Research Article

Sparse Representation Classification Based on Flexible Patches Sampling of Superpixels for Hyperspectral Images

Haifeng Sima [iD],[1] Pei Liu,[2] Lanlan Liu,[1] Aizhong Mi [iD],[1] and Jianfang Wang[1]

[1]*The School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China*
[2]*The School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo, China*

Correspondence should be addressed to Aizhong Mi; miaizhong@hpu.edu.cn

Aiming at solving the difficulty of modeling on spatial coherence, complete feature extraction, and sparse representation in hyperspectral image classification, a joint sparse representation classification method is investigated by flexible patches sampling of superpixels. First, the principal component analysis and total variation diffusion are employed to form the pseudo color image for simplifying superpixels computing with (simple linear iterative clustering) SLIC model. Then, we design a joint sparse recovery model by sampling overcomplete patches of superpixels to estimate joint sparse characteristics of test pixel, which are carried out on the orthogonal matching pursuit (OMP) algorithm. At last, the pixel is labeled according to the minimum distance constraint for final classification based on the joint sparse coefficients and structured dictionary. Experiments conducted on two real hyperspectral datasets show the superiority and effectiveness of the proposed method.

1. Introduction

Remote sensing image classification has become an important part of remote sensing applications, which can be used in urban planning, environmental monitoring, classification, crop management, and many other applications [1–4]. Hyperspectral images (HSI) contain hundreds-dimensional spectrum vectors, which may bring to higher accuracy for land cover recognition and classification. Therefore, hyperspectral remote sensing image classification has always been the concerning focus of researchers. At the same time, the supervised classification technique has been proved to be a more proper method for remote sensing classification [5]. It inputs a small number of representative marked areas as the training samples to training discriminate function and classifier and then computes the statistical characteristic of unlabeled samples and compares with labeled samples for predicting classification. Thus supervised learning is more efficient and realizes significantly improvement on the classification accuracy [6]. With the development of pattern recognition and the deepening research of compressed sensing, the support vector machine (SVM) [7], Bayesian [8], logistic regression

[9], and manifold learning[10] have achieved more ideal classification results. In order to improve the efficiency of data processing and the accuracy of classification, researchers commonly used two of many strategies: one is feature dimension reduction or optimal spectral feature selection before classification, such as principal components analysis [11], spectral derivative features [12], and context information [13]. The second is to establish optimized classification model, such as kernel optimization [14] and ensemble learning [15]. On the other hand, with the further research of compressed sensing, sparse classification based on compressed sensing has been widely concerned. It was firstly applied to face detection and the classification can be achieved by sparse modeling with minimum reconstruction error, which has brought outstanding enhancement in classification accuracy. Sparse representation technology has been applied in various fields of computer vision pattern recognition, such as image segmentation, image restoration, super resolution, and face recognition.

Recently, sparse representation has been used for hyperspectral image classification, and achieved certain results [16–21]. The sparse representation classification method maps

high-dimensional signals into a few combinations of dictionary atoms and their coefficients. This method can extract data source features and describe category information effectively while removing noise. It can achieve more accurate classification based on minimum reconstruction error. Chen et al. [17] proposed that an unknown pixel is expressed as a sparse vector whose nonzero entries correspond to the weights of the selected training samples. The sparse vector is recovered by solving a sparsity-constrained optimization problem, and it can directly determine the class label of the test pixel. Castrodad et al. [18] proposed a sparse representation method at the subpixel level based on the learned block-structured discriminative dictionaries. Chen et al. [22] proposed a sparse representation model based on sparsely representing a test sample in terms of all of the training samples in a feature space induced by a kernel function. Srinivas et al. [19] learn a discriminative graph-based classifier that captures interclass information for sparse representation vectors of each pixel in the local spatial neighbourhood of a central pixel. In [20], Fang et al. proposed a multiscale adaptive sparse representation model for dictionary estimation and defined an optimizable adaptive sets based on residual matrix to estimate sparse coefficients to determine the classification results, and the classification accuracy is improved a lot.

How to use context information is the key technology for accurate classification of hyperspectral images. In addition to the above neighbourhood information fusion methods, superpixel segmentation has become an important technology for features extraction and optimization in numerous applications of computer vision and digital image processing. The superpixels segmentation provides homogeneous regions of original image, and the complex structure of image is compressed and simplified for further analysis and processing. Based on this, Feng et al. [23] proposed a basic assumption that interior pixels of superpixels have similar markers, namely, that the sparse coefficients of homogeneous pixels have similar structure, and the pixel classification is regulated by contextual information exploited from superpixel with decision rule of majority voting. This method classifies the superpixels as a whole, and the overall classification effect is not fine enough. Zhang et al. [21] integrate spectral and spatial information into group sparse coding (GSC) via clusters which is an adaptive spatial partition derived from the mean-shift superpixels.

All these methods use the spatial context information in different levels and put forward effective sparse model hypothesis. The spectral-spatial methods are still suffer from the selection of adjacent region scales; overcomplete feature extraction and sparse representation model are still more arbitrary. They cannot provide accurate classification of boundary pixels and small regions in the image. To overcome these limitations, the authors consider that the superpixel segmentation can provide homogeneous regions of original image, and the complex structure of image is compressed and simplified for further analysis and processing.

Based on the above analyses, a joint sparse representing classification method is proposed based on flexible patches sampling of superpixels (SRC-FPSS). The flowchart of SRC-FPSS method is shown in Figure 1. Firstly, a group of relatively complete homogeneous superpixels are computed for reliable contextual information for sparse recovery. For each test pixel, its context information can be extracted from its corresponding superpixel. So we are sampling a set of pixels from inner pixels of corresponding superpixel sorted by similarity measure. Then, all the neighbour patches of sampling pixels assumes share a common sparsity pattern and the sparse coefficient of test pixel can be estimated by solving a sparse optimization problem. Then, the class label of the test pixel can be determined by the characteristics of the minimal total residuals.

