fixed  $D_s = 4$ . As shown in Figures 9 and 10, when  $D_s$  and  $d_s$  are fixed, the AA of the algorithm decreases dramatically if the smoothing coefficient  $h$  is oversize, it is because that the oversize smoothing coefficient will lead to excessive smoothing of the image, and thus, result in the wrong classification in some small scale regions. As shown in Figure 9a, with the parameter  $h$  increases, the tendency of OA increases first and then decreases, since the wrong classification of some small scale targets such as Oats which only contains 20 pixels, the average classification accuracy of the algorithm declines sharply. Similarly, if  $h$  is smaller,

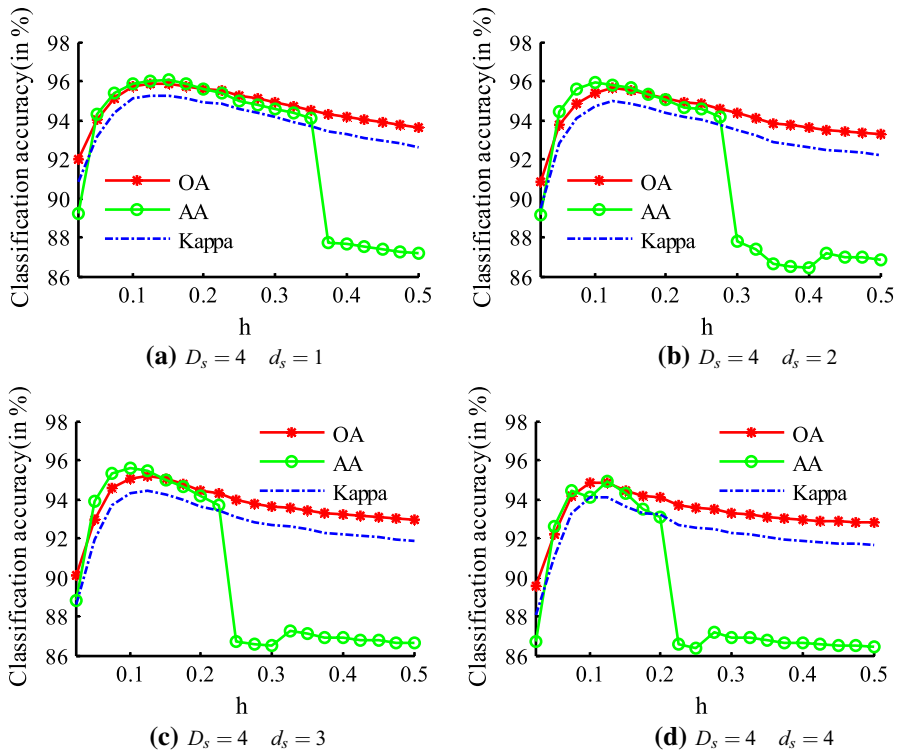


**Fig. 8** **a** Tree-band color composite of the Salinas image. **b, c** Ground truth data of the University of Salinas image (Color figure online)

the performance of the algorithm also declines obviously because the more edge detail ingredients are kept, the more noise will be remained.

At the same time, in the process of probability optimization, different  $D_s$  and  $d_s$  also have greater impact on the accuracy of the algorithm. Since the higher or the smaller range, which is determined by the the size of the neighborhood window and the search window, has a great impact on the the final classification accuracy. Moreover, Fig. 10e–h show the effect of different search window  $D_s$  on the classification accuracy. Figures 11 and 12 demonstrate the influence of different parameter settings on the algorithm accuracy of object recognition while using the first three principal components as the guided image. The experiments can reach the similar conclusion as Figs. 9 and 10:  $D_s$ ,  $d_s$ , and  $h$  must be in a specific range. Therefore,  $D_s = 44$ ,  $d_s = 1$ ,  $h = 0.1$  are set as the default parameters in this paper.

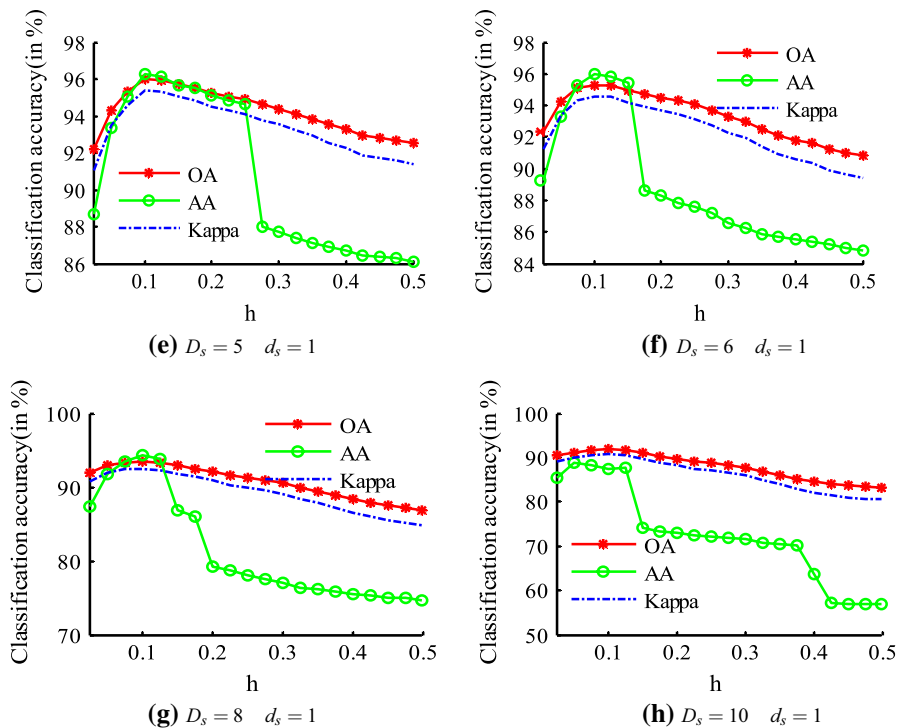
Spectral–spatial classification framework based on probability optimization can be combined with other edge smooth-keeping algorithm or the local optimization algorithm based on statistical methods, i.e., the methods based on statistics according to the local window pixels belonging to one kind of the quantity of optimization classification results [43], WLS (WLS: weighted further



