CrossMark

Spectral–Spatial Hyperspectral Image Classification Based on KNN

Kunshan Huang[1] · **Shutao Li**[1] · **Xudong Kang**[1] ·
Leyuan Fang[1]

Abstract   Fusion of spectral and spatial information is an effective way in improving the accuracy of hyperspectral image classification. In this paper, a novel spectral–spatial hyperspectral image classification method based on K nearest neighbor (KNN) is proposed, which consists of the following steps. First, the support vector machine is adopted to obtain the initial classification probability maps which reflect the probability that each hyperspectral pixel belongs to different classes. Then, the obtained pixel-wise probability maps are refined with the proposed KNN filtering algorithm that is based on matching and averaging nonlocal neighborhoods. The proposed method does not need sophisticated segmentation and optimization strategies while still being able to make full use of the nonlocal principle of real images by using KNN, and thus, providing competitive classification with fast computation. Experiments performed on two real hyperspectral data sets show that the classification results obtained by the proposed method are comparable to several recently proposed hyperspectral image classification methods.

This article is part of the Topical Collection on Hyperspectral Imaging and Image Processing.

✉ Shutao Li
    shutao_li@hnu.edu.cn

[1]   College of Electrical and Information Engineering, Hunan University, Changsha 410082, China

⌷ Springer

1 Introduction

Hyperspectral sensors capture more than one hundred spectral bands which provide rich spectral information regarding the physical nature of different materials. For instance, the Airborne Visible-Infrared Imaging Spectrometer (AVRIS) system can capture 224 spectral channels with a spectral resolution of around 10 nm, covering the wavelength from 0.4 to 2.5 μm. The wide spectral coverage and fine resolution of the hyperspectral data provide the capability to distinguish objects in the image more accurately. However, it also presents challenges to image classification due to the high dimensionality of the data. Specifically, traditional image processing tools for the analysis of gray-level or color images may be not appropriate for hyperspectral images. For instance, Hughes phenomenon may be produced for classification due to the well known curse of dimensionality, which means that the accuracy of classification algorithms may decrease significantly while the dimensionality of the data increases [1]. In order to make full use of the rich spectral information provided by the high spectral dimension, many different hyperspectral image classification algorithms have been developed in recent years [2].

In the first beginning, many algorithms were designed to classify each pixel of the hyperspectral image based on its spectrum only [2]. These methods are known as pixel-wise methods which can be divided into two categories: spectral feature extraction and spectral classification. The spectral feature extraction process only aims at reducing the spectral dimension of the data using linear and nonlinear transformations such as principal component analysis (PCA) [3] and independent component analysis (ICA) [4]. In addition to spectral feature extraction, spectral classification methods such as Bayesian estimation techniques [5], neural networks [6], decision trees [7], and genetic algorithms [8] also have been investigated to learn the class distributions in high-dimensional spaces by inferring the no-linear boundaries between classes in feature space. Among these methods, support vector machines [9] has shown robust classification performances when a limited number of training samples is available.

In recent years, it is found that the integration of spectral and spatial information in the image analysis can further improve the classification results. Specifically, a hyperspectral pixel is classified based on both the feature vector of this pixel and feature values extracted from the pixel's neighborhood. Morphological filters [10, 11] and other types of local filtering approaches [12–15] have been investigated to develop novel spatial feature extraction and classification methods. Zhang et al. [16, 17] investigated several frameworks which aim at combing multiple features to improve the classification accuracy. These methods have been demonstrated to show promising results in terms of classification accuracies. However, local processing techniques such as the recently proposed edge-preserving filtering based method [15] only consider the local neighborhoods. Although the local neighborhoods can be defined using different scales of filtering operations, this kind of methods cannot make full use of the deep and global spatial correlations among hyperspectral pixels.

Another approach which can make full of the spatial information is based on image segmentation [18]. Segmentation based classification usually consists of the following two steps: First, the hyperspectral image is segmented into non-overlapping homogeneous regions. Then, the classification result is obtained based on the pixel-wise classification, followed by major voting within the segmented regions. To make this approach applicable, accurate and automatic hyperspectral image segmentation is required. Different techniques have been successfully applied for hyperspectral image segmentation, such as watershed [19], partitional clustering [20], and hierarchical segmentation [21]. Although these approaches can usually lead to an improvement of classification accuracy, the segmentation algorithm may be time consuming.