The remainder of this paper is arranged as follows: the Section 2 is the detailed description of the proposed SRC-FPSS algorithm. Experimental results are presented in Section 3, and finally, conclusions are given in Section 4.

2. Proposed Method

Firstly, we use the PCA model to extract first three main components to composite the pseudo color image. Then, the simple linear iterative clustering (SLIC) [24] method is executed on total variation (TV) diffusion of the pseudo color image to compute superpixles. We assume that the test pixels have a great correlation with the internal pixels of superpixels, and the test pixels can be joint represented by them. The inner pixels are sampled at equal intervals according to a given sampling frequency N. All the sample pixels are extended to a set of patches to form the reconstruction matrix as joint spares representation of test pixels. At last the test pixels are classified according to the reconstruction error.

2.1. Superpixels Computing. Existing methods for superpixels computing always operate on low-dimensional feature space, such as nature images. These common superpixels segmentation methods may not be able to obtain a better result and time consuming when working on hyperspectral images contains hundreds-dimensional spectrums [25, 26], and the PCA method is an effective way to deal with the computation complexity of high-dimensional data [26]. Therefore, we choose the compression feature (first three principle components) to construct pseudo color image by principal component analysis. In order to deal with the complicated texture, many image enhancement strategies have been developed [27–33]. In this paper, we use the conduction function to implement the nonlinear coupling diffusion filtering on the pseudo color image, which is processed by the total variation model in [34]. The implicit smooth method is shown in formula (1):

$$\frac{\partial I_c(x, y, t)}{\partial t} = \text{div}\left(g\left(\sum_{i=1}^{C} |\nabla I_c| \right) |\nabla I_c| \right) \tag{1}$$

where $I_c(x, y, t)$ is a single band of the pseudo color image, div is the divergence of vertical and horizontal dimensions with four features $[I, I_x^2, I_x^2, I_x I_y]$, t is the iteration times, g is the gauss diffusion function, and ∇ denotes the gradient vector. The iteration time relies on the superpixels scale for proper smoothing of local area. After nonlinear diffusion
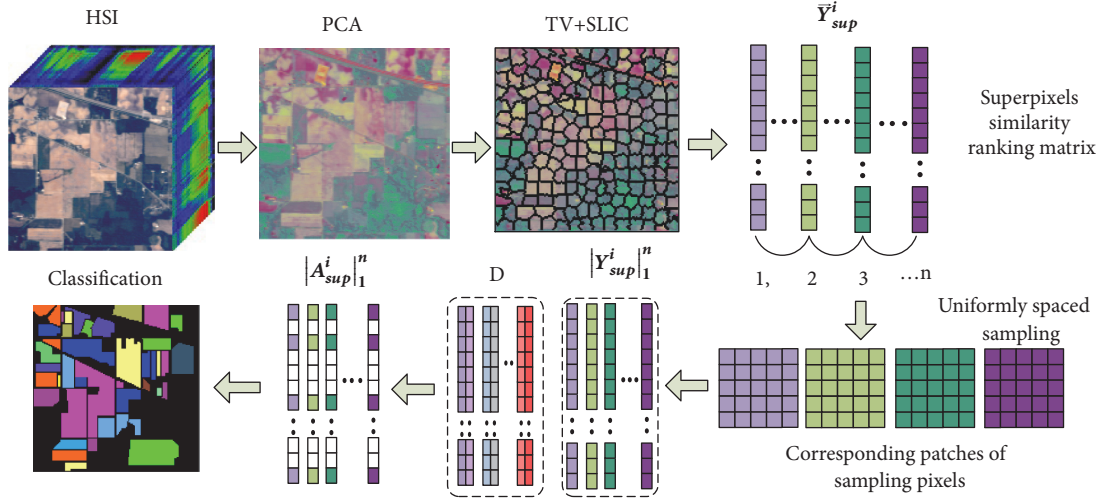
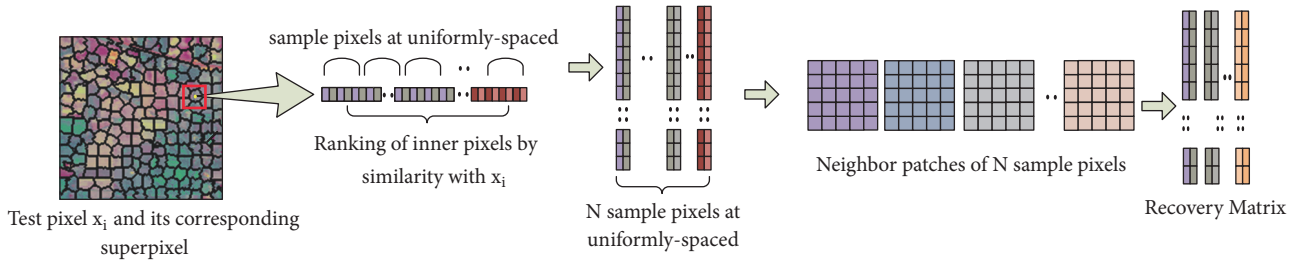FIGURE 1: The illustration of the proposed method.

FIGURE 2: The illustration of overcomplete sampling of superpixels for overcomplete recovery matrix.

filtering, the SLIC segmentation is carried out for superpixels computing. According to experience, larger superpixel is not proper for common feature computing and the inner pixels number of 150 lead to a good compactness and uniformity [35].

2.2. SRC-FPSS.
For the SRC algorithm, determining the specific categories of a test pixel relies on the residuals of the feature reconstruction computed by dictionaries and corresponding sparse vectors. It is necessary to update the sparse vector and the corresponding dictionary to compute the optimized sparse feature vector.