**Fig. 9** Indian Pine image: analysis of the influence of the parameters  $D_s$ ,  $d_s$ , and  $h$  when the first principal component is selected as the guidance image

squares) can realize blurring image under the condition of keeping the edge as much as possible by solving a weighted further squares optimization problem [44]. Normalized convolution filtering (NC: normalized convolution) [45] is an edge-preserving filter which is similar to bilateral filter. Bilateral filter (BF) [26] can get every pixel value through itself and other pixels in the neighborhood after weighting average because of its simple expression. Guided filter (GF) [26] is a filtering algorithm based on local linear model, having the function of the joint edge-preserving filtering.

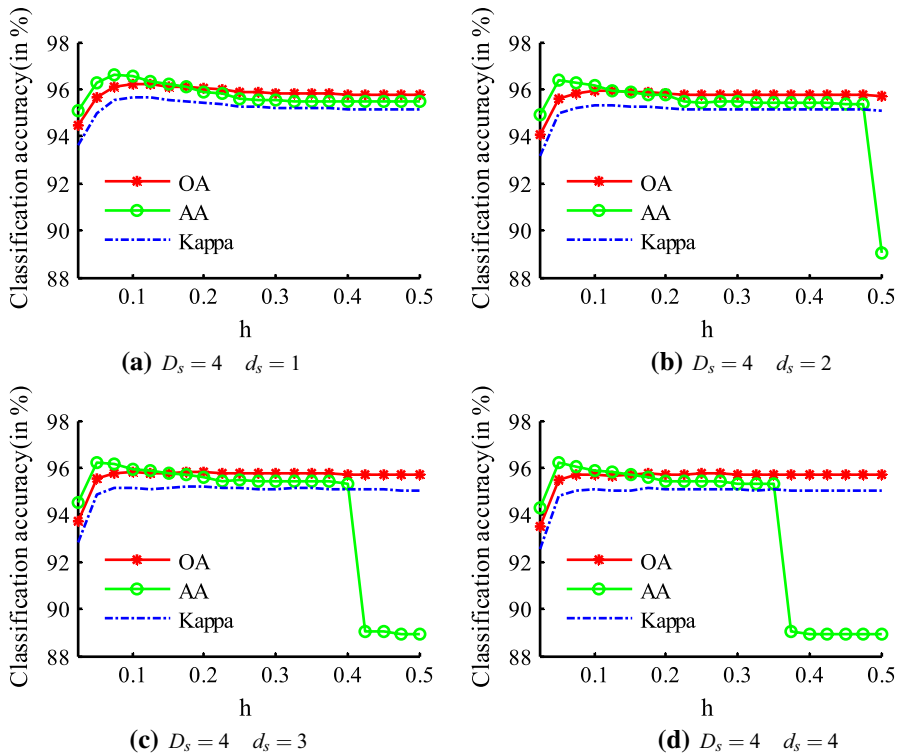
In Table 1, we compare the identification accuracy in the same probability optimization framework using other methods and the NLM filtering method in this paper. As shown in the Table 1, WLS, NC, BF and GF can obtain a similar OA: 94.93, 95.2, 94.93 and 95.2% because different edge smoothing algorithm can make the space smooth and get the probability image match the true edge. On the contrary, without taking the edge information into consideration, the classification accuracy of the statistics is limited, which can only reach to 89.33%.



**Fig. 10** Indian Pines image: analysis of the influence of the parameters  $D_s$ ,  $d_s$ , and  $h$  when the first principal component is selected as the guidance image

#### 4.2.2 Comparisons with other approaches

In this paper, the method based on the NLM filtering is compared with the widely used methods of classification and feature extraction algorithm, i.e., extended morphological profiles (EMP) [18], automatic extended attribute profiles (AEAP) [46], logistic regression and multi-level logistic (LMLL) [25], bilateral filter (BF) and guided filter (GF) [26]. The SVM is implemented by using the Gaussian kernel with fivefold cross-validation. Based on EMP, according to the first three principal components of HSI and two pixels with an increasing size of the morphological operator in turn, four morphological opening and closing and refactoring operations are performed, and the multi-scale morphological characteristics is constructed. Based on AEAP, features are built by using the default parameters and adopting the Profatran software provided by the original authors. The LMLL, BF and GF methods use the Matlab codes provided by the authors and the default parameters setting. The proposed NLM\_G method is based on the NLM filtering and the gray guided image, the proposed NLM\_C method is based on the NLM filtering and the color guided image, the SNLM\_G method is based on the SSIM-NLM filtering and the gray guided image, and the SNLM\_C method is

