HYDERSPECTRAL images (HSIs) are very important in remote sensing. An HSI generally consists of more than a hundred spectral bands which represent the reflected optical radiation from ground objects. Therefore, HSIs not only contain spatial information as other optical images but also provide abundant spectral information with its numerous and continuous spectral bands. Due to the above properties, HSIs have been used in many domains such as agriculture [1]–[3], military [4], [5], and environmental monitoring [6], [7].

Supervised HSI classification [8]–[14], which categorizes test pixels after training classifiers with some labeled samples, is a hot research topic for HSIs. Many spectral pixel-wise classification methods have been developed, such as the support vector machine (SVM) [15], the neural networks [16], and the adaptive artificial immune network [17]. In general, these methods mainly make use of the spectral information in the HSI for classification. Therefore, the obtained classification maps usually look noisy. Recently, many classification methods have been proposed to incorporate the spatial information of an HSI, based on the assumption that pixels within a local region represent the same ground objects and share the similar spectral characteristics [18]–[26]. The classification method proposed in [18] uses a composite kernel machines to combine spatial and spectral information for SVM classification. In [19], a Bayesian-based classification method is proposed to exploit the statistical dependence among neighboring pixels. In [20], spatial information is extracted based on mathematical morphology, and then, spatial and spectral information are incorporated into an SVM classifier. In [21], the extended SVM algorithm combines the spectral information of the HSI with the spatial information by using a composite kernel approach. In [22], spatial information is combined with spectral information by using a hypergraph approach. In [23], spatial and spectral information are exploited by a spatial–spectral similarity measure which maps the distance between two image patches in HSIs. In [24], the spectral information is characterized by fusion local and global probabilities, and spatial information is utilized by including a Markov random field regularizer. In [25], HSI segmentation based on hidden Markov random fields is combined with an SVM classifier to integrate spatial and spectral information. In [26], the classification method utilizes spatial and spectral information by integrating a spectral-domain local preserving scatter matrix and a spatial-domain local pixel neighborhood preserving scatter matrix.

The sparse representation (SR) approach has been extremely powerful in digital image processing applications, such as in fusion [27], interpolation [28], and face recognition [29]. During the last few years, SR-based classifiers have been applied to HSI classification [30]–[35]. The SR for HSI classification is based on the observation that HSI pixels from the same class can usually be represented by a linear combination of a few common pixels from the same class. Thus, a test pixel can be sparsely represented by a few atoms from a structured dictionary which consists of training pixels from all classes. The test pixel can be classified by the recovered sparse coefficients, which contain the positions of the selected atoms and their weight values. A joint sparse representation classification (JSRC) method is proposed in [36] to exploit the spatial context and shapes, and conform to the spatial structure of the HSI as far as possible. Finally, the label of the test pixel is determined by applying the joint sparse representation classifier to the first several PCs of pixels within the corresponding SA region. According to the experiments performed on several HSIs, the proposed SAJSRC method outperforms some widely used HSIs classification approaches.

I. INTRODUCTION

The sparse representation (SR) approach has been extremely powerful in digital image processing applications, such as in fusion [27], interpolation [28], and face recognition [29]. During the last few years, SR-based classifiers have been applied to HSI classification [30]–[35]. The SR for HSI classification is based on the observation that HSI pixels from the same class can usually be represented by a linear combination of a few common pixels from the same class. Thus, a test pixel can be sparsely represented by a few atoms from a structured dictionary which consists of training pixels from all classes. The test pixel can be classified by the recovered sparse coefficients, which contain the positions of the selected atoms and their weight values. A joint sparse representation classification (JSRC) method is proposed in [36] to exploit the spatial context and shapes, and conform to the spatial structure of the HSI as far as possible. Finally, the label of the test pixel is determined by applying the joint sparse representation classifier to the first several PCs of pixels within the corresponding SA region. According to the experiments performed on several HSIs, the proposed SAJSRC method outperforms some widely used HSIs classification approaches.

Index Terms—Classification, hyperspectral image (HSI), shape-adaptive algorithm, sparse representation.
Fig. 1. Proposed SAJSRC method for HSIs.

Fig. 2. (a) Sketch map of shape-adaptive region. Red block represents the central pixel, black blocks denote lengths of each direction, and gray area means the shape-adaptive region. (b) Examples of the obtained SA regions.

Fig. 3. (a) PC1 map of Indian Pines image. (b) Color map of the ground truth corresponding to this HSI. (c) Edges of (b). (d) Edges of the largest coverage regions of SA regions. (e) Coincident edges between (c) and (d). (f) Edges of the largest coverage regions of square windows.

Both the JSRC and NLW-SR methods are fixed-size window-based joint SR methods. However, the fixed-size window cannot sufficiently explore the spatial information in HSIs since the homogeneous area may need a large window while the heterogeneous area need a small window. In order to adaptively and sufficiently explore the spatial information for different types of spatial structures (e.g., homogeneous structure and heterogeneous structure), a new shape-adaptive joint sparse representation classification (SAJSRC) method is proposed in this paper. The shape-adaptive regions, which are constructed based on the local variations in the HSI, can sufficiently explore the spatial information as compared with the fixed-size window, and thus improve the JSRC. Experimental results show that the proposed method outperforms other HSI classification methods, especially the fixed-size window-based joint SR methods.