In this paper, a novel spectral–spatial hyperspectral image classification method is introduced based on KNN searching in a novel feature space [22]. The main contributions of the paper are twofold: The first contribution is the extension of KNN searching for the non-local filtering of images which can make full use of the spatial correlation among adjacent pixels. The second contribution is the extension of the KNN based filtering algorithm to spectral–spatial hyperspectral image classification. Specifically, the KNN based filtering algorithm is used to refine the initial probability maps obtained by pixel-wise classifier. The resulting classification map is obtained by assigning each pixel with the label which gives the highest probability. This probability optimization based scheme is similar to our previous works which optimize the probabilities by using edge-preserving filtering [15] and extended random walkers [23] (a global optimization method). In comparison with the two methods, the major advantage of the proposed KNN method is that it can make full use of the non-local spatial information of the hyperspectral image while does not need to solve a global energy optimization problem. In this work, it is shown that such a KNN based non-local filtering scheme is able to improve the classification accuracies effectively. Experiments performed on two real hyperspectral data sets demonstrate the effectiveness of the proposed method.

The rest of this paper is organized as follows. Section 2 describes the proposed KNN based image filtering algorithm. Section 3 introduces the proposed KNN based spectral-spatial classification method in detail. Section 4 gives the results and discussions. Finally, conclusions are given in Sect. 5.

2 KNN Based Non-local Image Filtering

The k-nearest-neighbor (KNN) classifier is one of the simplest and most widely used nonparametric classification methods. Although it has been successfully used for hyperspectral image classification, the KNN is usually utilized as a pixel-wise classifier in these researches which rely heavily on the optimal distance metric and feature space [24–26]. Different from these works, in this paper, KNN is used to search similar non-local pixels for image filtering rather than to achieve a direct classification of each pixel.

Rather than searching the nearest neighbors in the pixel value domain, the nonlocal principle can be implemented by computing the K nearest neighbors in the

feature space which includes both pixel value and spatial coordinates. Specifically, the feature vector $F(i)$ is defined as follows:

$$F(i) = (I(i), \lambda \cdot l(i), \lambda \cdot h(i)); \tag{1}$$

where $I(i)$ refers to the normalized pixel value, $l(i)$ and $h(i)$ refer to the normalized longitude and latitude of pixel i, respectively. λ controls the balance between pixel value and spatial coordinate in the KNN searching process. In order to do the KNN searching efficiently, the Fast Library for Approximate Nearest Neighbors (FLANN) is adopted to compute the K nearest neighbors in the defined feature space [27].

Generally, the KNN filtering process involves a guidance image I, an input image P, and an output image O. Both I and P are given beforehand according to the application. As shown in Fig. 1, P is one of the probability maps estimated with Support Vector Machines (SVM). I is the first principal component of the hyperspectral image. Given I and P, the KNN based nonlocal filtering method can be defined as follows:

$$O(i) = \frac{\sum P(j)}{K}, j \in \omega_i; \tag{2}$$

where ω_i refers to the K nearest neighbors of pixel i found in the feature space $F(i)$ defined in (1), $O(i)$ is the filtering output. As shown in Fig. 1, when $\lambda = 0$, the spatial distances between different pixels are not considered in the filtering operation, and thus, the KNN filtering cannot effectively transfer the spatial structures of I to P. By contrast, through modeling the spatial coordinates and pixel value in the same feature space, the spatial structures of the guidance image can be used to refine the boundaries in the input image (see Fig. 1d). This property makes it possible to apply the KNN filtering for spectral–spatial hyperspectral image classification.

3 Spectral–Spatial Hyperspectral Image Classification with KNN

In this section, the information about spatial structures defined by the KNN filtering algorithm mentioned above is used to improve the results of classification of a hyperspectral image. A probability optimization based spectral–spatial classification
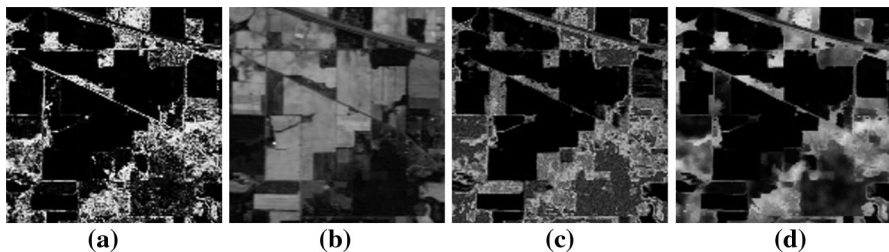

(a)   (b)   (c)   (d)