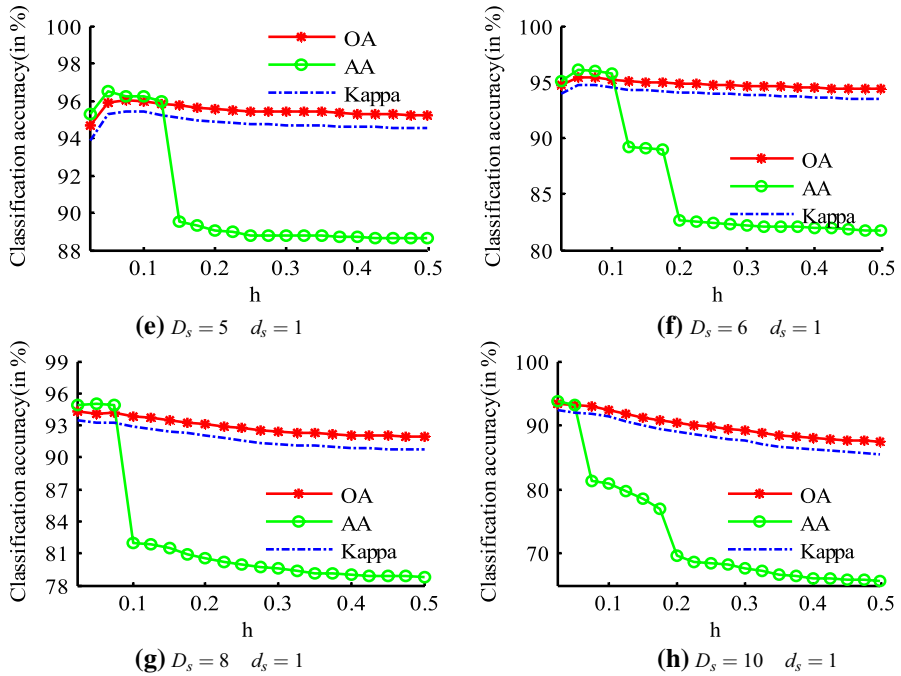


**Fig. 11** Indian Pines image: analysis of the influence of the parameters  $D_s$ ,  $d_s$ , and  $h$  when the color composite of the first three principal components is selected as the guidance image (Color figure online)

based on the SSIM–NLM filtering and the color guided image. All experiments adopt the default parameter settings.

Firstly, the experimental comparison is performed on the Indian Pines data set. Table 2 shows the number of training samples and testing samples in the experiment. Among them, the training samples accounted for 10% of the reference data are selected randomly. Figure 13 shows the classification maps based on different method and the corresponding classification accuracy. As shown in Fig. 13, the classification results based on EMP and AEAP still have an obvious mis-classification of similar noise. Compared with the SVM, the classification accuracy of NLM filtering has increased approximately from 79.81 to 96%. and compared with LMLL, BF and GF, the classification accuracy in ground and objects is higher. As shown in Table 2, the Kappa coefficient has increased significantly, and the NLM\_C method can obtain the best classification results in terms of OA, AA and Kappa coefficient.

Secondly, experiments are conducted on the University of Pavia image data set. Figure 14 shows the classification results base on different methods and the corresponding overall classification accuracy, respectively. Table 3 shows the different



**Fig. 12** Indian Pines image: analysis of the influence of the parameters  $D_s$ ,  $d_s$ , and  $h$  when the color composite of the first three principal components is selected as the guidance image (Color figure online)

**Table 1** The statistics-based method is applied with a  $7 \times 7$  windows

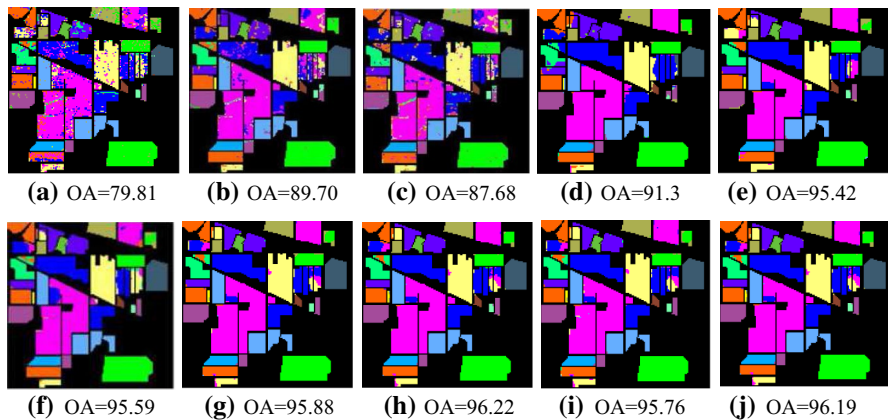
Methods	Statistics	WLS	NC	BF	GF	<i>NLM_G</i>	<i>NLM_C</i>	<i>SNLM_G</i>	<i>SNLM_C</i>
OA	89.33	94.93	95.2	95.42	95.29	95.88	96.22	95.76	96.19
AA	83.33	94.48	94.32	94.81	94.81	96.01	96.57	95.88	96.55
Kappa	87.81	94.17	94.49	94.75	94.6	95.25	95.64	95.12	95.62

The parameters of the WLS filter are set to be  $\alpha = 1.4$  and  $\lambda = 0.3$ . The parameters of the NC filter are set to be  $\delta_s = 3$  and  $\delta_r = 0.2$ . The parameters of the BF filter are set to be  $\delta_s = 3$  and  $\delta_r = 0.2$ . The parameters of the GF filter set to be  $\gamma = 3$  and  $\varepsilon = 0.01$

number of training samples and test samples (the training samples account for 6% of the reference data) and the classification accuracy, which includes single class accuracy (CA), AA, OA and Kappa coefficient. From Table 3, the OA of the proposed method is more significant than EMP, AEAP, LMLL, and BF and GF. Compared with the traditional SVM algorithm, the proposed method can significantly improve the classification accuracy.