First, the dictionary $\mathbf{D}$ is initialized with a few labeled training samples collected from the image randomly within N classes. Select M labeled samples from original datasets according to the proportion randomly to set up a structured dictionary $\mathbf{D}$ with M atoms, and define the atomic index set $\mathbf{B}(null)$; thus a test pixel can be represented as a linear combination of $\mathbf{D}$ from all classes.

Based on the superpixels segmentation, for arbitrary test pixel $\mathbf{x}_i$, its corresponding superpixel can be arranged in a matrix $\mathbf{Y}_{sup}^i = [\mathbf{Y}_1, \dots \mathbf{Y}_2, \dots \mathbf{Y}_{N_p}] \in \mathbb{R}^{D * N_p}$. N_p is the pixels number of the superpixels and D is the band dimension of spectral space. All pixels within $\mathbf{Y}_{sup}^i$ provide complementary and associate information to the same test pixel $\mathbf{x}_i$. In order to get a joint area consist of similar materials with $\mathbf{x}_i$, we compute the similarity between $\mathbf{x}_i$ and inner pixels of $\mathbf{Y}_{sup}^i$ by formula (2) and sort them as $\overrightarrow{Y}_{sup}^i$ in descending order.

$$\|\mathbf{x}_i - \mathbf{x}_{inner}\| = \sum_{k=1}^{d} \|\mathbf{x}_i(k) - \mathbf{x}_{inner}(k)\| \qquad (2)$$

where k is the band counting as $= 1, 2 \dots D$. Next, we sample N pixels $[\mathbf{x}_i^1, \mathbf{x}_i^2, \dots \mathbf{x}_i^N]$ at uniformly spaced from $\overrightarrow{Y}_{sup}^i$ and choose their neighbour windows to construct overcomplete recovery matrix for $\mathbf{x}_i$, which can be employed to estimate the joint sparse vector for more accurate classification. The visual illustration of overcomplete sampling of superpixels for recovery matrix is shown in Figure 2. The scale of neighbour windows of sample pixels can be set as $S = \sqrt{N_p/2}$ (if S is even, $S = S + 1$). Then the joint representation can be expressed as

$$\mathbf{Y}_{sup}^i\Big|_1^N = \mathbf{D} A_{sup}^i\Big|_1^N \qquad (3)$$

$\mathbf{Y}_{sup}^i\Big|_1^N$ is the recovery matrix consisting of N neighbour windows for joint sparse representation of test pixel $\mathbf{x}_i$ and $\|A_{sup}^i\|_1^N = [\alpha^1, \dots \alpha^n \dots \alpha^N]$ is the joint sparse coefficient.
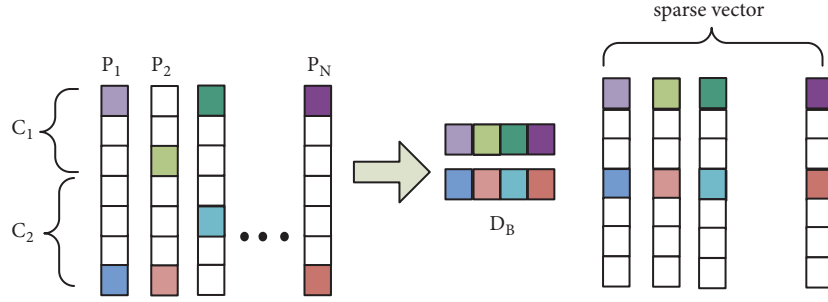
FIGURE 3: The illustration of joint sparse representation of two class.

Then the sparse coefficient can be computed by solving following optimization problem:

$$A^i_{sup}\big|^N_1 = \arg\min \left\| \mathbf{Y}^i_{sup}\big|^N_1 - \mathbf{D}A^i_{sup}\big|^N_1 \right\|_F$$

$$subject\ to\ \left\| A^i_{sup}\big|^N_1 \right\|_{row,0} \leq K_0 \tag{4}$$

where K_0 is the upper bound of sparse level, which means the maximum number of selected atoms in the dictionary. Given the redundant structure dictionary $\mathbf{D}$, solving the sparse representation vector of each pixel is the key for classification. The OMP algorithm concentrates on finding the most relevant atoms to the current residual signal and updating new representative atom set in each iteration process. When the dictionary atoms are selected, the test pixel is projected onto the spanned subspace of the selected atoms, then recomputed sparse coefficients, and updated the residual error until the termination condition is reached. The advantage of OMP algorithm is that it is able to select the best-matched atoms to given signal from the dictionary in each iteration of the approximation process. In order to enhance the impact of homogeneous regions for the sparse representation of the i-th iteration, it is necessary to calculate pixel residuals for all pixels in recover matrix and consist a correlation matrix $\mathbf{R}_M$. For any pixel i in superpixels $\mathbf{Y}^i_{sup}$, calculate the residual correlation matrix as

$$\mathbf{R}_M = \mathbf{D}^T\mathbf{Y}^i_{sup}\big|^N_1 \tag{5}$$

Select atoms with maximum correlation values in matrix $\mathbf{R}_M$ for each classes and patches, then sum the correlation for all N patches of each class as $V_1..V_C$ and select the max V and incorporate corresponding atoms indexes of each patches into the index set $\mathbf{B}$ of selected representative atoms. Thus, the joint sparse coefficient is estimated as

$$A^i_{sup}\big|^N_1 = \left(\mathbf{D}^T_B\mathbf{D}_B\right)^{-1}\mathbf{D}^T_B\mathbf{Y}^i_{sup}\big|^N_1 \tag{6}$$

The illustration of the joint sparse representation of two classes is displayed in Figure 3.