Fig. 1 An example of KNN filtering. **a** Input image P, **b** guidance image I, **c** filtering result O with $\lambda = 0$, **d** filtering result O with $\lambda = 3$
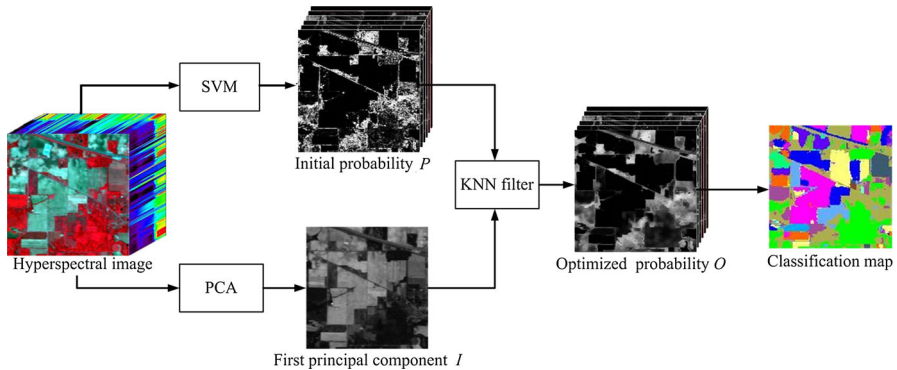
Fig. 2 A schematic of the proposed KNN based spectral–spatial hyperspectral image classification method

scheme is adopted here for hyperspectral images based on KNN filtering. The schematic of the proposed classification method is given in Fig. 2. Specifically, the proposed method consists of the following steps:

1. SVM classification: Given a d-dimensional hyperspectral image $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_i) \in \mathbb{R}^{d \times i}$ and τ training samples $T_\tau \equiv \{(\mathbf{x}_1, c_1), \ldots, (\mathbf{x}_\tau, c_\tau)\} \in (\mathbb{R}^d \times \mathcal{L}_C)^\tau$, a pixel-wise classification is performed on the hyperspectral image with pixel-wise SVM [28], where $\mathcal{L}_C = \{1, \ldots, N\}$ be a set of labels and N is the number of classes in the hyperspectral image.

2. KNN filtering: In this step, the principal component analysis method is first adopted to compute a one band representation of the hyperspectral image, i.e., I. Here, the first principal component is adopted to be the guidance image because it gives an optimal representation of the hyperspectral image in the mean squared sense, and thus, contains most of the salience information in the hyperspectral image (see Fig. 2). Then, based on (2), the proposed KNN filtering is performed on the the initial probability map P_n for each class n, with I serving as the guidance image. Finally, the classification result can be easily obtained by assigning each pixel with the label which gives the highest probability.

4 Experimental Results and Discussion

4.1 Experiments Performed on the Indian Pines Image

4.1.1 Data Set

In this experiment, the proposed classification algorithm is tested on a hyperspectral image of a rural area (the Indian Pines image). The Indian Pines image was

Table 1 Class names, number of training and test samples, global and class-specific classification accuracies in percentage for the Indian Pines image

Class		Samples		Classification methods					
No.	Name	Train	Test	SVM	EMP	LMLL	LBP	EPF	KNN
1	Corn-no till	10	36	75.01	79.31	95.01	91.31	100	100
2	Corn-min till	90	1338	76.49	84.79	84.13	91.52	97.95	96.39
3	Corn	79	751	77.28	89.31	89.15	87.57	99.11	95.61
4	S-N till	67	170	49.54	83.33	100	100	62.04	81.86
5	S-M till	71	412	92.16	96.58	99.03	98.31	97.87	97.62
6	S-C till	74	656	94.57	97.76	100	99.85	97.75	99.39
7	Alfalfa	10	18	91.67	76.47	100	100	100	100
8	G-pasture	70	408	99.27	100	100	100	100	99.55
9	G-trees	10	10	64.29	76.92	100	100	100	100
10	G-pasture-m	78	894	69.45	80.38	98.42	85.31	87.02	89.76
11	Hay	102	2353	84.54	95.85	88.64	93.98	95.02	96.11
12	Oats	75	518	72.95	84.09	98.66	96.56	93.91	95.72
13	Wheat	68	137	96.43	100	100	100	100	100
14	Woods	81	1184	95.38	99.41	90.23	100	99.39	99.73
15	Bldg-g	76	310	62.11	97.52	100	93.15	87.51	79.85
16	Towers	64	29	95.65	97.83	100	100	97.83	93.55
OA (%)		–	–	81.47	91.83	92.36	93.61	94.76	**95.39**
AA (%)		–	–	81.05	89.97	**96.45**	96.09	94.71	95.32
Kappa (%)		–	–	78.84	90.63	91.25	92.64	93.98	**94.71**
Time (s)		–	–	7.35	8.81	16.13	95.24	9.43	15.53

Bold numbers refer to the highest accuracies among different methods

collected by the AVIRIS sensor over the Indian Pines region in Northwestern Indiana in 1992. This scene, with a size of 145 by 145 pixels, and a spatial resolution of 20 m per pixel, was acquired over a mixed agricultural area. It is composed of 220 spectral channels in the wavelength range from 0.4 to 2.5 µm. Before classification, some spectral bands (no. 104–108, 150–163, and 220) were removed from the data set due to noise and water absorption, leaving a total of 200 spectral channels to be used in the following experiments. The Indian Pines data contains 16 classes, which are detailed in Table 1. For illustrative purposes, Fig. 3 shows the three band false color composite and the ground-truth map available for the scene. In order to test classification performances, 10 % of the samples were randomly selected from the reference data as training samples in our experiments.