This paper is organized as follows. The window-based JSRC method for HSIs are introduced in Section II. The SAJSRC method for HSIs is proposed in Section III. The experimental results and discussions are presented in Section IV. Finally, Section V summarizes the paper and future works are suggested.

II. WINDOW-BASED JSRC METHOD FOR HSIS CLASSIFICATION

For HSIs, pixels from the same class commonly represent the same ground materials and share the same spectral characteristics. So an HSI pixel $s^j$, which belongs to class $j$, can be compactly represented by

$$s^j = D^j \alpha^j$$  \hspace{1cm} (1)
where columns of matrix $\mathbf{D}^j$ are HSI pixels from class $j$ and vector $\alpha^j$ records the corresponding weights. In the ISRC method, the full $\mathbf{D} = [\mathbf{D}^1 \mathbf{D}^2 \ldots \mathbf{D}^J]$ consists of all the subdictionaries of total $J$ classes. The ISRC method assumes that the total $n$ pixels in the $\sqrt{n} \times \sqrt{n}$ window are from the same class $j$. Therefore, the joint signal matrix $\mathbf{S} = [s_1^j s_2^j \ldots s_n^j]$ which consists of those $n$ pixels in the window can be represented as

$$\mathbf{S} = [s_1^j s_2^j \ldots s_n^j] = [\mathbf{D}^1 \alpha^j \mathbf{D}^2 \alpha^j \ldots \mathbf{D}^J \alpha_n^j] = [\mathbf{D}^1 \ldots \mathbf{D}^j \ldots \mathbf{D}^J] \cdot [0 \ldots \Psi^j \ldots 0]^T = \mathbf{D} \Psi. \quad (2)$$

Since the number of columns in subdictionary $\mathbf{D}^j$ is much less than the number of columns in dictionary $\mathbf{D}$, the number of rows in $\Psi^j$ is much less than the number of rows in $\Psi$. So, $\Psi$ is a sparse vector matrix with only a few nonzero rows. $\Psi$ can be obtained by solving the following optimization problem:

$$\hat{\Psi} = \arg \min_{\Psi} \|\Psi\|_{\text{row},0} \text{ s.t. } \mathbf{S} = \mathbf{D} \cdot \Psi \quad (3)$$

where $\|\Psi\|_{\text{row},0}$ means the number of nonzero rows of $\Psi$. Problem in (7) can be relaxed to an inequality problem

$$\tilde{\Psi} = \arg \min_{\tilde{\Psi}} \|\mathbf{D} \tilde{\Psi} - \mathbf{S}\|_2 \text{ s.t. } \|\Psi\|_{\text{row},0} \leq \eta \quad (4)$$

where $\eta$ is the given sparsity level [38]. The simultaneous orthogonal matching pursuit (SOMP) method [39], [40] can be used to solve the problem in (4). Once $\Psi$ is obtained, the reconstruction residual errors-based classification of the central pixel $s$ can be described as

$$\text{Class}(s) = \arg \min_{j=1,2,\ldots,J} r_j^2(s) \quad (5)$$

where $r_j^2(s) = \|\mathbf{S} - \mathbf{D}^j \cdot \Psi^j\|_2$, $j = 1, 2, \ldots, J$ is the corresponding reconstruction residual errors of the $j$th class.

### III. SAJSRC Method for HSIs Classification

The SAJSRC method consists of two parts: 1) the shape-adaptive algorithm-based spatial information exploration and 2) the incorporation of the spatial information into the joint SR classifier. The proposed classification method is demonstrated in Fig. 1, and the detailed description is given below.

#### A. Spatial Information Exploration

For an $M \times N \times B$ HSI image, the PC1 (the size is $M \times N$) of the HSI can represent the spatial variance of the image [41], so the HSI is first transformed to a set of principal components (PCs) by the PCA algorithm. Then, the shape-adaptive algorithm [42] is used in the PC1 to construct a shape-adaptive (SA) smooth region for each pixel in PC1.

As shown in Fig. 2(a), an SA region is a polygon and its shape and size are determined by the positions of its eight polygon vertices. The eight directions $\{\theta_k = k\pi/4\}_{k=1,2,\ldots,8}$ are known and the eight lengths $\{h_{\theta_k}\}_{k=1,2,\ldots,8}$ from the central pixel to vertices need to be computed. Represent the PC1 by $z(x)$, which is corrupted by noise $\varepsilon(x)$, where $x$ denotes the spatial variable. The model is

$$z(x) = y(x) + \varepsilon(x) \quad (6)$$

where $y(x)$ represents the ideal noiseless image. In the shape-adaptive algorithm, the local polynomial approximation (LPA) filtering [43] technique is first used to directionally estimate $y(x)$ with the different-scale neighboring information in $z(x)$ by filtering. Then, the intersection of confidence intervals (ICI) rule [42] is used to adaptively determine the optimal length of $h_{\theta_k}$ by finding the similar and continuous-scale LPA estimates of neighboring information.