Update R_M with new sparse coefficients as follows:

$$\mathbf{R}_M = \mathbf{D}^T\left(\mathbf{Y}^i_{sup}\big|^N_1 - \mathbf{D}_BA^i_{sup}\big|^N_1\right) \tag{7}$$

Input:
  $x_i \in Y_{sup}, D = \{D_1, D_2, ...D_C\}, B(null), K,$
  while $iter < K$
(1) $\overrightarrow{Y}^i_{sup} \longleftarrow Y_{sup}$
(2) $\mathbf{Y}^i_{sup}\big|^N_1 \longleftarrow \overrightarrow{Y}^i_{sup}$
(3) $R^i_M \longleftarrow D^T\mathbf{Y}^i_{sup}\big|^N_1;$
(4) $D_B \longleftarrow B, B \longleftarrow Index(\max(R^i_M)\big|^C_1)$
(5) $A^i_{sup}\big|^N_1 \longleftarrow (\mathbf{D}^T_B\mathbf{D}_B)^{-1}\mathbf{D}^T_B\mathbf{Y}^i_{sup}\big|^N_1$
(6) $\mathbf{R}_M \longleftarrow \mathbf{D}^T(\mathbf{Y}^i_{sup}\big|^N_1 - \mathbf{D}_BA^i_{sup}\big|^N_1)$
(7) iter=iter+1
  end while
Output:
  $C_i \longleftarrow \arg\min \|Y^i_{sup}\big|^N_1 - D_cA^i_{sup}\big|^N_1\|_2$

ALGORITHM 1: SRC-FPSS algorithm.

The procedure is iterated until the number of iterations$\leq$K is satisfied. At last, output A^i_{sup}, and compute the class of pixels x_i as

$$C_i = \arg\min \left\| Y^i_{sup}\big|^N_1 - D_cA^i_{sup}\big|^N_1 \right\|_2 \tag{8}$$

The label of the input pixels are determined by the minimal representation error between $Y^i_{sup}\big|^N_1$ and its approximation recovered with $A^i_{sup}\big|^N_1$ and subdictionary $\mathbf{D}c$. The description of SRC-FPSS is summarized in Algorithm 1.

3. Experiment and Analysis

3.1. Datasets and Quantitative Metrics. In this section, we evaluate the proposed approach on two real hyperspectral datasets to verify its effeteness: AVIRIS Indian Pines image and Salinas image.

The Indian Pine image was collected from Indiana, USA, in June 1992, and it was provided by Purdue remote sensing image processing laboratory. The image is of size 145 × 145 × 220 with a spectral coverage ranging from 0.2 to 2.4 m and 20-m spatial resolution. For removals of noise and water absorption bands from the original data, 200 bands are retained as experimental data. The Indian Pine dataset contains 16 typical classes with 10249 samples. For each class, the numbers of training and test pixels are given in Table 1.

Table 1: Training and test set for 16 classes in Indian Pines dataset.

Class No.	name	Sample training	test
1	Alfalfa	6	48
2	Corn-no-till	144	1290
3	Corn-min	84	750
4	Corn	24	210
5	Grass/Pasture	50	447
6	Grass/Trees	75	672
7	Grass/Pasture-mowed	3	23
8	Hay-windrowed	49	440
9	Oats	2	18
10	Soybeans-no-till	97	871
11	Soybeans-min	247	2221
12	Soybeans-clean	62	552
13	Wheat	22	190
14	Woods	130	1164
15	Building-Grass-Trees-Drives	38	342
16	Stone-steel Towers	10	85
	Total	1043	9323

Table 2: Training and test set for 16 classes in Salinas dataset.

Class No.	name	Sample training	test
1	Brocoli_weeds_1	20	1989
2	Brocoli_weeds_2	37	3689
3	Fallow	20	1956
4	Fallow_rough_plow	14	1380
5	Fallow_smooth	27	2651
6	Stubble	40	3919
7	Celery	36	3543
8	Grapes_untrained	113	11158
9	Soil_vinyard_develop	62	6141
10	Corn_senesced_weeds	33	3245
11	Lettuce_romaine_4wk	11	1057
12	Lettuce_romaine_5wk	19	1908
13	Lettuce_romaine_6wk	9	907
14	Lettuce_romaine_7wk	11	1059
15	Vinyard_untrained	73	7195
16	Vinyard_trellis	18	1789
	Total	543	53586

The Salinas image was also acquired by the AVIRIS sensor over Salinas River Basin in California. The image size is of $512 \times 217 \times 224$ and the spatial resolution is 3.7 meters per pixel. Similar to the India pine image, the water and noise bands were removed and 204 bands are retained for test. According to the image there are 16 different categories with 54129 samples. For each class, the numbers of training and test pixels are given in Table 2.

Three commonly preferred performance indexes overall accuracy (OA), average accuracy (AA), and the Kappa coefficient (Kappa) are adopted to evaluate the quality of classification results in the experiments.

3.2. Experiments and Performance Analysis. In this section, we choose five state of art classification approaches for comparison: PSRC [17], JSRM [17], MJSR [20], MASR [20], and BTC [36]. For each class of two test data, about 10% and 1% of the labeled samples were chosen from Indian Pine and Salinas for training and the rest are used for testing. We take ten runs of classification estimation to estimate the average accuracy to avoid any bias.

The scale parameter setting is as follows, JSRM using a single scale for two datasets India pine (7*7), Salinas (11*11), MJSR, and MASA using multiscale for India pine (3-13) and Salinas (3-15). The parameters for the PSRC and BTC were set to the default values reported in [17, 36]. In the proposed method, the superpixels number is set by experience for getting a couple of compactness superpixels. Where the superpixels number of India pine image is 240 and the superpixels number of Salinas image is 540. The flexible and adjustable sampling frequency in each iteration is set to be $N=10$ for India pine image and Salinas image. More atoms cannot bring the improvement of accuracy but time

consuming. In this paper, the sparsity degree of $K = 2$ for testing on other sparsity degrees performs poorly.