4.1.2 Influence of Parameters to the Classification Performance

In this experiment, the influence of the two parameters λ and K to the performance of the proposed classification algorithm is analyzed. Figure 4 shows the classification maps and overall accuracies obtained by the proposed algorithm with
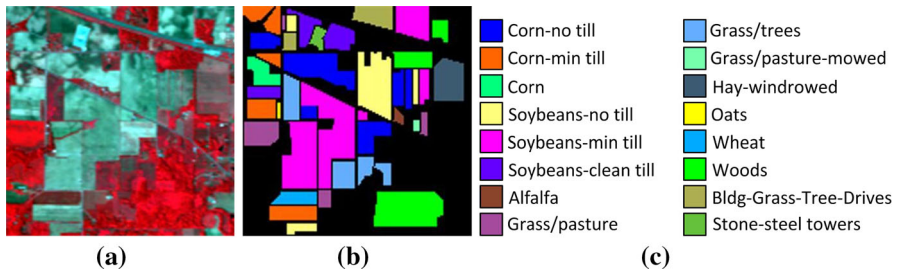
Fig. 3 Indian Pines data set: **a** three-band color composite; **b** ground-truth classification map; **c** Color code of different classes
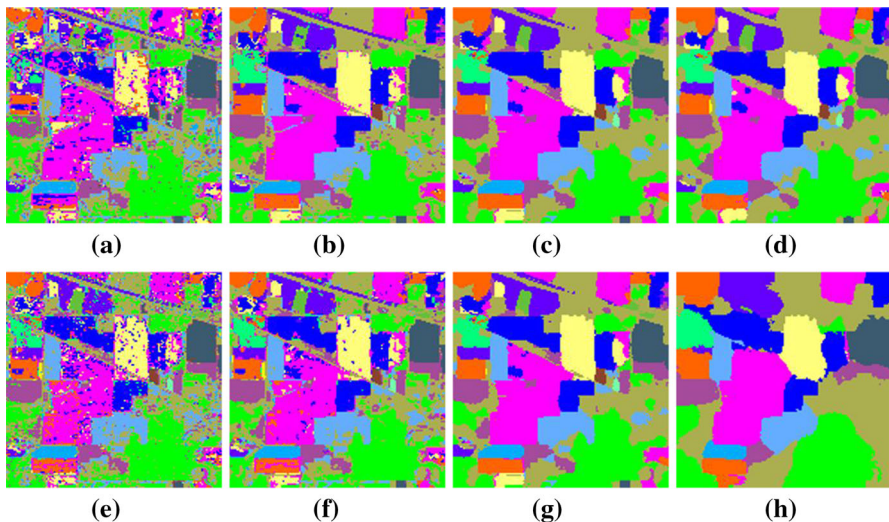

Fig. 4 Classification results obtained with **a** $\lambda = 0$ (OA = 71.79 %), **b** $\lambda = 1$ (OA = 94.29 %), **c** $\lambda = 5$ (OA = 96.24 %), **d** $\lambda = 100$ (OA = 95.44 %), **e** $K = 5$ (OA = 89.82 %), **f** $K = 10$ (OA = 92.61 %), **g** $K = 40$ (OA = 96.24 %), and **h** $K = 400$ (OA = 86.94 %) [K is fixed as 40 for (**a**–**d**), and λ is fixed as 5 for (**e**–**h**)]

different values of λ and K. From Fig. 4a, it can be seen that the result contains serious "noise" and the corresponding accuracy is quite low (OA = 71.19 %) when the coefficient λ is set to be 0. The reason is that the spatial coordinates are not considered in the KNN filtering operation when $\lambda = 0$. Furthermore, the classification result tends to be oversmoothed and the corresponding accuracy decreases when λ is very large. It means that pixel value and spatial coordinates are both important factors for the improvement of classification accuracy. The parameter K has similar influences to the classification result. For example, when K is quite large, the proposed filtering method may lead to the oversmooth of classification result, and thus, decreases the accuracy dramatically (OA = 86.94 %). However, when K is relatively small, it means that only a small number of non-local pixels are considered in the averaging operation. In this situation, the classification accuracy

also cannot be effectively improved (OA = 89.82 %). In this paper, $K = 40$ and $\lambda = 5$ are set to be the default parameters which gives the best performance in this experiment (OA = 96.24 %). In order to ensure the optimal of the two parameters, an adaptive setting scheme will be researched in the future.

