Hyperspectral Image Classification Using Weighted Joint Collaborative Representation

Mingming Xiong, Qiong Ran, Wei Li, Member, IEEE, Jinyi Zou, and Qian Du, Senior Member, IEEE

Abstract—Recently, representation-based classifiers have gained increasing interest in hyperspectral image (HSI) classification. In this letter, based on our previously developed joint collaborative representation (JCR) classifier, an improved version, which is called weighted JCR (WJCR) classifier, is proposed. JCR adopts the same weights when extracting spatial and spectral features from surrounding pixels. Differing from JCR, WJCR attempts to utilize more appropriate weights by considering the similarity between the center pixel and its surroundings. Experimental results using two real HSIs demonstrate that the proposed WJCR outperforms the original JCR and other traditional classifiers, such as the support vector machine (SVM), the SVM with a composite kernel, and simultaneous orthogonal matching pursuit.

Index Terms—Collaborative representation based classifier, hyperspectral image (HSI) classification, nearest regularized subspace (NRS) classifier, sparse representation based classifier, spectral–spatial information.

I. INTRODUCTION

HYPERSPECTRAL image (HSI) classification, which aims at categorizing pixels into one of several land-use land-over classes, is an important application in the remote sensing field [1]–[3]. To date, numerous HSI classification techniques have been proposed [4]–[6]. Among these approaches, the support vector machine (SVM) [7] is capable of discriminating two classes by fitting an optimal separating hyperplane to the training data within a multidimensional feature space, and has shown excellent performance in HSI classification even with limited training samples. An improved SVM exploited the properties of Mercers conditions to construct a composite kernel (CK) for the combination of both spectral and spatial information, which is referred to as SVM-CK [8].

Recently, a sparse representation-based classifier has been proposed for HSI classification. In [9], two joint sparsity models have been proposed to incorporate the contextual information: In the one called simultaneous orthogonal matching pursuit (SOMP), pixels in a small neighborhood around the test pixel are simultaneously represented by a linear combination of labeled samples, whereas in the other one called OMP with smoothing (OMP-S), the smoothing constraint is imposed to force the vector Laplacian of the approximations to zero. In our previous work, a collaborative representation based classifier, named nearest regularized subspace (NRS) has been proposed [10]. NRS couples nearest subspace classification with the distance-weighted Tikhonov regularization and attempts to find a regularized collaborative representation for each testing sample via the linear combination of labeled samples within each class.

Similar to others, the NRS classifier was originally designed to be a pixel-wise classifier, i.e., only the spectral signature has been exploited while ignoring the spatial information at neighboring locations [11]. However, with the advancement of sensor technology, HSIs with high spatial resolution are gradually becoming available, and more and more classification methods have used joint spatial–spectral features to improve the classification performance [12], [13]. Joint collaborative representation (JCR) [14] can be viewed as an extension of previously proposed NRS with spatial information, by adjusting the contribution of labeled samples based on their similarity to the test sample. JCR seeks to incorporate the contextual information during classification; specifically, neighbors near the test pixel are simultaneously represented via a joint collaborative model based on linear combinations of labeled samples. In [14], we have observed that the joint collaborative model by consideration of neighboring pixels have much better performance than the original NRS.

However, we notice that JCR takes the surrounding pixels with the same weights, which is suboptimal, particularly in heterogeneous regions where the central pixel and neighboring pixels do not belong to the same class. Under such a case, only these neighboring pixels that are associated with the central pixels should be taken into consideration. Nevertheless, removal of the irrelevant pixels is not easy, which may increase additional computational complexity. Therefore, in this letter, we propose a simple but effective method to describe the contribution from a neighboring pixel neighboring pixel with adaptive weights. In the resulting WJCR, more appropriate weights are determined by using a Gaussian kernel function. The WJCR provides the benefit of efficiently extracting more accurate spectral–spatial features, which is particularly useful to data with a heterogeneous image scene.
II. PROPOSED CLASSIFIERS

Both NRS and JCR are to linearly represent each pixel under test using the available labeled samples with representation coefficients being constrained to have a minimum $\ell_2$ norm, and the pixel is assigned to the class producing the minimum representation residual. Consider a data set with training samples $X = \{x_i\}_{i=1}^N$ in $\mathbb{R}^d$ ($d$-dimensional feature space) and class labels $\omega_i \in \{1, 2, \ldots, C\}$, where $C$ is the number of classes, $n_l$ is the total number of training samples for the $l$th class, and $\sum_{l=1}^C n_l = n$. Let $Y = \{y_1, y_2, \ldots, y_m\}$ in $\mathbb{R}^d$ represent hyperspectral pixels in a local region where the central pixel is $y$.