The classification evaluation of six approaches of India pines is shown in Table 3 and we display the classification map and corresponding overall accuracy results in Figure 4. From the visual results, we can observe that the pixel wise sparse representation classifier provides much noise estimation of the classification. The JSRM, MJSR, and MASR incorporating the contextual information from adjacent area displayed smoother appearance and performed better on the quantitative comparison. These methods bring a big improvements on classification accuracy compared to PSRC. BTC-WLS is a lightweight sparsity-based classification technique and also provides a smooth appearance than PSRC, JSRM, and MJSR. As can be seen, the proposed method supervised by superpixels outperforms the comparison methods on visual effect. For the quantitative comparison of OA, AA, and the Kappa coefficient in Table 3, the SRC-FPSS also perform better than the other compared methods except AA compared with BTC-WLS.

The size of Salinas image is a relatively bigger than India pines and the average accuracy (OA, AA, and Kappa) results are in Table 4. The classification map and corresponding overall accuracy results are displayed in Figure 5.

As can be seen from Table 4, the classification accuracy of PSRC is on a low level. With the contextual spatial information, the accuracy of joint sparse representation model has been greatly improved and the average accuracy improves nearly 25%. On this basis, the multiscale joint sparse representation and the multiscale adaptive sparse representation can improve the classification accuracy rate by roughly 5%. With homogeneous superpixel constraint, the classification accuracy of SRC-FPSS has been generally improved: (1) the overall accuracy is improved to (99.48%),

TABLE 3: The classification accuracy (average on ten runs with 10% random samples) for India Pines image on six test approaches.

class	PSRC	JSRM	MJSR	MASR	BTC-WSL	SRC-FPSS
C1	39.84	84.55	83.74	94.31	99.07	92.68
C2	54.29	96.42	93.64	98.19	93.02	97.85
C3	52.21	92.41	93.44	97.50	89.91	97.90
C4	41.31	92.64	88.41	96.09	90.92	95.77
C5	85.36	94.33	94.41	96.24	94.32	98.01
C6	89.75	93.71	97.92	99.95	100.00	99.89
C7	74.67	85.33	84.00	100.00	98.15	100.00
C8	96.28	99.85	99.77	99.85	100.00	99.92
C9	24.07	31.48	59.26	66.67	100.00	70.37
C10	69.37	91.70	93.22	95.47	92.26	98.10
C11	70.88	96.73	97.13	98.98	99.17	98.87
C12	39.76	92.57	88.23	94.38	99.19	95.57
C13	91.85	80.98	93.84	98.55	100.00	98.37
C14	90.48	98.54	99.56	100.00	98.92	99.97
C15	41.11	91.64	89.15	97.60	95.48	95.39
C16	88.10	88.49	86.90	96.43	99.20	96.43
OA	69.29	94.77	94.16	98.03	96.43	98.29
AA	65.58	88.21	91.05	95.64	96.85	95.94
Kappa	64.94	94.03	94.31	97.75	95.91	98.05

TABLE 4: The classification accuracy (average on ten runs with 1% random samples) for Salinas images on six test approaches.

class	PSRC	JSRM	MJSR	MASR	BTC-WSL	SRC-FPSS
C1	98.11	100.00	100.00	100.00	100.00	100.00
C2	98.47	99.83	99.89	99.97	100.00	100.00
C3	96.32	98.72	99.95	99.89	100.00	98.77
C4	98.34	98.64	99.32	93.66	99.92	95.92
C5	96.97	99.25	99.10	94.06	99.92	99.29
C6	99.73	100.00	100.00	97.98	100.00	99.92
C7	99.32	99.94	99.88	97.97	100.00	100.00
C8	75.82	88.07	95.07	99.43	100.00	99.69
C9	98.73	99.97	99.78	100.00	100.00	100.00
C10	92.23	97.98	98.46	98.14	100.00	99.71
C11	98.42	99.51	100.00	99.11	100.00	99.70
C12	99.40	99.95	99.95	95.90	100.00	99.89
C13	96.90	99.20	99.89	92.99	99.15	99.31
C14	95.87	98.82	99.51	95.96	99.48	99.51
C15	64.53	79.98	91.93	99.15	88.75	98.70
C16	98.66	99.36	99.71	99.36	100.00	98.78
OA	88.79	94.50	97.67	98.46	98.53	99.48
AA	94.24	9745	98.90	99.20	97.72	99.32
Kappa	85.42	93.87	97.40	98.28	98.36	99.42

(2) the average accuracy is improved to 99.32%, and the Kappa coefficients are improved to 99.42% in the proposed method.

Figures 6 and 7 show the corresponding relationship between sampling frequency N and the classification accuracy indexes (OA, AA, and Kappa coefficient) on India pines and Salinas image. These indexes are obtained by averaging the results conducting five independent runs. As can be observed in the two figures, the three indexes of the proposed classifier generally improve with the increase of sampling frequency.