1) LPA Filtering: Define a candidate length set $H = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ ($\lambda_1 < \lambda_2 < \ldots < \lambda_m$) of $h$ ($h \in H$). First, construct a set of LPA convolution kernels $g_{n,h_{\theta_k}}$ by $g(x, h_{\theta_k}) = \Phi^{-1} w_h \phi(\nu/v) [1 \ldots 0]^T$. $\nu = \Phi(\theta_k) \Delta x$ denotes the rotated difference of coordinates between pixel $x$ and neighborhood pixels, in which $\Phi(\theta_k)$ denotes the rotation operator, $w_h(\nu) = w(\nu/h)$ is a basic window function satisfying: $w(\nu) \geq 0$, $w(0) = \max\{w(\nu), \int w(\nu) d\nu = 1\$. $\Phi$ is a vector of two-dimensional (2-D) polynomials $\phi_n = \nu^m / n!$, $n = 0, \ldots, \omega$, and $\Phi = \sum_n w_h(\nu) \phi(\nu) \phi(\nu)^T$. Then, the LPA filtering results can be obtained by $g_{n,h_{\theta_k}} = z \otimes g_{n,h_{\theta_k}}$, where “$\otimes$” denotes the convolution operation.

2) ICI Rule: For point $x$, the confidence intervals $E(x)_{h_{\theta_k}}$ of the LPA filtering results $\hat{y}(x)_{h_{\theta_k}}$ of the LPA filtering results $\hat{y}(x)_{h_{\theta_k}}$ are

$$E(x)_{h_{\theta_k}} = \left[\hat{y}(x)_{h_{\theta_k}} - \Gamma \cdot \text{std} \left(\hat{y}(x)_{h_{\theta_k}}\right), \hat{y}(x)_{h_{\theta_k}} + \Gamma \cdot \text{std} \left(\hat{y}(x)_{h_{\theta_k}}\right)\right] \quad (7)$$

where $\Gamma > 0$ is a threshold parameter related to the noise standard of images and $\text{std}(\hat{y}(x)_{h_{\theta_k}})$ is the standard deviation of $\hat{y}(x)_{h_{\theta_k}}$. Based on the confidence intervals, the length in direction $\theta_k$ of point $x$ can be determined by

$$h(x)_{\theta_k} = \lambda_i \text{ s.t. } \left\{ E(x)_{h=\lambda_i} \cap E(x)_{h<\lambda_i} \neq \emptyset \right\} = E(x)_{h=\lambda_i} \cap E(x)_{h>\lambda_i} = \emptyset \quad (8)$$

where $\lambda_i \in H$.

The real examples of the obtained SA regions are shown in Fig. 2(b). In addition, the superiority of the SA regions over the fixed-size square window in terms of avoiding dissimilar pixels from other classes are shown in Fig. 3. In ideal cases, the SA regions should not cross the edges of the HSI since pixels on both sides of the edges are generally dissimilar. The exact edges of HSI is hard to obtain, so we use the edges of the ground truth for test. Let $\Omega_i$ (i.e., the spatial coordinate set) be the SA region of pixel $s_i$. We define a largest coverage region for $n$ pixels as

$$\Omega_{\text{max}} = \Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_n. \quad (9)$$

For pixels in each separate color region in ground truth map, the corresponding largest coverage region is computed by (9). The edges of all largest coverage regions are shown in Fig. 3(d). As can be seen in Fig. 3(e), the edges of largest coverage regions are consistent with the edges of the ground truth map in most
cases, which means that the SA barely cross the edges of the ground truth map and can effectively avoid pixels from other classes. The similar definition of the largest coverage region can be used in the fixed-size square windows, and the edges of all largest coverage regions are shown in Fig. 3(f). Obviously, some edges in Fig. 3(f) cross each other, which means that the square windows may search pixels from the different classes.

B. Incorporation of Spatial Information Into Joint SR Classifier

High-dimensional HSI pixels result in a heavy computation burden for the joint SR classifier. For an efficient processing of the neighboring pixels in the shape-adaptive region, the known PCA is utilized to reduce the dimension of HSIs [44]. Consequently, the incorporation of spatial information into the joint SR classifier can be described as follows.

First, stack the first $L$ PCs $[PC_1\ PC_2\ldots\ PC_L]$ into a dimensionality-reduced image $I$ with the size of $M \times N \times L$. Next, construct a dictionary $D_1 = [D_1^1\ D_1^2\ldots\ D_1^J]$, where subdictionary $D_1^j (j = 1, 2, \ldots, J)$ consists of training pixels randomly chosen from class $j$ in image $I$. For each test pixel $s$ of image $I$, the corresponding SA region searches neighboring pixels and stacked obtained pixels into a joint signal matrix $S_{SA}$. The joint signal matrix can be represented by a few common atoms of the dictionary $D_1$:

$$
\Psi = \arg \min \| D_1 \Psi - S_{SA} \|_2 \text{ s.t. } \| \Psi \|_{row,0} \leq \eta.
$$

(10)
The sparse coefficients matrix $\Psi$ can be obtained by the SOMP [40] algorithm. The final classification of the test pixel $s$ can be represented as

$$\text{Class}(s) = \arg \min_{j=1,2,\ldots,J} \| S_{SA} - D_1^j \cdot \Psi_j \|_2.$$ (11)

It is noticed that in order to effectively solve (9) by the SOMP algorithm, the columns of $S_{SA}$ and $D_1$ should be normalized to have unit norm. The implementation details of the proposed SAJSRC algorithm is summarized in Algorithm 1.