A. JCR

In HSI, neighboring pixels may belong to the same material with high probability, and their spectral signatures are highly correlated. To exploit such local continuity, it is straightforward to spatially average all the pixels with its nearest neighbors. For the test sample $y$ located at the center of a local window, $\tilde{y} = (1/m) \sum_{i=1}^m y_i$, where $\tilde{y}$ represents the averaged value for the central pixel. For the training samples, $\tilde{x}_{i,j} = (1/m) \sum_{i=1}^m x_i$, where $\tilde{x}_{i,j}$ represents the averaged value for the pixel that belongs to the class $l$. The weight vector $\alpha_l$ for the linear combination is calculated according to

$$\alpha_l = \arg\min_{\alpha_l^T} \left\| \tilde{y} - \tilde{x}_{l,1} \alpha_l^T \right\|_2^2 + \lambda \left\| \Gamma_l \tilde{y} \alpha_l^T \right\|_2^2$$

(1)

where $\Gamma_l$, $\tilde{y}$ is a biasing Tikhonov matrix specific to each class $l$ and the averaged test sample, and $\lambda$ is a global regularization parameter that balances the minimization between the residual and regularization terms. Note that $\alpha_l^T$ is a certain representation of $\alpha_l$ with a size of $n_l \times 1$. Specifically, the regularization term is designed in the form of

$$\Gamma_l, \tilde{y} = \begin{bmatrix} \left\| \tilde{y} - \tilde{x}_{l,1} \right\|_2 & 0 \\ \vdots & \ddots \\ 0 & \left\| \tilde{y} - \tilde{x}_{l,n} \right\|_2 \end{bmatrix}$$

(2)

Then, the weight vector $\alpha_l$ can be estimated in a closed-form solution, i.e.,

$$\alpha_l = \left( \tilde{x}_l^T \tilde{x}_l + \lambda \Gamma_l^T \Gamma_l \tilde{y} \right)^{-1} \tilde{x}_l^T \tilde{y}$$

(3)

After that, the residual between the approximation and averaged test pixel is $r_l(y) = \left\| \tilde{x}_l \alpha_l - \tilde{y} \right\|_2$. Moreover, the class label of test pixel $y$ is finally determined by

$$\text{class}(y) = \arg\min_{l=1, \ldots, C} r_l(y).$$

(4)

In doing this, the smoothness among the neighbors is considered in the process of estimating weight vector $\alpha_l$, which causes the approximations to have similar spectral characteristics to its surrounding neighbors.

B. Proposed WJCR

In WJCR, the similarity between the center pixel and a neighboring pixel is measured to determine the contribution from the neighboring pixel adaptively. We notice that JCR actually takes equal weights for all the surrounding pixels, which may be suboptimal for heterogeneous regions, as shown in Fig. 1. Fig. 1(a) illustrates a $5 \times 5$ region with 25 spatial neighborhood pixels, where the same color indicates the similar material. In Fig. 1(b), JCR assumes that these weights are uniform, whereas the proposed WJCR does not. Thus, we expect a process to reduce the weight values of these neighbors obviously distinctive from the central pixel. Here, the Gaussian kernel function [15] is employed for adaptive weight assignment

$$w_i' = \exp \left( -\frac{\left\| y - y_i \right\|^2}{2\sigma^2} \right)$$

(5)

where $w_i'$ represents the relationship between the central pixel $y$ and its $i$th neighbor $y_i$, and $\sigma$ is a parameter set by the median value of $\left\| y_i - \tilde{y} \right\|_2$, $i = 1, 2, \ldots, m$, where $\tilde{y} = (1/m) \sum_{i=1}^m y_i$ is the mean of all available data [16]. There are two major reasons for choosing Gaussian kernel function: 1) its output is positively proportional to spectral similarity (which means two very similar pixels have a larger distance (close to one) as wanted); and 2) its output is exponentially decreased when two pixels are very dissimilar, which means the contribution from a dissimilar pixel will be significantly suppressed as needed. Thus, the Gaussian kernel function is preferred here, rather than other distance metrics such as the Euclidean distance or its inverse. Thus, for all the pixels in $Y$, we calculate a weight vector $w = \{w_1, w_2, \ldots, w_m\}$ after normalizing each weight

$$w_i = \frac{w_i'}{\sum_{i=1}^m w_i'}$$

(6)

After obtaining the weight vector, the newly weighted central pixel can be represented as

$$\tilde{y} = Y w$$

(7)

which can be used in (1). The similar computation is also implemented for the training samples, and the following process is the same as the aforementioned JCR.