4. Conclusion

Aiming at the problems of the imperfect utilization of the context information, this paper puts forward a sparse representation classification algorithm based on overcomplete
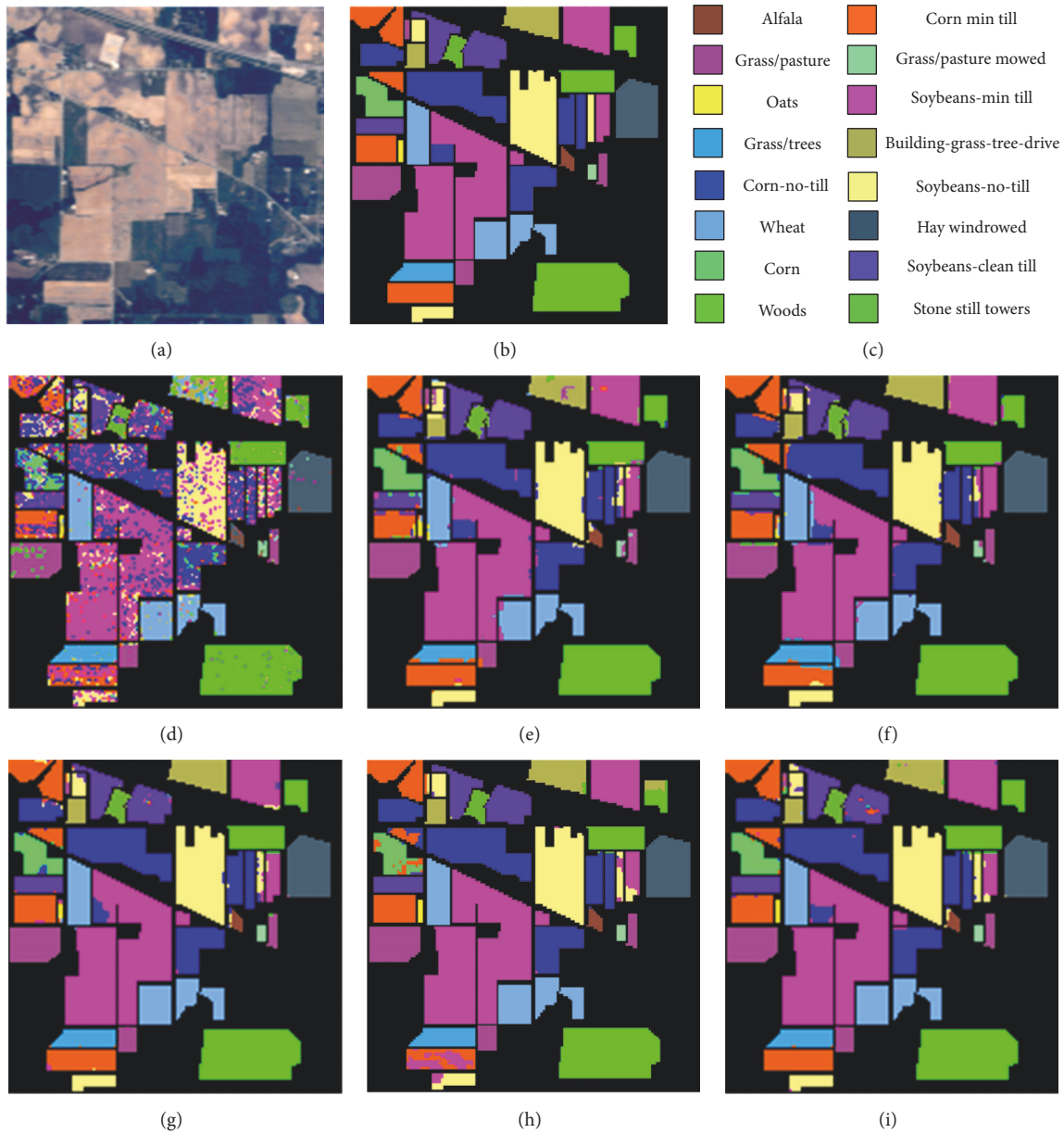
FIGURE 4: Classification maps by different methods for the AVIRIS Indian Pines image with overall accuracies (OA in %). (a) Three-band color image by PCA. (b) Ground-truth image. (c) Color legend of corresponding classes. (d) PSRC (68.54%). (e) JSRM (94.15%). (f) MJSR (94.00%). (g) MASR (98.25%). (h) BTC-WSL (95.58%). (i) SRC-FPSS (98.35%).

sampling on superpixels for hyperspectral image. The main contribution of this paper is that collaborative sampling on spatial and spectral in superpixels exploits implicit context information for test pixels, and the joint sparse optimal of sampling patches fuses spectral and spatial structure information effectively of HSI data point. Thus, the joint sparse norm improves the category feature extraction and representation of pixels, providing significant enhancement of classification performance in the iterative process. The proposed SRC-FPSS was tested on two hyperspectral images and obtained better classification performance. In addition, we will introduce the discriminative learning algorithms in the proposed model in our further work.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

