Hyperspectral Images Classification via Weighted Spatial-Spectral Principle Component Analysis

Fang HE, Wei-min JIA, Qiang YU, Rong WANG and Shan LIU
Xi'an Research Institute of High-tech, Xi'an 710025, China

Keywords: Hyperspectral images classification, Weighted spatial and spectral principle component analysis, reconstruct.

Abstract. In order to improve the Hyperspectral images (HSI) classification accuracy and to preprocess HSI by fully using the spatial and spectral information, a new spatial-spectral dimensionality reduction method, Weighted Spatial and Spectral Principle Component Analysis (WSSPCA), was proposed. This algorithm reconstructed the HSI by using the physical characters of HSI. Then, principle Component Analysis (PCA) was utilized to reduce dimensionality of HSI. The new method not only can lower the influence of singular point in HSI, but also reduce the redundancy between bands, which improves the HSI classification accuracy efficiently. The benchmark tests on Indian Pines and PaviaU demonstrate that the performance of WSSPCA is better than PCA and LPP when 10%, 5% samples in each class. The best values of kappa coefficient obtained by WSSPCA are 89.61% and 95.59% respectively on the HSI datasets, exceeding the baseline 20.5% and 19.38%.

Introduction

Hyperspectral remote sensing technology has been rapidly developed since in the 1980 of the 20th century [1], and widely used for environmental science, geological research, precision agriculture, and military [2], [3]. The hyperspectral image (HSI) is the remote sensing image obtained by imaging spectrometer. The main characteristic of HSI is fusing traditional of space dimension and spectrum dimension information for one, forming "maps one" of super dimension data [4]. The high of spectrum dimension number and spectrum resolution for features brings great opportunity for HSI classification, while also proposes challenges to traditional of image classification recognition algorithms [5].

Hyperspectral images are made up of a large number of single band images. The rapid expansion of the amount of data increased the complexity of data processing, reducing the efficiency of data processing [6]. As the spectral resolution improves, the redundancy between adjacent bands is increasing, which is seriously reducing the accuracy of traditional algorithm for HSI classification. When the sample is limited, the HSI classification will experience "dimensionality curse" [7]. Dimensionality reduction (DR), which searches a low-dimensional representation that contains useful information for high dimension data, is an important part of HSI classification preprocessing [6]. Meanwhile, DR can reduce the amount of data and look for a simpler and clearer representation of data, which makes for HSI classification [8].

In recent years, the traditional DR methods are linear discriminant analysis (LDA) [9], nonparametric weighted feature extraction (NWFE) [10], principle component analysis (PCA) [11][12][13], locality preserving projection (LPP) [14], and neighborhood preserving embedding (NPE) [15]. However, the above DR methods simply use the spectral information of HSI and ignore the spatial information which can not reveal the intrinsic relationship between different samples.

As the distribution of samples in HSI is spatial continuity, this paper proposes a new DR method of fusion of hyperspectral image space and spectral information, weighted spatial spectral principle component analysis (WSSPCA). This new algorithm uses the neighborhood space between the pixels to reconstruct HSI and PCA is used to reduce the dimension of reconstructed image.
HSI Reconstructed Based on Weighted Spatial and Spectral Information

The samples of HSI are spatial continuity that is the local contiguous region of neighboring pixels are made up of the same material and belong to the same class. Using the weighted spatial and spectral information to reconstruct the HSI can smooth HSI.

In HSI, setting the coordinates of pixel is \(x_i\), the neighborhood space of \(x_i\) is:

\[
N(x_i) = \{x \triangleq (p, q)\}, \quad p \in \{p_i - a, p_i + a\}, \quad q \in \{q_i - a, q_i + a\}
\]

where \(a = (w-1)/2\), \(w\) is odd number, which presents the window scale of neighborhood space. The pixels \(x_i, x_{i+1}, x_{i+2}, \ldots, x_n\) are in the neighbor space of \(N(x_i)\). The number of nearest neighbor point is \(s = w^2 - 1\).

The reconstructed pixel can be gotten by using the following rule:

\[
\hat{x}_i = \frac{\sum_{x \in N(x_i)} v_i x_j}{\sum_{x \in N(x_i)} v_j} = x_i + \frac{\sum_{s=1}^{w^2} v_i x_{i,k}}{1 + \sum_{s=1}^{w^2} v_i}
\]

where the weight of any pixel \(x_{i,k}\) to the center pixel \(x_i\) in the neighbor space of is \(v_i = \exp(-\gamma_0 |x_i - x_{i,k}|)\). The parameter \(\gamma_0\) is the spectral factor which can be used to measure the degree of interaction between the different pixels in the same neighborhood space. The physical meaning of \(\gamma_0\) is the closer the spectral curves between the pixels, the more interaction with each other, whereas the smaller.

This algorithm can adjust the size of neighborhood space by the neighbors window scale and gives the smaller weight to reduce the interference of singular points by the spectral factor, achieving to smooth the HSI. Parameter \(w\) and \(\gamma_0\) are important for the final result, thus we used experimental research method to select.

The pixels in the edges or corners are preprocessed as Fig.1. Each square grid represents a pixel in HSI, the light gray grid as the center pixel and the dark gray grid as the filled method. When the pixel is at the edge or corner location, its neighbor pixel is filled with.