III. EXPERIMENTS AND ANALYSIS

A. Experimental Data

In this section, we evaluate the classification performance of the proposed WJCR by comparing with several traditional
methods, such as NRS, JCR, SVM, SVM-CK, and SOMP. The first experimental data set in our experiments was acquired by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor during a flight campaign over Pavia, northern Italy. The image has 115 spectral bands with a spectral coverage from 0.43 to 0.86 $\mu m$ and a spatial resolution of 1.3 m. The university area in our experiments has 103 spectral bands after removing the noisy and water absorption bands, and image scene used in the experiment has 610 $\times$ 340 pixels.

The second experimental hyperspectral data set employed was acquired using the HyMap sensor. The scene is about an area close to Purdue University cropped into a subimage with 377 $\times$ 512 pixels and 126 bands spanning the wavelength interval of 0.45 to 2.5 $\mu m$ with a spatial resolution of 3.5 m. There are six classes (i.e., Road, Grass, Shadow, Soil, Tree, and Roof) for this subimage scene. In our experiments, training samples are randomly selected from the ground-truth map multiple times to ensure fair assessment, and all the remaining labeled pixels as tested. The numbers of labeled samples are 1287, 1114, 219, 379, 1351, and 1285 for each class, respectively.

### B. Classification Performance

We first investigate the regularization parameter $\lambda$ of the proposed WJCR. As a global parameter, the adjustment of $\lambda$ is important to the classification performance. Its sensitivity along with the varied window size is shown in Fig. 2. In Fig. 2, we also include the classification accuracy versus varying windows. Since we adaptively assign the weights to the neighbors, even when the window size is very large, only samples highly correlated to the central pixel are used to contribute the representation. This means WJCR is relatively robust to the selection of window size. As shown in Fig. 2(a), a window size of 21 $\times$ 21 actually provides the best performance for the University of Pavia data. For subscene of Purdue data, the window size can be as large as 13 $\times$ 13. The optimal $\lambda$ is $10^{-3}$ and $10^{-2}$ for these two...
TABLE I
CLASSIFICATION ACCURACY (%) FOR THE UNIVERSITY OF PAVIA DATA WITH 30 TRAINING SAMPLES PER CLASS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Asphalt</td>
<td>30</td>
<td>Train</td>
<td>66.13</td>
<td>86.64</td>
<td>79.56</td>
<td>77.81</td>
<td>93.93</td>
<td>97.87</td>
</tr>
<tr>
<td>Bare soil</td>
<td>30</td>
<td>Test</td>
<td>186.49</td>
<td>71.55</td>
<td>80.15</td>
<td>76.15</td>
<td>95.64</td>
<td>98.72</td>
</tr>
<tr>
<td>Bitumen</td>
<td>30</td>
<td>SVM</td>
<td>2090.00</td>
<td>83.34</td>
<td>85.72</td>
<td>70.51</td>
<td>93.81</td>
<td>99.19</td>
</tr>
<tr>
<td>Bricks</td>
<td>30</td>
<td>SVM-CK</td>
<td>3064.00</td>
<td>92.92</td>
<td>93.86</td>
<td>94.22</td>
<td>92.85</td>
<td>95.79</td>
</tr>
<tr>
<td>Grass</td>
<td>30</td>
<td>SOMP</td>
<td>1345.00</td>
<td>99.03</td>
<td>99.63</td>
<td>99.93</td>
<td>99.55</td>
<td>100.00</td>
</tr>
<tr>
<td>Meadows</td>
<td>30</td>
<td>NRS</td>
<td>5029.00</td>
<td>93.36</td>
<td>92.68</td>
<td>87.08</td>
<td>81.31</td>
<td>91.79</td>
</tr>
<tr>
<td>Metal sheets</td>
<td>30</td>
<td>JCR</td>
<td>1330.00</td>
<td>96.89</td>
<td>95.86</td>
<td>93.23</td>
<td>90.27</td>
<td>98.42</td>
</tr>
<tr>
<td>Shadow</td>
<td>30</td>
<td>WJCR</td>
<td>3682.00</td>
<td>99.11</td>
<td>98.59</td>
<td>93.61</td>
<td>81.48</td>
<td>89.84</td>
</tr>
<tr>
<td>Trees</td>
<td>30</td>
<td>Overall Accuracy</td>
<td>947.00</td>
<td>99.68</td>
<td>99.89</td>
<td>65.58</td>
<td>99.47</td>
<td>98.81</td>
</tr>
<tr>
<td>Time(s)</td>
<td></td>
<td></td>
<td>78.20</td>
<td>89.81</td>
<td>73.29</td>
<td>80.17</td>
<td>94.13</td>
<td>97.90</td>
</tr>
</tbody>
</table>