Thirdly, the experiment is performed on the Salinas image data set. Figure 15 shows the classification results with different methods and the corresponding classification accuracy respectively. Similarly, to analyze the performance of different methods, the training samples randomly selected account for 2% of the reference





**Fig. 13** Classification results (Indian Pines image) obtained by (a) the SVM method (full feature space), (b) the EMP method, (c) the AEAP method, (d) the LMLL method, (e) the BF method, (f) the GF method, (g) the NLM\_G method, (h) the NLM\_C method, (i) the SNLM\_G method, and (j) the SNLM\_C method. The value of OA is given in percent

data, which is shown in Table 4. As can be seen, the overall classification of the proposed method is more significant than the EMP, AEAP, LMLL, BF and GF. Compared with the traditional SVM method, the proposed methods can significantly improve the classification accuracy. For example, for the Vinyard\_U, the classification accuracy can be improved from 64.2 to 89.39%. The proposed NLM filtering-based methods can further improve the accuracy of ground-objects classification both in the three data set of experiments.

The influence of different training samples to the performance of the proposed method is analyzed in Fig. 16. Experiments are conducted on the Indian Pines data set, the University of Pavia data set and the Salinas data set, respectively. Since the different NLM filtering methods tend to share the similar classification performance, this part chooses the proposed NLM\_G method for hyperspectral image classification. Figure 16 shows the classification accuracy of the proposed NLM\_G algorithm in the Indian Pines data set (the number of training samples increases from 1 to 10%), the University of Pavia data set (the number of training samples increases from 1 to 6%) and the Salinas data set (the number of training samples increases from 1 to 2%), respectively. As shown in Fig. 16a, as the training sample increases, the proposed method can always significantly enhance the classification accuracy when compared with SVM. Similar conclusion can be obtained when the experiments are performed on the the University of Pavia and the Salinas data set.

#### 4.2.3 Computational complexity

The computational complexity of the proposed methods and other aforementioned methods are reported in Table 5. Experiments are performed on a 2.9 GHz CPU notebook with 8GB memory by using Matlab. Table 5 shows the computing time of the proposed NLM\_G, NLM\_C, SNLM\_G and SNLM\_C methods. Through the

**Table 2** Number of training (train) and test (test) samples of the Indian Pines image and classification accuracy (in percent) for the SVM, EMP, AEAP, LMLL, BF, GF, NLM\_G, NLM\_C, SNLM\_G, and SNLM\_C (methods)