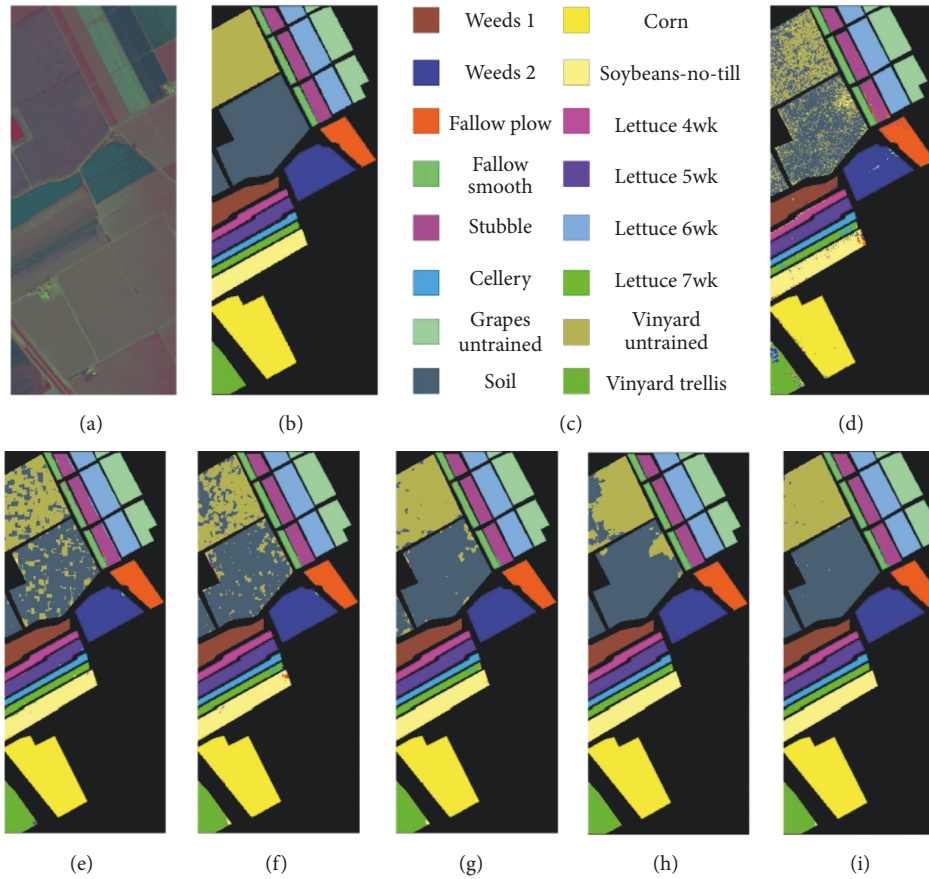The authors declare that they have no conflicts of interest.

Figure 5: Classification maps by different methods for the Salinas image with overall accuracies (OA in %). (a)Three-band color image by PCA. (b) Ground-truth image. (c) Color legend of corresponding classes. (d) PSRC (86.48%). (e) JSRM (92.07%). (f) MJSR (93.64%). (g) MASR (98.06%). (h) BTC-WSL (95.99%). (i) SRC-FPSS (99.35%).
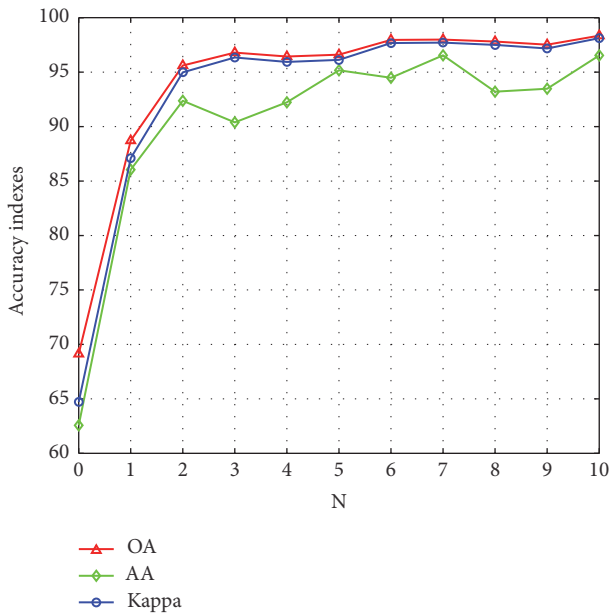

Figure 6: Effect of the sampling frequency N and the classification accuracy indexes on Indian Pines.
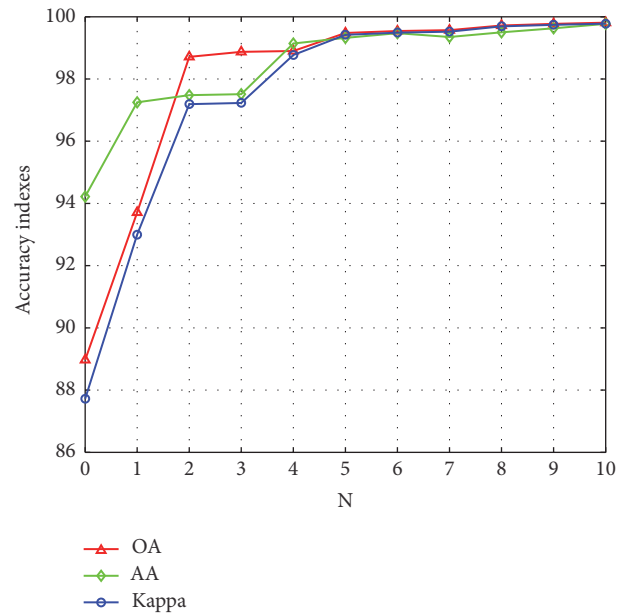

Figure 7: Effect of the sampling frequency N and the classification accuracy indexes on Salinas.

Acknowledgments

References

[1] G. Cavallaro, M. Riedel, J. A. Benediktsson et al., "Smart data analytics methods for remote sensing applications," in *Proceedings of the IGARSS 2014 - 2014 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1405–1408, July 2014.

[2] B. Yousefi, S. Sojasi, C. I. Castanedo et al., "Mineral identification in hyperspectral imaging using Sparse-PCA," in *Proceedings of the SPIE Commercial + Scientific Sensing and Imaging*, p. 986118, 2016.

[3] J. Y. Wu, S. Sfarra, and Y. Yao, "Sparse Principal Component Thermography for Subsurface Defect Detection in Composite Products," *IEEE Transactions on Industrial Informatics*, pp. 1-1, 2018.

[4] B. Yousefi, C. Ibarra Castanedo, G. Beaudoin, X. P V Maldague, F. Huot, and M. Chamberland, "Mineral Identification in LWIR Hyperspectral Imagery Applying Sparse Spectral Clustering," in *Proceedings of the Quantitative InfraRed Thermography Asia*, 2017.

[5] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, 2013.

[6] N. Alajlan, Y. Bazi, F. Melgani, and R. R. Yager, "Fusion of supervised and unsupervised learning for improved classification of hyperspectral images," *Information Sciences*, vol. 217, pp. 39–55, 2012.