Algorithm 1. SAJSRC

Input: 1). The hyperspectral image; 2). The training samples; 3). The number of significant principal components $L$;

Output: Classification results of all test pixels; 1): Decompose the HSI with the PCA algorithm; 2): Construct a SA region for each test pixel by applying the SA algorithm to the PC1; 3): Stack the first $L$ PCs of HSI to a new image $I$; 4): Construct the structured dictionary $D_1$ and normalize all columns of $D_1$ to have an unit norm; 5): for each test pixel $s$ in the new image $I$ do 6): Construct a joint signal matrix $S_{SA}$ by grouping pixels within the SA region, and normalize the columns of $S_{SA}$ to have unit norm; 7): Calculate the joint sparse vector matrix $\Psi$ from (9); 8): Determine the class of test pixel $s$ by (10); 9): Turn to the next test pixel; 10): end for.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we show the effectiveness of the proposed SAJSRC method on four real HSIs, i.e., the AVIRIS Indian Pines, the AVIRIS Salinas, the ROSIS-03 Pavia University, and the Houston data. For the shape-adaptive algorithm in our method, the candidate lengths are $H = \{1,2,3,5,7,9\}$. The sparsity level is chosen from $\eta = 3$ to $\eta = 7$ to solve (9) and the number of PCA bands is set as $L = 10$. According to our experiments, this parameters setting is able to obtain high classification accuracies for four real HSIs.

The classification results of the SAJSRC method are visually and quantitatively compared with those obtained by some widely used classification methods, i.e., SVM [15], SVM-CK [18], OMP [45], LORSAL-MLL [19], JSRC [36], and NLW-SR [37]. The SVM-CK, LORSAL-MLL, JSRC, and NLW-SR methods take advantage of the spatial information for HSI classification, while the SVM and OMP methods are pixel-wise classification methods without considering the spatial information. The parameters of the SVM method are determined by the 10-fold cross-validation technique. For the SVM-CK method, the standard deviation and the weight of the textural content is $\mu = 0.6$ and the weight of the spectral content is $1 - \mu$. The parameters of the SVM classifier used in the SVM-CK method are determined by the five-fold cross-validation technique. For the JSRC method, the window sizes for the Indian Pines image, the Salinas image, the Pavia University image, and the Houston image are $7 \times 7$, $15 \times 15$, $11 \times 11$, and $5 \times 5$, respectively. For the NLW-SR method, the size of the nonlocal weighting patch is $7 \times 7$. The neighborhood window sizes for the Indian Pines image, the Salinas image, the Pavia University image, and the Houston image are $9 \times 9$, $15 \times 15$, $13 \times 13$, and $7 \times 7$, respectively. The thresholds of the nonlocal weights are $\omega_1 = 0.14, \omega_2 = 0.88$. The parameters of the LORSAL-MLL and OMP method are given in [19] and [36], respectively.

The computer used in the experiments has an Intel(R) Core(TM)2 CPU 2.00 GHz and 4 GB RAM, and the software platform is MATLAB.

A. AVIRIS Data Set: Indian Pines Image

This HSI was collected by the Airborne/Visible Infrared Imaging Specrometer (AVIRIS) sensor over the agricultural Indian Pine test site in northwestern Indiana. The image consists of 220 data channels across the spectral range from 0.2 to 2.4 $\mu$m and, each band contains 145 $\times$ 145 pixels with a spatial resolution of 20 m. For this HSI, we remove 20 water absorption bands, i.e., channels no. 104–108, 150–163, and 220 [46]. The false color composite of the Indian Pines image is shown in Fig. 4(a). The reference of this image contains 16 classes of materials. The reference map is shown in Fig. 4(b). For this image, we randomly choose 10% of the labeled pixels as training samples and the remainder as test samples [36], as shown in Table I.

The classification maps obtained by various classification methods are shown in Fig. 4(c)–(i). The quantitative results averaged over 10 runs, including three important metrics [overall (OA), average accuracy (AA), and the Kappa coefficient ($\kappa$)], as well as the running time, are presented in Table II. As can be observed, the spectral–spatial-based classification methods (SVM-CK, LORSAL-MLL, JSRC, NLW-SR, and SAJSRC) significantly outperform the pixel-wise methods, i.e., SVM and OMP. In particular, the proposed SAJSRC method has the best performance in terms of OA, AA, and $\kappa$. Moreover, the classification accuracies of the SAJSRC method are higher than the JSRC method for almost all classes. Especially for some classes, the improvement is significant. For example, in Table II, the classification accuracy of the 13th class increases from 74.54% to 99.07%. Similar improvements can be found in the classification accuracy of the 5th, 7th, 9th, and 16th classes. The reason is that the shape-adaptive local regions are more reasonable for detailed regions than the fix-sized square windows to sufficiently exploit spatial information.