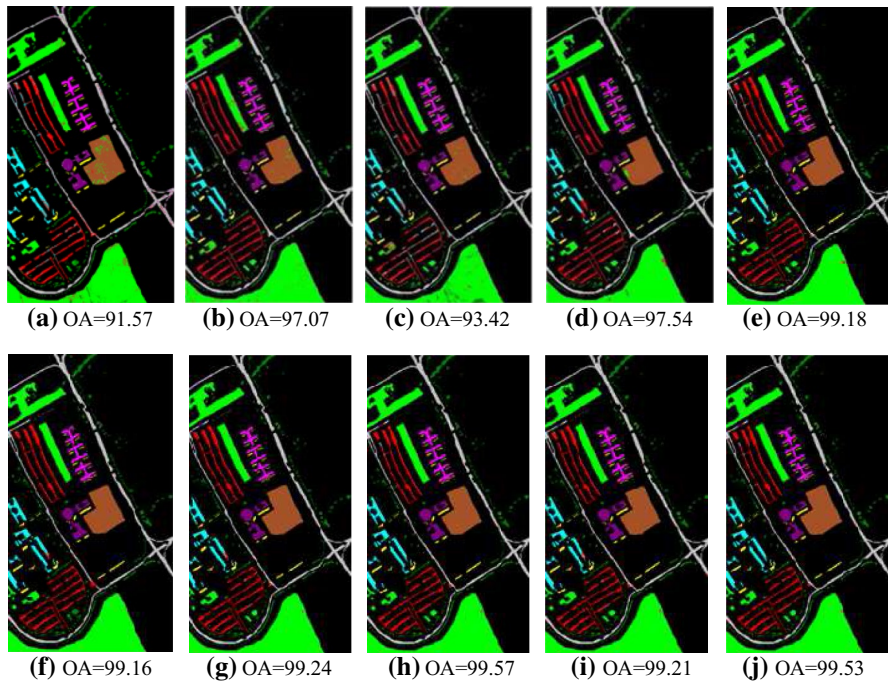
Class	Train	Test	SVM	EMP	AEAP	LMLL	BF	GF	NLM_G	NLM_C	SNLM_G	SNLM_C
Alfalfa	25	21	48.6	95.0	<b>100</b>	90.5	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Corn-N	83	1345	74.7	88.9	82.4	89.4	95.3	95.21	94.58	<b>95.39</b>	94.51	<b>95.39</b>
Corn-M	78	752	65.6	82.7	78.5	90.6	94.0	94.8	96.04	<b>96.27</b>	96.16	96.14
Corn	68	169	45.8	80.4	64.3	<b>99.4</b>	73.2	72.2	80.86	78.60	80.09	78.97
Grass-M	79	404	83.9	84.1	85.6	98.5	98.7	<b>99.0</b>	97.24	98.23	96.28	98.23
Grass-T	78	652	96.7	96.4	95.9	<b>100</b>	99.5	99.7	99.69	99.39	99.24	99.39
Grass-P	14	14	66.7	36.8	60.9	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Hay-W	66	412	99.0	99.5	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Oats	10	10	50.0	76.9	55.6	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Soybean-N	81	891	73.2	83.6	80.1	88.3	88.4	88.4	93.41	<b>94.90</b>	92.78	<b>94.9</b>
Soybean-M	99	2356	83.7	88.5	93.4	84.5	<b>98.4</b>	98.2	95.49	95.17	95.62	95.21
Soybean-C	73	520	75.0	85.6	85.4	<b>98.5</b>	94.3	92.9	97.9	97.91	96.79	97.36
Wheat	70	135	95.7	<b>100</b>	98.5	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Woods	90	1175	95.2	99.2	98.4	91.6	99.6	99.4	<b>100</b>	99.65	99.91	99.65
Buildings	65	321	64.0	97.5	75.8	<b>99.4</b>	85.1	85.1	87.19	91.69	88.89	91.69
Stone	46	47	90.2	97.9	90.4	<b>100</b>	90.4	92.2	93.75	97.92	93.88	97.92
OA			79.81	89.7	87.67	91.31	95.42	95.29	95.88	<b>96.22</b>	95.76	96.19
AA			75.50	87.06	84.09	95.66	94.81	94.81	96.01	<b>96.57</b>	95.88	96.55
Kappa			76.99	88.16	85.92	90.07	94.75	94.60	95.25	<b>95.64</b>	95.12	95.62

Bold values indicate the best classification result

**Table 3** Number of training (train) and test (test) samples of the university of Pavia image and classification accuracy (in percent) for the SVM, EMP, AEAP, LMLL, BF, GF, NLM\_G, NLM\_C, SNLM\_G, and SNLM\_C (methods)

Class	Train	Test	SVM	EMP	AEAP	LMLL	BF	GF	NLM_G	NLM_C	SNLM_G	SNLM_C
Asphalt	286	6345	96.9	99.6	98.1	94.8	99.2	98.7	99.33	99.89	99.55	<b>99.9</b>
Meadows	286	18363	97.7	99	98.9	99	99.8	99.7	99.78	<b>99.91</b>	99.81	99.89
Gravel	285	1814	75.0	96.5	73	90.1	94.9	98.6	<b>99.83</b>	98.31	99.49	98.52
Trees	285	2779	87.0	98.8	75.7	98.8	99.2	99.1	99.89	<b>100</b>	99.96	<b>100</b>
Sheets	285	1060	96.3	96.9	96.8	<b>100</b>	99.4	99	99.91	99.91	99.81	99.91
Soil	285	4744	81.4	86	90.1	98.5	96.4	98.6	98.3	<b>99.31</b>	97.81	99.02
Bitumen	285	1045	67.5	95.8	96.0	98.9	<b>100</b>	<b>100</b>	98.21	97.57	98.12	97.48
Bricks	285	3397	88.8	<b>98.5</b>	90.5	94.7	95.2	97.8	96.65	98.31	96.61	98.22
Shadows	285	662	<b>100</b>	<b>100</b>	<b>100</b>	99.2	99.8	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
OA			91.57	97.07	93.42	97.54	99.18	99.16	99.24	<b>99.57</b>	99.21	99.53
AA			87.85	96.79	91.02	96.67	99.05	99.05	99.1	<b>99.24</b>	99.02	99.22
Kappa			88.72	96.05	91.21	97.12	98.89	98.87	98.97	<b>99.42</b>	98.93	99.36