4.1.3 Comparison of Different Classification Methods

Figure 5a–e shows the classification results obtained with the Support Vector Machines method (SVM) [28], the Extended Morphological Profiles method (EMP) [11], the logistic regression and multilevel logistic method (LMLL) [29], the loopy belief propagation method (LBP) method [30], the edge-preserving filtering method (EPF) [15], and the proposed K nearest neighbors based method (KNN), respectively. The SVM classification is performed with the Gaussian radial basis function (RBF) kernel, using the LIBSVM library [28]. The optimal parameters C and γ were determined by fivefold cross validation. The default parameters given in [15, 29, 30] are adopted for the LMLL, LBP and EPF methods. As shown in this figure, all spectral–spatial methods can reduce significantly the noise in the classification map, resulting in more homogeneous and meaningful regions in the classification map. For example, with the proposed KNN method, the classification result of pixel-wise SVM can be improved significantly, because of the noise reduction. In order to evaluate the improvement
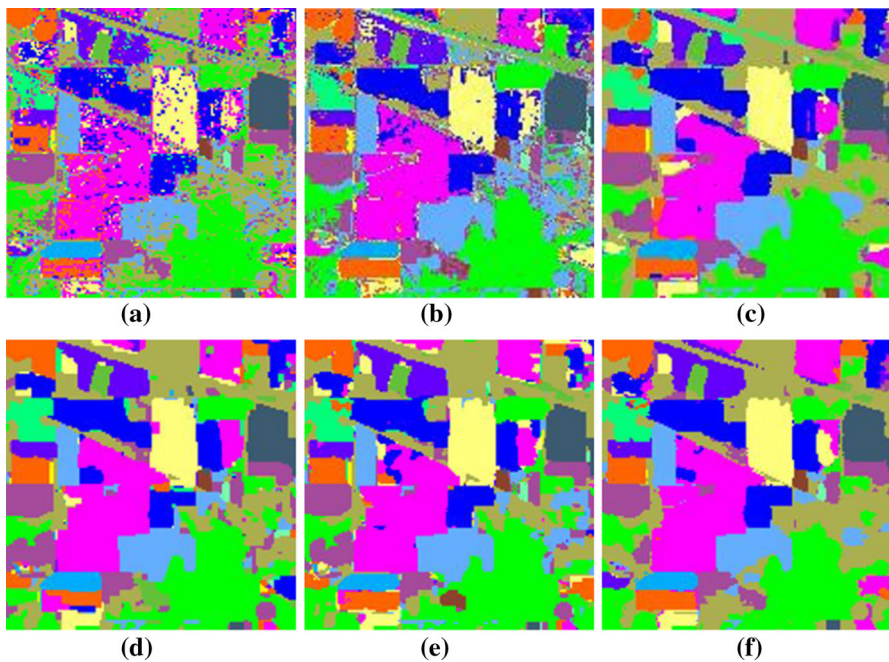

Fig. 5 Indian Pines data set: **a–e** classification map for the SVM method [28] (OA = 81.72 %), the EMP method [11] (OA = 91.83 %), the LMLL method [29] (OA = 92.88 %), the LBP method [30] (OA = 92.17 %), the EPF method [15] (OA = 95.48 %) and the proposed KNN method (OA = 96.23 %)

more objectively, the number of training and test samples, and the global and individual classification accuracies of different classification methods are presented in Table 1. Three measures of accuracy are used: (1) Overall accuracy (OA) which measures the percentage of correctly classified pixels; (2) Average accuracy (AA) which measures the mean of the percentage of correctly classified pixels for each class; (3) Kappa coefficient (kappa) which measures the percentage of agreement (correctly classified pixels) corrected by the number of agreements that would be expected purely by chance. Furthermore, the accuracies are calculated as an average after 10 repeated experiments. Table 1 shows that the proposed method can effectively improve the OA, AA, and Kappa of SVM. Furthermore, the individual classification accuracies are also improved by the proposed KNN method for almost all of the classes. For example, the accuracy of the corn-no till class has been improved from 75.01 to 100 %. Compared with the EMP, LMLL method, LBP method, and the EPF methods, the proposed KNN method gives a comparable performances for OA, AA, and Kappa. It means that the proposed KNN method can effectively improve classification accuracy.