[7] Z. Xue, P. Du, and H. Su, "Harmonic analysis for hyperspectral image classification integrated with PSO optimized SVM," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2131–2146, 2014.

[8] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new bayesian approach with active learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3947–3960, 2011.

[9] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image classification using soft sparse multinomial logistic regression," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 2, pp. 318–322, 2013.

[10] B. Du, L. Zhang, L. Zhang, T. Chen, and K. Wu, "A discriminative manifold learning based dimension reduction method for hyperspectral classification," *International Journal of Fuzzy Systems*, vol. 14, no. 2, pp. 272–277, 2012.

[11] S. Prasad and L. M. Bruce, "Limitations of principal components analysis for hyperspectral target recognition," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 4, pp. 625–629, 2008.

[12] R. Archibald and G. Fann, "Feature selection and classification of hyperspectral images with support vector machines," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 4, pp. 674–677, 2007.

[13] Z. Xue, J. Li, L. Cheng, and P. Du, "Spectral-spatial classification of hyperspectral data via morphological component analysis-based image separation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 70–84, 2015.

[14] Z. He and J. Li, "Multiple data-dependent kernel for classification of hyperspectral images," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1118–1135, 2015.

[15] H. Su and P. Du, "Multiple classifier ensembles with band clustering for hyperspectral image classification," *European Journal of Remote Sensing*, vol. 47, no. 1, pp. 217–227, 2014.

[16] P. Du, Z. Xue, J. Li, and A. Plaza, "Learning Discriminative Sparse Representations for Hyperspectral Image Classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 1089–1104, 2015.

[17] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3973–3985, 2011.

[18] A. Castrodad, Z. Xing, J. B. Greer, E. Bosch, L. Carin, and G. Sapiro, "Learning discriminative sparse representations for modeling, source separation, and mapping of hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4263–4281, 2011.

[19] U. Srinivas, Y. Chen, V. Monga, N. Nasrabadi, and T. Tran, "Exploiting Sparsity in Hyperspectral Image Classification via Graphical Models," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 3, pp. 505–509, 2012.

[20] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral-spatial hyperspectral image classification via multiscale adaptive sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 12, pp. 7738–7749, 2014.

[21] X. Zhang, Q. Song, Z. Gao, Y. Zheng, P. Weng, and L. C. Jiao, "Spectral-Spatial Feature Learning Using Cluster-Based Group Sparse Coding for Hyperspectral Image Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4142–4159, 2016.

[22] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 1, pp. 217–231, 2013.

[23] J. Feng, Z. Cao, and Y. Pi, "Polarimetric contextual classification of PolSAR images using sparse representation and superpixels," *Remote Sensing*, vol. 6, no. 8, pp. 7158–7181, 2014.

[24] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2281, 2012.

[25] L. Fang, S. Li, W. Duan, J. Ren, and J. A. Benediktsson, "Classification of Hyperspectral Images by Exploiting Spectral–Spatial Information of Superpixel via Multiple Kernels," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6663–6674, 2015.

[26] B. Yousefi, S. Sfarra, C. Ibarra Castanedo, and X. P. V. Maldague, "Comparative analysis on thermal non-destructive testing imagery applying Candid Covariance-Free Incremental Principal Component Thermography (CCIPCT)," *Infrared Physics & Technology*, vol. 85, pp. 163–169, 2017.

[27] K. Ahi, "A Method and System for Enhancing the Resolution of Terahertz Imaging," *Measurement*, 2018.

[28] L. Pezzati, P. Targowski, B. Yousefi, S. Sfarra, and X. P. Maldague, "Quantitative assessment in thermal image segmentation for artistic objects," in *Proceedings of the SPIE Optical Metrology*, p. 1033108, Munich, Germany, 2017.

[29] K. Ahi, S. Shahbazmohamadi, and N. Asadizanjani, "Quality control and authentication of packaged integrated circuits using enhanced-spatial-resolution terahertz time-domain spectroscopy and imaging," *Optics and Lasers in Engineering*, vol. 104, pp. 274–284, 2018.

[30] B. Yousefi, J. Fleuret, H. Zhang, X. P. V. Maldague, R. Watt, and M. Klein, "Automated assessment and tracking of human body thermal variations using unsupervised clustering," *Applied Optics*, vol. 55, no. 34, pp. D162–D172, 2016.

[31] H. Zhang, S. Sfarra, A. Osman et al., "An Infrared-Induced Terahertz Imaging Modality for Foreign Insert Detection in A Glass Fiber-Skinned Lightweight Honeycomb Composite Panel," *IEEE Transactions on Industrial Informatics*, pp. 1-1.

[32] K. Ahi, "Mathematical modeling of thz point spread function and simulation of THz imaging systems," *IEEE Transactions on Terahertz Science and Technology*, vol. 7, no. 6, pp. 747–754, 2017.

[33] H. Zhang, S. Sfarra, A. Osman et al., "Eddy current pulsed thermography for ballistic impact evaluation in basalt-carbon hybrid composite panels," *Applied Optics*, vol. 57, no. 18, p. D74, 2018.

[34] M. Rousson, T. Brox, and R. Deriche, "Active unsupervised texture segmentation on a diffusion based feature space," in *Proceedings of the CVPR 2003: Computer Vision and Pattern Recognition Conference*, pp. 699–704, Madison, WI, USA.

[35] A. Schick, M. Fischer, and R. Stiefelhagen, "An evaluation of the compactness of superpixels," *Pattern Recognition Letters*, vol. 43, no. 1, pp. 71–80, 2014.

[36] M. A. Toksöz and I. Ulusoy, "Hyperspectral image classification via basic thresholding classifier," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 7, pp. 4039–4051, 2016.