Next, the influences of the DR process on the SAJSRC method are analyzed. The analysis focuses on two aspects: the effect of the number of dimensionality reduced bands $L$ and the robustness of using different DR approaches. In this experiment, the number of training samples and test samples are shown in Table I. The PCA-based DR approach is used for different classifiers and $L$ is ranged from 2 to 60, and the obtained...
overall accuracies are shown in Fig. 5(a). The corresponding results of using the minimum noise fraction (MNF) based [47] DR approach are shown in Fig. 5(b). Fig. 5(c) represents the results of using the independent component analysis (ICA)- based [48] DR approach. As observed in Fig. 5, the proposed SAJSRC consistently outperforms other classification methods with different DR approaches and various $L$. For the three DR approaches, the SAJSRC method can obtain high and stable OAs when $L \geq 10$. Therefore, the proposed SAJSRC is robust for different DR approaches and the optimal number of dimensionality-reduced bands is $L = 10$. We use the PCA-based DR process for the SAJSRC method just because that the PCA-based DR approach has the comparable performance with other DR approaches and has been used in the spatial information exploration once, and thus, no extra computation is needed.

In addition, the influences of different numbers of training samples on the SVM, SVM-CK, OMP, LORSAL-MLL, JSRC, NLW-SR, and SAJSRC methods are analyzed. In this experiment, different percentages of labeled samples, ranging from 1% to 30%, are randomly selected as training samples, and the remaining samples are used as test samples. Fig. 6 shows the overall accuracies (obtained by averaging over 10 runs) for different classification methods with different number of training samples. As shown in Fig. 6, the classification accuracies of all methods generally increase as the number of training samples increases, and the proposed SAJSRC method can consistently outperform other approaches when the number of training samples is larger than 3% reference data. As the number of training samples is quite limited (1%–3%), the SAJSRC method cannot obtain the highest classification accuracy. The reason is that when the number of training samples is quite limited, training samples will not be able to construct a complete dictionary $\mathbf{D}$, and thus, the representation of signals upon $\mathbf{D}$ is not accurate. In this situation, the classification accuracy of the proposed SAJSRC method, as well as other SR-based methods (JSRC and NLW-SR), may decrease significantly. The LORSAL-MLL is an active learning-based method, and the number of training samples will be gradually increased based on an active learning mechanism. So, the LORSAL-MLL outperforms other methods as the number of training samples is very small.

The candidate set $H = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ ($\lambda_1 < \lambda_2 \cdots < \lambda_m$) is a very important parameter to determine the shape and size of the shape-adaptive regions. So, we analyze this parameter for the proposed SAJSRC method by varying the values in $H$ and record the corresponding classification accuracies. In the experiment, $\lambda_1$ is fixed to be 1, the step size between $\lambda_{i-1}$ and $\lambda_i$ is set to be 1, and the maximum value $\lambda_m$ is ranged from 1 to 9. The classification accuracies with respect to different $\lambda_m$ are shown in Fig. 7. As shown in Fig. 7, the proposed SAJSRC method has a lower classification accuracies when $\lambda_m = 1, 2$. It is because that as $\lambda_m = 1, 2$, the SA regions are too small to sufficiently explore the spatial information in homogeneous areas. When $\lambda_m \geq 3$, the SA regions in homogeneous areas can be large enough, and SA regions in heterogeneous areas can be small enough, so the classification accuracies are high and stable.
TABLE III
NUMBER OF TRAINING AND TEST SAMPLES OF SIXTEEN CLASSES IN SALINAS IMAGE

<table>
<thead>
<tr>
<th>Class</th>
<th>Name</th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Weeds_1</td>
<td>20</td>
<td>1989</td>
</tr>
<tr>
<td>2</td>
<td>Weed_2</td>
<td>37</td>
<td>3689</td>
</tr>
<tr>
<td>3</td>
<td>Fallow</td>
<td>20</td>
<td>1956</td>
</tr>
<tr>
<td>4</td>
<td>Fallow plow</td>
<td>14</td>
<td>1380</td>
</tr>
<tr>
<td>5</td>
<td>Fallow smooth</td>
<td>27</td>
<td>2651</td>
</tr>
<tr>
<td>6</td>
<td>Stubble</td>
<td>40</td>
<td>3919</td>
</tr>
<tr>
<td>7</td>
<td>Celery</td>
<td>36</td>
<td>3543</td>
</tr>
<tr>
<td>8</td>
<td>Grapes</td>
<td>113</td>
<td>1158</td>
</tr>
<tr>
<td>9</td>
<td>Soil</td>
<td>62</td>
<td>6141</td>
</tr>
<tr>
<td>10</td>
<td>Corn</td>
<td>33</td>
<td>3245</td>
</tr>
<tr>
<td>11</td>
<td>Lettuce 4wk</td>
<td>11</td>
<td>1057</td>
</tr>
<tr>
<td>12</td>
<td>Lettuce 5wk</td>
<td>19</td>
<td>1908</td>
</tr>
<tr>
<td>13</td>
<td>Lettuce 6wk</td>
<td>9</td>
<td>907</td>
</tr>
<tr>
<td>14</td>
<td>Lettuce 7wk</td>
<td>11</td>
<td>1059</td>
</tr>
<tr>
<td>15</td>
<td>Vineyard untrained</td>
<td>73</td>
<td>7195</td>
</tr>
<tr>
<td>16</td>
<td>Vineyard trellis</td>
<td>18</td>
<td>1789</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>543</td>
<td>53586</td>
</tr>
</tbody>
</table>