Bold values indicate the best classification result

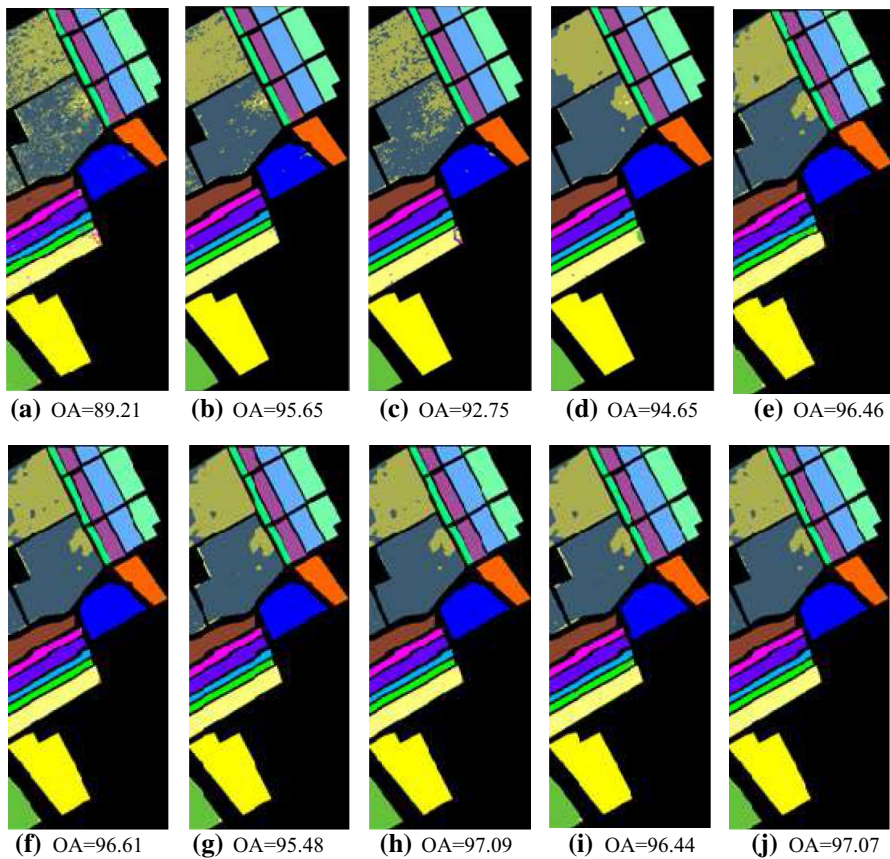


**Fig. 14** Classification results (University of Pavia image) obtained by (a) the SVM method (full feature space), (b) the EMP method, (c) the AEAP method, (d) the LMLL method, (e) the BF method, (f) the GF method, (g) the NLM\_G method, (h) the NLM\_C method, (i) the SNLM\_G method, and (j) the SNLM\_C method. The value of OA is given in percent

idea of integral graph, the NLM computational complexity is  $O(ND^2)$ , in which  $N$  is the number of image pixels and  $D$  is the size of search window. The proposed algorithm can realize a rapid classification of ground-object targets. As we can see, the proposed methods require much less computational time, when compared with other methods such as SVM, EMP, LMLL, BF and GF (most of them are more than 100 s), which demonstrates the high efficiency of the classification strategy.

## 5 Conclusion

This paper proposes a kind of spectral–spatial classification methods via the NLM filtering. The main contribution in this paper is that the proposed methods use the principal component analysis to extract spatial information of HSI, and further recognize detailed spatial characteristic of the ground-objects identification according to the NLM filtering. Experiments are conducted on three typical



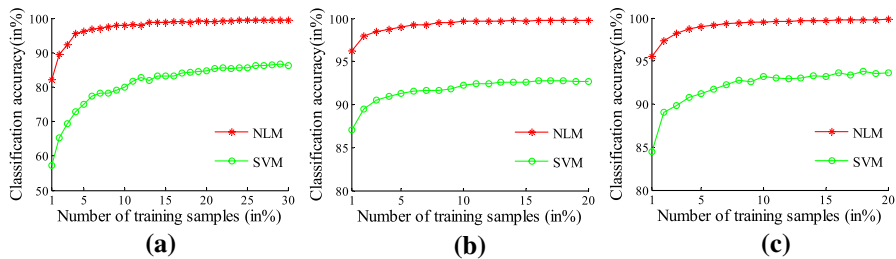
**Fig. 15** Classification results (Salinas image) obtained by (a) the SVM method(full feature space), (b) the EMP method, (c) the AEAP method, (d) the LMLL method, (e) the BF method, (f) the GF method, (g) the NLM\_G method, (h) the NLM\_C method, (i) the SNLM\_G method, and (j) the SNLM\_C method. The value of OA is given in percent

databases. According to the experimental results, the proposed methods can effectively combine the spectral and spatial information by using the NLM filtering, which can make the contour of ground-object consistent with the real ground-object contour, and remove the mis-classification which is similar to the pepper and salt noise. Furthermore, the computational burden can be reduced with the proposed methods. Moreover, compared with some common spectral–spatial classification algorithms, the proposed filter algorithm based on the NLM has a better treatment effect and can effectively realize the ground-object classification of HSI accurately.

**Table 4** Number of training (train) and test (test) samples of the Salinas image and classification accuracy (in percent) for the SVM, EMP, AEAP, LMLL, BF, GF, NLM\_G, NLM\_C, SNLM\_G, and SNLM\_C (methods)