TABLE II
CLASSIFICATION ACCURACY (%) FOR THE HYMAP DATA WITH TEN TRAINING SAMPLES PER CLASS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Road</td>
<td>10</td>
<td>Train</td>
<td>1287.00</td>
<td>90.13</td>
<td>96.27</td>
<td>95.88</td>
<td>83.45</td>
<td>92.23</td>
</tr>
<tr>
<td>Grass</td>
<td>10</td>
<td>Test</td>
<td>1114.00</td>
<td>98.29</td>
<td>97.76</td>
<td>89.50</td>
<td>95.78</td>
<td>99.55</td>
</tr>
<tr>
<td>Shadow</td>
<td>10</td>
<td>SVM</td>
<td>219.00</td>
<td>99.72</td>
<td>99.72</td>
<td>89.30</td>
<td>98.17</td>
<td>88.58</td>
</tr>
<tr>
<td>Soil</td>
<td>10</td>
<td>SVM-CK</td>
<td>579.00</td>
<td>93.67</td>
<td>98.15</td>
<td>98.15</td>
<td>97.10</td>
<td>99.47</td>
</tr>
<tr>
<td>Tree</td>
<td>10</td>
<td>SOMP</td>
<td>1351.00</td>
<td>93.63</td>
<td>95.56</td>
<td>99.85</td>
<td>99.11</td>
<td>99.48</td>
</tr>
<tr>
<td>Roof</td>
<td>10</td>
<td>NRS</td>
<td>1285.00</td>
<td>63.35</td>
<td>76.03</td>
<td>90.35</td>
<td>80.78</td>
<td>92.45</td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>87.01</td>
<td>91.96</td>
<td>94.09</td>
<td>90.52</td>
<td>95.81</td>
<td>97.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time(s)</td>
<td></td>
<td></td>
<td>2</td>
<td>29</td>
<td>19</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

Fig. 5. Thematic maps resulting from classification for the ROSIS University of Pavia data set using 30 training samples per class. (a) Pseudocolor image. (b) Ground-truth map. (c) SVM. (d) SVM-CK. (e) SOMP. (f) NRS. (g) JCR. (h) WJCR.

data, respectively. For other methods in our experiments, leave-one-out cross-validation strategy based on available training samples is considered for parameter tuning.

Figs. 3 and 4 show the classification accuracy versus different numbers of training samples per class for these two experimental data. It is obvious that the proposed WJCR outperforms other methods. We can mainly focus on the comparison between WJCR and JCR since the former can be viewed as a direct extension of the latter. For the University of Pavia data, when the number of training samples changes from 10 to 50, WJCR is always superior to JCR, which confirms the effectiveness of adaptive weights.

We then list the overall classification accuracy and accuracy for each class using aforementioned methods in Tables I and II.
The number of training data is randomly selected 20 times, and the results are averaged. In Table I, the accuracy of the proposed WJCR is 97.90%, with an improvement of 3% compared with that of JCR, and we observe that most of the highest accuracy levels of each class are provided by the proposed WJCR. We also provide the computing time (in seconds) in the tables. All experiments were carried out using MATLAB on an Intel Core 2 Duo CPU machine with 4 GB of RAM. It should be noted that SVM is implemented in the SVM package that uses MEX function to call C program in MATLAB. Fig. 5 further illustrates the classification maps for these data. The results are consistent with those in Table I. The standardized McNemar’s test [17] has been employed to verify the statistical significance in accuracy improvement of the proposed methods. As listed in Table III, the \( z \) values of McNemar’s test larger than 1.96 and 2.58 mean that two results are statistically different at the 95% and 99% confidence levels, respectively.

### IV. Conclusion

In this letter, an extension of the previously developed JCR was proposed. Adaptive weights were assigned to the surrounding neighbors based on spectral similarity in the proposed WJCR, which can generate more accurate spectral–spatial features in HSI classification. Experimental results indicated that the proposed WJCR was generally better than JCR, and some other classifiers, such as NRS, SVM, SVM-CK, and SOMP. It should be noted that, similar to other spectral–spatial classifiers, JCR and WJCR require training samples with spatial information being available.

### REFERENCES