B. AVIRIS Data Set: Salinas Image

The Salinas image was captured by the AVIRIS sensor over Salinas Valley, California. The size of this image is $512 \times 217$ with a spatial resolution of 3.7 m/pixel and each pixel has 224 spectral bands. Similar to the Indian Pines image, 20 water absorption bands (108–112, 154–167, and 224) are ignored. The false color composite and the reference map of the Salinas image are shown in Fig. 8(a) and (b), respectively. In the experiments, we randomly select only 1% of the labeled pixels as the training samples and remaining 99% of the labeled pixels as the
The classification maps of different classification methods are shown in Fig. 8(c)–(i). The OA, AA, Kappa, and computing time (obtained by averaging over 10 runs) of different classification methods are presented in Table IV. As shown in this table, the proposed SAJSRC method outperforms other approaches in terms of visual quality of classification maps and quantitative metrics, i.e., OA, AA, and $\kappa$. From Table IV, the gains (OA) of the SAJSRC method for other methods are more than 3% and the running time of the SAJSRC method is much less than other SR-based spectral–spatial methods (JSRC and NLW-SR). Although the LORSAL-MLL method is faster than the SAJSRC method when the number of training samples is very small, the OA of the LORSAL-MLL method has the approximate 5% decline than the OA of the SAJSRC method.

C. ROSIS Urban Data: Pavia University Image

The Pavia University image, which covers an urban area surrounding the University of Pavia, Italy, was captured by the Reflective Optics System Imaging Spectrometer (ROSIS-03) sensor. The size of this image is $610 \times 340 \times 115$ with a spatial resolution of 1.3 m/pixel and a spectral coverage ranging from 0.43 to 0.86 $\mu$m. In the experiments, the 12 very noisy
channels were removed and 9 labeled classes are used as the ground truth. The false color composite and the reference map are shown in Fig. 9(a) and (b), respectively. We randomly select 300 pixels for each class as the training pixels and then, the rest of the labeled pixels are used as the test pixels (see Table V).

For this image, the visual classification maps are shown in Fig. 9(c)–(i). The quantitative classification results and the running time (obtained by averaging over 10 runs) of different classification methods are shown in Table VI. From Table VI, the proposed SAJSRC method achieves the highest OA and spends less time in classification than other spectral–spatial approaches, i.e., LORSAL-MLL, JSRC, and NLW-SR.

D. AVIRIS Data Set: Houston Image

The Houston image acquired over the University of Houston campus and the neighboring urban area has 144 spectral bands.

![False color composite of the Houston image.](image)

![Ground truth of the Houston image.](image)

![Classification maps of various methods.](image)

**TABLE VII**

<table>
<thead>
<tr>
<th>Class</th>
<th>Name</th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Grass_healthy</td>
<td>125</td>
<td>1126</td>
</tr>
<tr>
<td>2</td>
<td>Grass_stressed</td>
<td>125</td>
<td>1129</td>
</tr>
<tr>
<td>3</td>
<td>Grass_synthetic</td>
<td>70</td>
<td>627</td>
</tr>
<tr>
<td>4</td>
<td>Tree</td>
<td>124</td>
<td>1120</td>
</tr>
<tr>
<td>5</td>
<td>Soil</td>
<td>124</td>
<td>1118</td>
</tr>
<tr>
<td>6</td>
<td>Water</td>
<td>33</td>
<td>292</td>
</tr>
<tr>
<td>7</td>
<td>Residential</td>
<td>127</td>
<td>1141</td>
</tr>
<tr>
<td>8</td>
<td>Commercial</td>
<td>124</td>
<td>1120</td>
</tr>
<tr>
<td>9</td>
<td>Road</td>
<td>125</td>
<td>1127</td>
</tr>
<tr>
<td>10</td>
<td>Highway</td>
<td>123</td>
<td>1104</td>
</tr>
<tr>
<td>11</td>
<td>Railway</td>
<td>124</td>
<td>1111</td>
</tr>
<tr>
<td>12</td>
<td>Parking_lot1</td>
<td>123</td>
<td>1110</td>
</tr>
<tr>
<td>13</td>
<td>Parking_lot2</td>
<td>47</td>
<td>422</td>
</tr>
<tr>
<td>14</td>
<td>Tennis_court</td>
<td>43</td>
<td>385</td>
</tr>
<tr>
<td>15</td>
<td>Running_track</td>
<td>66</td>
<td>594</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1503</td>
<td>13 526</td>
</tr>
</tbody>
</table>