Class	Train	Test	SVM	EMP	AEAP	LMLL	BF	GF	NLM_G	NLM_C	SNLM_G	SNLM_C
Weeds_1	67	1942	99.9	99.9	<b>100</b>	99.9	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Weeds_2	67	3659	99.6	99.8	<b>100</b>	99.6	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Fallow	67	1909	95.7	96.7	98.3	95.7	98.7	98.9	98.91	<b>99.27</b>	99.17	99.17
Fallow_P	69	1325	97.5	97.7	97.8	97.5	97.4	97.4	99.02	<b>99.85</b>	99.02	99.77
Fallow_S	67	2611	98.3	98.1	98.7	98.3	99.2	99.4	99.62	<b>100</b>	99.32	<b>100</b>
Stubble	67	3892	<b>100</b>	99.9	99.9	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	99.95	<b>100</b>
Celery	68	3511	99.0	99.7	85.6	99.0	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Graps	69	11202	81.9	92.1	<b>99.3</b>	81.9	95.8	96.1	93.71	93.29	93.72	93.29
Soil	68	6135	99.5	99.6	96.4	99.5	99.6	99.7	98.52	<b>100</b>	98.46	<b>100</b>
Corn	68	3210	87.3	94.1	99.4	87.3	96.0	95.5	94.34	<b>99.53</b>	94	99.38
Lettuce_4wk	68	1000	96.6	<b>100</b>	99.4	96.6	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Lettuce_5wk	67	1860	96.9	96.4	94.5	96.9	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Lettuce_6wk	67	849	96.9	96.8	99.9	96.9	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Lettuce_7wk	67	1003	89.2	96.4	96.9	89.2	99.1	89.3	99.3	<b>100</b>	99.1	<b>100</b>
Vinyard_U	70	7198	64.2	88.8	75.3	64.2	84.9	85.5	88.4	<b>89.39</b>	88.4	<b>89.39</b>
Vinyard_T	67	1740	95.0	93.7	97.5	95.0	99.9	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
OA			89.21	95.65	92.75	94.65	96.46	96.61	96.48	<b>97.09</b>	96.44	97.07
AA			93.59	96.86	96.21	97.44	98.17	98.24	98.24	<b>98.83</b>	98.2	98.81
Kappa			88.01	95.15	91.92	94.04	96.05	96.23	96.08	<b>96.75</b>	96.03	96.73

Bold values indicate the best classification result



**Fig. 16** OA of the proposed NLM\_G method with different numbers of train samples on different images: **(a)** Indian Pines. **(b)** University of Pavia. **(c)** Salinas

**Table 5** Computing time (in seconds) of the proposed algorithms

Data sets	SVM	EMP	LMLL	BF	GF	NLM_G	NLM_C	SNLM_G	SNLM_G
Indian Pines	140.59	37.35	71.16	143.63	145.77	25.38	26.75	25.40	26.87
University	256.12	120.96	549.51	235.25	259.05	68.93	84.48	69.01	84.68
Salinas	130.64	29.01	175.37	130.81	131.36	33.16	39.58	33.29	39.79

**Acknowledgements** This work was supported by the National Natural Science Foundation of China under Grant 51704115, by the Key Laboratory Open Fund Project of Hunan Province University under Grants 17K040 and 15K051, by the Hunan Provincial Natural Science Foundation under Grant 2016JJ2064, by the Fund of Education Department of Hunan Province under Grant 16C0723, and by the Science and Technology Program of Hunan Province under Grant 2016TP1021. The authors would like to thank the Dr. S. Li and the reviewers for their insightful comments and suggestions which have greatly improved this work.

## References

1. Kanning, M., Siegmann, B., & Jarmer, T. (2016). Regionalization of uncovered agricultural soils based on organic carbon and soil texture estimations. *Remote Sensing*, 8(11), 927.
2. Chi, J., & Crawford, M. M. (2014). Spectral unmixing-based crop residue estimation using hyperspectral remote sensing data: A case study at Purdue university. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), 2531–2539.
3. Clark, M. L., & Roberts, D. A. (2012). Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. *Remote Sensing*, 4(6), 1820–1855.
4. Yang, S., Qiao, Y., Yang, L., Jin, P., & Jiao, L. (2014). Hyperspectral image classification based on relaxed clustering assumption and spatial laplace regularizer. *IEEE Geoscience and Remote Sensing Letters*, 11(5), 901–905.
5. Ma, L., Crawford, M. M., & Tian, J. (2010). Local manifold learning-based-nearest-neighbor for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11), 4099–4109.
6. Goel, P. K., Prasher, S. O., Patel, R. M., Landry, J. A., Bonnell, R. B., & Viau, A. A. (2003). Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. *Computers and Electronics in Agriculture*, 39(2), 67–93.
7. Chen, C., & Ho, P. G. P. (2008). Statistical pattern recognition in remote sensing. *Pattern recognition*, 41(9), 2731–2741.