4.2 Experiments Performed on the Botswana Image

4.2.1 Data Set

In this experiment, the proposed classification algorithm is tested on a hyperspectral image of a woodlands area (the Botswana image). The Botswana image was collected by the NASA EO-1 satellite over Okavango Delta, Botswana in May 31, 2001. This scene, with a size of 1476 by 256 pixels, and a spatial resolution of 30m
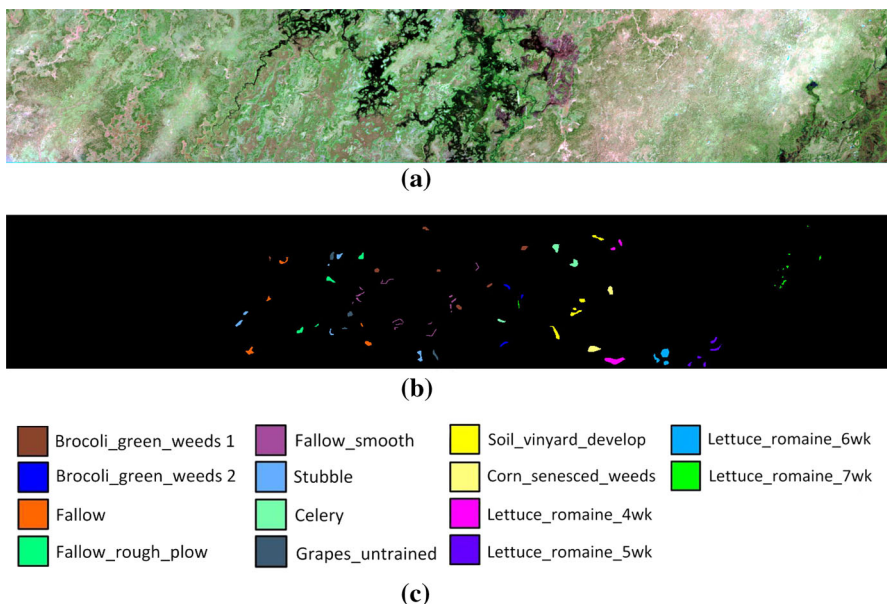

(a)

(b)

Brocoli_green_weeds 1	Fallow_smooth	Soil_vinyard_develop	Lettuce_romaine_6wk
Brocoli_green_weeds 2	Stubble	Corn_senesced_weeds	Lettuce_romaine_7wk
Fallow	Celery	Lettuce_romaine_4wk	
Fallow_rough_plow	Grapes_untrained	Lettuce_romaine_5wk	

(c)

Fig. 6 Botswana data set: **a** three-band color composite; **b** ground-truth classification map; **c** color code of different classes
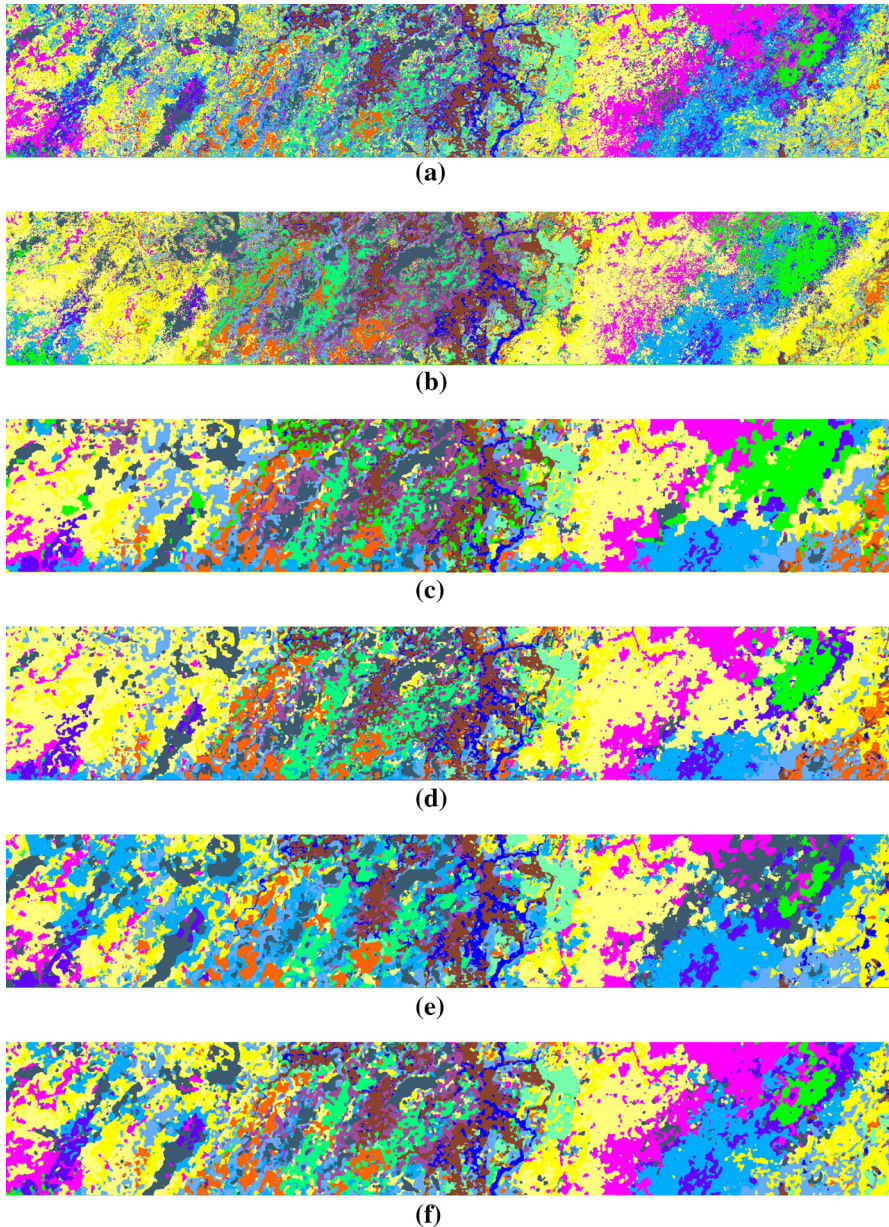
Fig. 7 Botswana data set: **a–e** classification map for the SVM method [28] (OA = 91.76 %), the EMP method [11] (OA = 96.34 %), the LMLL method [29] (OA = 97.37 %), the LBP method [30] (OA = 97.11 %), the EPF method [15] (OA = 97.13 %) and the proposed KNN method (OA = 98.81 %)