**TABLE VIII**

<table>
<thead>
<tr>
<th>Class</th>
<th>SVM</th>
<th>SVM-CK</th>
<th>LORSAL-MLL</th>
<th>OMP</th>
<th>JSRC</th>
<th>NLW-SR</th>
<th>SAJSRC (L = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>97.47</td>
<td>97.69</td>
<td>97.68</td>
<td>98.27</td>
<td>98.09</td>
<td>99.96</td>
<td>97.99</td>
</tr>
<tr>
<td>2</td>
<td>98.32</td>
<td>99.11</td>
<td>96.97</td>
<td>98.10</td>
<td>96.46</td>
<td>98.76</td>
<td>94.21</td>
</tr>
<tr>
<td>3</td>
<td>99.37</td>
<td>99.93</td>
<td>99.97</td>
<td>99.68</td>
<td>96.33</td>
<td>99.91</td>
<td>99.95</td>
</tr>
<tr>
<td>4</td>
<td>98.01</td>
<td>99.11</td>
<td>97.59</td>
<td>96.70</td>
<td>97.01</td>
<td>99.60</td>
<td>99.60</td>
</tr>
<tr>
<td>5</td>
<td>96.04</td>
<td>99.73</td>
<td>99.95</td>
<td>98.48</td>
<td>99.82</td>
<td>99.96</td>
<td>99.96</td>
</tr>
<tr>
<td>6</td>
<td>99.83</td>
<td>96.92</td>
<td>98.15</td>
<td>97.95</td>
<td>85.62</td>
<td>98.63</td>
<td>99.26</td>
</tr>
<tr>
<td>7</td>
<td>91.23</td>
<td>88.08</td>
<td>98.42</td>
<td>86.90</td>
<td>80.55</td>
<td>90.75</td>
<td>95.78</td>
</tr>
<tr>
<td>8</td>
<td>86.23</td>
<td>92.05</td>
<td>88.18</td>
<td>89.82</td>
<td>95.45</td>
<td>94.11</td>
<td>97.81</td>
</tr>
<tr>
<td>9</td>
<td>86.99</td>
<td>90.95</td>
<td>90.42</td>
<td>79.37</td>
<td>93.79</td>
<td>95.61</td>
<td>97.15</td>
</tr>
<tr>
<td>10</td>
<td>91.42</td>
<td>93.93</td>
<td>90.79</td>
<td>89.68</td>
<td>99.78</td>
<td>99.42</td>
<td>99.52</td>
</tr>
<tr>
<td>11</td>
<td>91.67</td>
<td>94.42</td>
<td>91.26</td>
<td>82.77</td>
<td>100</td>
<td>98.25</td>
<td>98.73</td>
</tr>
<tr>
<td>12</td>
<td>87.05</td>
<td>92.61</td>
<td>91.32</td>
<td>81.94</td>
<td>98.11</td>
<td>96.54</td>
<td>99.66</td>
</tr>
<tr>
<td>13</td>
<td>78.16</td>
<td>82.70</td>
<td>81.45</td>
<td>35.55</td>
<td>90.17</td>
<td>72.28</td>
<td>83.89</td>
</tr>
<tr>
<td>14</td>
<td>97.42</td>
<td>95.84</td>
<td>99.86</td>
<td>98.18</td>
<td>99.38</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>99.49</td>
<td>98.99</td>
<td>100</td>
<td>98.40</td>
<td>98.49</td>
<td>99.53</td>
<td>99.83</td>
</tr>
<tr>
<td>OA</td>
<td>93.06</td>
<td>94.89</td>
<td>94.64</td>
<td>97.70</td>
<td>95.72</td>
<td>96.85</td>
<td>97.84</td>
</tr>
<tr>
<td>AA</td>
<td>93.25</td>
<td>94.81</td>
<td>94.82</td>
<td>88.78</td>
<td>95.31</td>
<td>96.25</td>
<td>97.56</td>
</tr>
</tbody>
</table>

κ = 0.925 0.945 0.942 0.889 0.954 0.966 0.977
in the 380 to 1050 nm region. Each band of the Houston image has 349 × 1905 pixels. The false color composite of the Houston image is shown in Fig. 10(a). Fig. 10(b) represents the ground truth of the image. About 10% labeled pixels are randomly chosen as the training samples and the remainder is the test samples. The detailed number of the training and test samples is shown in Table VII.

The classification maps of various classification methods are shown in Fig. 10(c)–(i) and the quantitative classification results (OA, AA, κ) of various classification methods can be seen in Table VIII. The results in Table VIII are obtained by averaging on five runs. According to the results shown in Table VIII, the proposed SAJSRC method has the best performance in terms of OA, AA, and κ. Especially, the proposed SAJSRC method has a more than 2% gain of OA over the window-based JSRC method, which confirm the effectiveness of using the shape-adaptive neighborhood instead of the fix-sized window.

E. Computational Complexity

This section aims at analyzing the computational complexity of the SAJSRC method. For an $M \times N \times B$ HSI, the time complexity of the PCA process is $O(B^3)$. According to the literature in [43] and the concept of the shape-adaptive algorithm, the convolution operator is the most time-consuming step of the shape-adaptive algorithm and its complexity for all direction $\{b_{k}\}$ and all lengths $\{h_{i}\} = \lambda_{i}, i = 1, 2, \ldots, m$ is $O(mkMN \log(MN))$, where $m$ is the dimension of the candidate length vector $H$ and $k$ denotes the total number of directions. The SOMP is used for the JSRC model. For each joint signal matrix in the SOMP, the number of the basic scalar multiplication with respect to the main step in the SOMP is $\sum_{i} (nLT - i) + 2(i^2L + i^3 + iLN)i = 1, 2, \ldots, \eta$, where $n$ is the number of the bands in signals, $L$ is the number of atoms in the dictionary $D$, and $\eta$ is the sparsity level in joint SR model. Since that $\eta$ is set to be comparably small in our method, the computational complexity of the SOMP for the whole HSI is upper-bounded by $O(MN(\eta LN T + 2 \eta^2L + \eta^2LN))$.