8. Bali, N., & Mohammad-Djafari, A. (2008). Bayesian approach with hidden Markov modeling and mean field approximation for hyperspectral data analysis. *IEEE Transactions on Image Processing*, 17(2), 217–225.
9. Yang, H. (1999). A back-propagation neural network for mineralogical mapping from AVIRIS data. *International Journal of Remote Sensing*, 20(1), 97–110.
10. Fauvel, M., Benediktsson, J. A., Chanussot, J., & Sveinsson, J. R. (2008). Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11), 3804–3814.
11. Heikkinen, V., Tokola, T., Parkkinen, J., Korpela, I., & Jaaskelainen, T. (2010). Simulated multi-spectral imagery for tree species classification using support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3), 1355–1364.
12. Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259.
13. HUGES
14. Prasad, S., & Bruce, L. M. (2008). Limitations of principal components analysis for hyperspectral target recognition. *IEEE Geoscience and Remote Sensing Letters*, 5(4), 625–629.
15. Villa, A., Benediktsson, J. A., Chanussot, J., & Jutten, C. (2011). Hyperspectral image classification with independent component discriminant analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 49(12), 4865–4876.
16. Falco, N., Benediktsson, J. A., & Bruzzone, L. (2014). A study on the effectiveness of different independent component analysis algorithms for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), 2183–2199.
17. Benediktsson, J. A., Pesaresi, M., & Amason, K. (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9), 1940–1949.
18. Benediktsson, J. A., Palmason, J. A., & Sveinsson, J. R. (2005). Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3), 480–491.
19. Falco, N., Benediktsson, J. A., & Bruzzone, L. (2015). Spectral and spatial classification of hyperspectral images based on ICA and reduced morphological attribute profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 53(11), 6223–6240.
20. Erturk, A., Gullu, M. K., & Erturk, S. (2013). Hyperspectral image classification using empirical mode decomposition with spectral gradient enhancement. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5), 2787–2798.
21. Demir, B., & Bruzzone, L. (2016). Histogram based attribute profiles for classification of very high resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(4), 2096–2107.
22. Tarabalka, Y., Fauvel, M., Chanussot, J., & Benediktsson, J. A. (2010). SVM-and MRF-based method for accurate classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 7(4), 736–740.
23. Zhang, B., Li, S., Jia, X., Gao, L., & Peng, M. (2011). Adaptive Markov random field approach for classification of hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 8(5), 973–977.
24. Li, J., Huang, X., Gamba, P., Bioucas-Dias, J. M., Zhang, L., Benediktsson, J. A., et al. (2015). Multiple feature learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 53(3), 1592–1606.
25. Li, J., Bioucas-Dias, J. M., & Plaza, A. (2013). Spectral-spatial classification of hyperspectral data using loopy belief propagation and active learning. *IEEE Transactions on Geoscience and Remote Sensing*, 51(2), 844–856.
26. Kang, X., Li, S., & Benediktsson, J. A. (2014). Spectral-spatial hyperspectral image classification with edge-preserving filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5), 2666–2677.
27. Chen, Y., Nasrabadi, N. M., & Tran, T. D. (2011). Hyperspectral image classification using dictionary-based sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 49(10), 3973–3985.
28. Fang, L., Li, S., Kang, X., & Benediktsson, J. A. (2014). Spectral-spatial hyperspectral image classification via multiscale adaptive sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 52(12), 7738–7749.



29. Lu, T., Li, S., Fang, L., Ma, Y., & Benediktsson, J. A. (2016). Spectral spatial adaptive sparse representation for hyperspectral image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, 54(1), 373–385.
30. Zhang, H., Li, J., Huang, Y., & Zhang, L. (2014). A nonlocal weighted joint sparse representation classification method for hyperspectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), 2056–2065.
31. Liu, H., Yang, C., Pan, N., Song, E., & Green, R. (2010). Denoising 3D MR images by the enhanced non-local means filter for Rician noise. *Magnetic Resonance Imaging*, 28(10), 1485–1496.
32. Coup, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., & Barillot, C. (2008). An optimized block-wise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Transactions on Medical Imaging*, 27(4), 425–441.
33. Fontes, F. P. X. D., Barroso, G. A., Coup, P., & Hellier, P. (2011). Real time ultrasound image denoising. *Journal of Real-Time Image Processing*, 6(1), 15–22.
34. Rahman, M. M., Arefin, M. G., PK, M. K., & Uddin, D. M. S. (2012). Optimal nonlocal means algorithm for denoising ultrasound image. *Computer Engineering and Intelligent Systems*, 3(3), 56–64.
35. Kelm, Z. S., Blezek, D., Bartholmai, B., & Erickson, B. J. (2009). Optimizing non-local means for denoising low dose CT. In *Proceedings of the IEEE ISBI* (pp. 662–665).
36. Froment, J. (2014). Parameter-free fast pixelwise non-local means denoising. *Image Processing (Online)*, 4, 300–326.
37. Mahmoudi, M., & Sapiro, G. (2005). Fast image and video denoising via nonlocal means of similar neighborhoods. *IEEE Signal Processing Letters*, 12(12), 5839–842.
38. Wang, J., Guo, Y., Ying, Y., Liu, Y., & Peng, Q. (2006). Fast non-local algorithm for image denoising. In *Image Processing (IEEE ICIP)* (pp. 1429–1432).
39. Jia, M., Gong, M., Zhang, E., Li, Y., & Jiao, L. (2014). Hyperspectral image classification based on nonlocal means with a novel class-relativity measurement. *IEEE Geoscience and Remote Sensing Letters*, 11(7), 1300–1304.
40. Buades, A., Coll, B., & Morel, J. M. (2005). A non-local algorithm for image denoising. In *Image Processing (CVPR)* (pp. 60–65).
41. Zhang, W. G., & Zhang, Q. (2011). SAR image despeckling combining target detection with improved nonlocal means. *Electronics Letters*, 47(12), 724–725.
42. Buades, A., Coll, B., & Morel, J. M. (2005). A non-local algorithm for image denoising. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2(7), 60–65.
43. Gonzalez, R. C., & Woods, R. E. (2001). *Digital Image Processing* (2nd ed.). Boston, MA: Addison-Wesley.
44. Farbman, Z., Fattal, R., Lischinski, D., & Szeliski, R. (2008). Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics (TOG)*, 27(3), 67.
45. Gastal, E. S. L., & Oliveira, M. M. (2008). Domain transform for edge-aware image and video processing. *ACM Transactions on Graphics (TOG)*, 30(4), 1.
46. Marpu, P. R., Pedergrana, M., Mura, M. D., Benediktsson, J. A., & Bruzzone, L. (2013). Automatic generation of standard deviation attribute profiles for spectral spatial classification of remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 10(2), 293–297.