per pixel, was acquired to study the impact of flooding on vegetation in this area. It is composed of 242 spectral channels in the wavelength range from 0.4 to 2.5 μm. Before classification, uncalibrated and noisy bands that cover water absorption

features were removed, and the remaining 145 bands were included as candidate spectral features: [10–55, 82–97, 102–119, 134–164, 187–220]. The Botswana data consists of 14 identified classes representing the land cover types in seasonal swamps, occasional swamps, and drier woodlands, which are detailed in Table 1. For illustrative purposes, Fig. 6 shows the three band false color composite and the ground-truth map available for the study area. Similar to the experiments performed on the Indian Pines data set, 10 % of the samples were randomly selected from the reference data as training samples in our experiments.

4.2.2 Comparison of Different Classification Methods

Figure 7a–e shows the classification results obtained by the SVM, EMP, LMLL, LBP, EPF, and KNN methods, respectively. Furthermore, the number of training and test samples, and the global and individual classification accuracies for different methods are presented in Table 2. The accuracies are calculated as an average after 10 repeated experiments. It can be seen that the proposed method can effectively improve the OA, AA, and Kappa of SVM. Moreover, the KNN method shows the best classification performance in terms of OA, AA, and Kappa.

Table 2 Class names, number of training and test samples, global and class-specific classification accuracies in percentage for the Botswana image

Class		Samples		Classification methods					
No.	Name	Train	Test	SVM	EMP	LMLL	LBP	EPF	KNN
1	Water	23	247	99.61	100	100	100	99.61	99.41
2	Hippo	23	78	95.01	98.73	97.44	100	98.73	100
3	Grasses1	25	226	98.21	100	100	100	100	100
4	Grasses2	23	192	87.79	97.44	100	100	87.67	91.23
5	Reeds1	23	246	88.64	93.09	93.09	91.09	92.17	99.76
6	Riparian	23	246	80.24	93.83	88.21	79.69	98.28	97.32
7	F-scars	23	236	99.13	100	100	100	100	100
8	Island	23	180	97.78	100	100	100	100	100
9	W-lands	23	291	89.46	93.62	99.66	98.97	100	99.38
10	S-lands	23	224	88.84	86.33	100	100	94.12	98.71
11	Grasslands	24	282	94.91	98.48	93.24	93.26	99.26	97.89
12	S-mopane	23	158	93.71	98.67	100	100	100	100
13	M-mopane	23	245	94.31	96.85	100	96.34	99.19	99.88
14	E-soils	23	72	94.74	100	95.83	100	100	99.47
OA (%)		–	–	92.68	96.34	97.57	96.49	97.64	**98.32**
AA (%)		–	–	93.02	96.87	97.37	97.11	97.79	**98.43**
Kappa (%)		–	–	92.06	96.03	97.68	96.19	97.44	**98.58**
Time (s)		–	–	32.17	41.53	93.42	1569.99	40.91	57.49