V. Conclusion

In this paper, a new SAJSRC method is proposed for HSI. Unlike previous SR-based classification approaches, the proposed SAJSRC method constructs a shape-adaptive local region for each test pixel instead of using a fixed square window. The shape-adaptive regions, which are constructed based on the spatial structure of the HSI, can be used to effectively and adaptively exploit spatial information. Then, the spatial information is incorporated into the joint SR model. Furthermore, to efficiently classify the HSI, the PCA data of the image are used for the joint SR instead of the original data. Experiments on three real HSIs demonstrate that the proposed SAJSRC method outperforms several widely used classification approaches.

In this paper, the shape-adaptive local regions are constructed based on a dimensionality reduced map. Another shape-adaptive region searching method, which uses the original HSI and the full spectral information, will be explored in our future work.

ACKNOWLEDGMENT

The authors would like to gratefully thank the editors and reviewers. Their valuable comments and suggestions greatly helped us improve the quality and presentation of this paper.

REFERENCES


Wei Fu (S’14) received the B.Sc. degree in electrical engineering from Hunan University, Changsha, China, in 2012. Currently, he is pursuing the Ph.D. degree in electrical engineering at Hunan University, Changsha, China.

His research interests include hyperspectral image processing, sparse representation, and synthetic aperture radar image analysis.

Shutao Li (M’07–SM’15) received the B.S., M.S. and Ph.D. degrees from Hunan University, Changsha, China, in 1995, 1997, and 2001, respectively, all in electrical engineering.

In 2001, he joined the College of Electrical and Information Engineering, Hunan University. From May 2001 to October 2001, he was a Research Associate with the Department of Computer Science, Hong Kong University of Science and Technology, Kowloon, Hong Kong. From November 2002 to November 2003, he was a Postdoctoral Fellow with the Royal Holloway College, University of London, Surrey, U.K., working with Prof. John Shawe-Taylor. From April 2005 to June 2005, he was a Visiting Professor with the Department of Computer Science, Hong Kong University of Science and Technology. Currently, he is a Full Professor with the College of Electrical and Information Engineering, Hunan University. He has authored or coauthored more than 160 refereed papers. His research interests include compressive sensing, sparse representation, image processing, and pattern recognition.

Dr. Li is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and a member of the Editorial Board of the Information Fusion and the Sensing and Imaging. He was a recipient of two 2nd-Grade National Awards at the Science and Technology Progress of China, in 2004 and 2006, respectively.

Leyuan Fang (S’10–M’14) received the B.S. degree in electrical engineering from Hunan University of Science and Technology, Xiangtan, China, in 2008. He joined the College of Electrical and Information Engineering, Hunan University, Changsha, China, in 2008, for the Ph.D. degree program.

From September 2011 to September 2012, he has been a Visiting Ph.D. Student at the Department of Ophthalmology, Duke University, Durham, NC, USA, supported by the China Scholarship Council. His research interests include sparse representation and multiresolution analysis in remote sensing and medical image processing.

Mr. Fang has won the Scholarship Award for Excellent Doctoral Student granted by the Chinese Ministry of Education in 2011.
Xudong Kang (S’13) received the B.Sc. degree in automation from Northeast University, Shenyang, China, in 2007. Currently, he is pursuing the Ph.D. degree in electrical engineering from Hunan University, Changsha, China.

From 2012 to 2013, he was working as a Visiting Ph.D. Student in electrical engineering at the University of Iceland, Reykjavik, Iceland. His research interests include image fusion, image super-resolution, pansharpening, and hyper-spectral image classification.

Mr. Kang was the recipient of the Scholarship Award for Excellent Doctoral Student granted by Chinese Ministry of Education, the Presidential Scholarship granted by Hunan University, and the National Graduate Scholarship.

Jón Atli Benediktsson (S’84–M’90–SM’99–F’04) received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, Iceland, in 1984, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1987 and 1990, respectively.

Currently, he is Pro Rector of Academic Affairs and a Professor of Electrical and Computer Engineering at the University of Iceland. He is a Co-founder of the biomedical startup company Oxymap. His research interests include remote sensing, biomedical analysis of signals, pattern recognition, image processing, and signal processing, and he has published extensively in those fields.

Prof. Benediktsson is the 2011–2012 President of the IEEE Geoscience and Remote Sensing Society (GRSS) and has been on the GRSS Administrative Committee since 2000. He was the Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS) from 2003 to 2008 and has served as an Associate Editor of TGRS since 1999 and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS since 2003. He is a Fellow of SPIE. He is a member of Societas Scientiarum Islandica and Tau Beta Pi. He received the Stevan J. Kristof Award from Purdue University in 1991 as an Outstanding Graduate Student in remote sensing. In 1997, he was the recipient of the Icelandic Research Councils Outstanding Young Researcher Award; in 2000, he was granted the IEEE Third Millennium Medal; in 2004, he was a corecipient of the University of Iceland’s Technology Innovation Award; in 2006, he received the yearly research award from the Engineering Research Institute of the University of Iceland; and in 2007, he received the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society. He is a corecipient of the 2012 IEEE Transactions on Geoscience and Remote Sensing Best Paper Award, and in 2013 the corecipient of the GRSS Highest Impact Paper Award. In 2013, he received the joint IEEE/VFI Award as Electrical Engineer of the Year.