Bold numbers refer to the highest accuracies among different methods

5 Conclusions

Although hyperspectral imaging provides rich spectral information, increasing the capability to distinguish different objects in a scene, the large number of spectral channels presents challenges to image classification. Instead of processing each pixel independently without considering information about spatial structures, the proposed KNN based image filtering algorithm can incorporate spatial information into classifier, and thus, the pixel-wise classification accuracies can be improved significantly, especially in areas where structural information is important to distinguish between classes. The proposed method has two main contributions: First, a simple yet effective feature vector construction methodology combining the values and spatial coordinates of different pixels is applied for the joint filtering of images. Second, the proposed KNN based filtering algorithm is applied for spectral–spatial hyperspectral image classification. In the experiments, it was shown that the proposed spectral–spatial classifier can lead to competitive classification accuracies when compared to other previously proposed spectral-spatial classification techniques. In conclusion, the proposed KNN based classification method succeeded in taking advantage of spatial and spectral information simultaneously.

References

1. Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, *14*(1), 55–63.
2. Plaza, A., Benediktsson, J. A., Boardman, J. W., Brazile, J., Bruzzone, L., Camps-Valls, G., et al. (2009). Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, *113*(Supplement 1), S110–S122.
3. 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.
4. 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.
5. Jackson, Q., & Landgrebe, D. A. (2002). Adaptive bayesian contextual classification based on Markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(11), 2454–2463.
6. Ratle, F., Camps-Valls, G., & Weston, J. (2010). Semisupervised neural networks for efficient hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(5), 2271–2282.
7. Bittencourt, H.R., de Oliveira Moraes, D.A., & Haertel, V. (2007). A binary decision tree classifier implementing logistic regression as a feature selection and classification method and its comparison with maximum likelihood. In *Proceedings of IEEE International Geoscience Remote Sensing Symposium* (pp. 1755–1758).
8. Stavrakoudis, D. G., Galidaki, G. N., Gitas, I. Z., & Theocharis, J. B. (2012). A genetic fuzzy-rule-based classifier for land cover classification from hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(1), 130–148.
9. Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(8), 1778–1790.

10. 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.

11. Plaza, A., Martinez, P., Plaza, J., & Perez, R. (2005). Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(3), 466–479.

12. Kang, X., Li, S., & Benediktsson, J. A. (2014). Spectral-spatial hyperspectral image feature extraction with image fusion and recursive filtering. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(6), 3742–3752.

13. 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.

14. Demir, B., & Erturk, S. (2010). Empirical mode decomposition of hyperspectral images for support vector machine classification. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(11), 4071–4084.

15. 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.

16. Zhang, L., Zhang, L., Tao, D., & Huang, X. (2012). On combining multiple features for hyperspectral remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(3), 879–893.

17. Zhang, L., Zhang, Q., Zhang, L., Tao, D., Huang, X., & Du, B. (2015). Ensemble manifold regularized sparse low-rank approximation for multiview feature embedding. *Pattern Recognition*, *48*(10), 3102–3112.

18. Fauvel, M., Tarabalka, Y., Benediktsson, J. A., Chanussot, J., & Tilton, J. C. (2013). Advances in spectral–spatial classification of hyperspectral images. *Proceedings of the IEEE*, *101*(3), 652–675.

19. Tarabalka, Y., Chanussot, J., & Benediktsson, J. A. (2010). Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recognition*, *43*(7), 2367–2379.

20. Tarabalka, Y., Benediktsson, J. A., & Chanussot, J. (2009). Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(8), 2973–2987.

21. Tarabalka, Y., Benediktsson, J. A., Chanussot, J., & Tilton, J. C. (2010). Multiple spectral–spatial classification approach for hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(11), 4122–4132.

22. Chen, Q., Li, D., & Tang, C. K. (2013). KNN matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(9), 2175–2188.

23. Kang, X., Li, S., Fang, L., Li, M., & Benediktsson, J. A. (2015). Extended random walker-based classification of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(1), 144–153.

24. Ma, L., Crawford, M. M., & Tian, J. (2010). Local manifold learning-based k-nearest-neighbor for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(11), 4099–4109.

25. Blanzieri, E., & Melgani, F. (2008). Nearest neighbor classification of remote sensing images with the maximal margin principle. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(6), 1804–1811.

26. Jia, X., & Richards, J. A. (2005). Fast k-NN classification using the cluster-space approach. *IEEE Geoscience and Remote Sensing Letters*, *2*(2), 225–228.

27. Muja, M., Lowe, D.G. (2009). Fast approximate nearest neighbors with automatic algorithm configuration. In *International Conference on Computer Vision Theory and Application* (pp. 331–340). INSTICC Press.

28. Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 27:1–27:27.

29. Li, J., Bioucas-Dias, J. M., & Plaza, A. (2011). Hyperspectral image segmentation using a new Bayesian approach with active learning. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(10), 3947–3960.

30. 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.