![Figure 1. Neighbor space of pixel under different conditions.](image)

The Algorithm of Weighted Spatial and Spectral Information Principle Component Analysis

PCA, a classical and effective method for feature extraction, can reduce the correlation between HSI data by using several principal components which are unrelated to present the original energy information of multiple bands, when applied in HSI DR. WSSPCA integrates the spatial and spectral information of HSI to reduce the dimension, which is helpful for HSI classification.

The procedures of WSSPCA are as follows. The reconstructed HSI can be represented as \(\hat{X}\) whose size is \(n \times d\), where \(n\) is the total pixel in each band and \(d\) is the number of all bands. Then the mean of each image in HSI is \(\mu = \frac{1}{n}\sum_{i=1}^{n} \hat{X}_i\).

The covariance matrix of \(\hat{X}\) is:
\[ S = \frac{1}{n} \sum_{i=1}^{n} (\hat{x}_i - \mu)(\hat{x}_i - \mu)^T \]  

(3)

Then the eigenvalues and eigenvectors of the covariance matrix are:

\[ S_{\nu i} = \hat{x}_i x_i^T, \quad (i = 1, 2, \ldots, n) \]  

(4)

Choosing \( k \) components of HSI to form the feature space: \( \{w_1, w_2, \ldots, w_k\} \)

The reconstructed data are projected on the feature space:

\[ y = W^T (\hat{x} - \mu) \]  

(5)

The nearest neighbor classifier is used to classify the projected data and the classification accuracy evaluation index is counted to measure the result.

**HSI Classification by Weighted Spatial and Spectral Information Principle Component Analysis**

The concrete steps of HSI classification by weighted spatial and spectral information principle component analysis is as the follows:

- **Input:** HSI data \( X = \{x_1, \ldots, x_n\}^T \), the scale parameter \( w \) and spectral factor \( \gamma_0 \).
- **Output:** the classification accuracy evaluation index.
- **Step 1:** the data set’s neighboring space is selected according to Eq. 1;
- **Step 2:** for any pixel \( x_i \) in dataset \( X \), the \( x_i \) is reconstructed according to Eq.2 to get the new data \( \hat{X} \);
- **Step 3:** the covariance matrix of \( \hat{X} \) is counted according to Eq.3 and computing the eigenvalues and eigenvectors of the covariance matrix. \( k \) components of HSI are chosen to form the feature space.
- **Step 4:** the train and test samples are chosen from the data set according to certain rules. The reconstructed data are projected on the feature space and the nearest neighbor classifier is used to classify the projected data. The classification accuracy evaluation index is counted.

![Figure 2](image1.png)

(a) False color  
(b) Corresponding ground truth  

Figure 2. The Indian-pines hyperspectral image.

![Figure 3](image2.png)

(a) False color  
(b) Corresponding ground truth  

Figure 3. The PaviaU hyperspectral image.

**Experiment Result and Analysis**

**Hyperspectral Data Sets**

**Indian Pines.** The data set was acquired by the AVIRIS sensor in 1992. The image has 220 bands, in which 20 bands are discarded because of the water and atmospheric affection. Each band has 145×145 pixels. The total number of samples is 10249 ranging from 20 to 2455 in each class. The false color composition of bands 50, 27, and 17 and the ground-truth map are shown in Fig.2 (a) and Fig.2 (b).

**PaviaU.** The data set was acquired in 2001 by the ROSIS instrument over the city of Pavia, Italy. This image scene corresponds to the University of Pavia and has the size of 610×340 pixels and 115 bands. After removing 12 water absorption and noisy bands, 103 bands were used in the
experiments. The data contains 9 ground-truth classes. The false color composition of bands 50, 27, and 17 and the ground-truth map are shown in Fig.3 (a) and Fig.3 (b).

**Experimental Result and Analysis**

WSSPCA, PCA, and LPP are used to reduce the dimension of HSI and the nearest neighbor classifier is used to classify the data after DR. The result without DR is acted as the baseline. The best parameter of WSSPCA is chosen by experimental analysis. The weight matrix $W$ of LPP is constructed by the heart method. The HSI classification accuracy evaluations are overall accuracy (OA), average accuracy (AA), and kappa coefficient ($k$). OA is the percentage of correctly classified pixels in the test set, whereas AA is the mean of class-specific accuracy values. $k$ is the percentage of agreement corrected by the number of agreements that would be expected purely by chance.

**HSI Classification on Indian Pines Data Set.** The nearest window scale and spectral factor are two main parameters in WSSPCA, which can be chosen from experiment. We select randomly 10% of samples per class for training, and the remaining samples are used for testing. For very small classes, we take a minimum of ten training samples per class. Fig.4 shows the relationship between classification accuracy and different parameter on Indian Pines dataset. As shown in Fig.4a, when the neighbor window scale is 9, AA, OA, and kappa coefficient can get the optimal result. From Fig.4b, when the spectral factor is 1, AA, OA, and kappa coefficient reached the maximum value. The best parameters on Indian Pines data set are $w = 9, \gamma = 1.0$.

(a) The relationship between Classification Accuracy (%) and neighborhood scale

(b) The relationship between Classification Accuracy (%) and spectral factor

Figure 4. The relationship between Classification Accuracy (%) and different parameter on Indian Pines dataset.

To compare the performance of different algorithms, we repeated 10 times classification experiments, and computed the mean of AA, OA, and kappa coefficient. Fig.5 shows the relationship between classification accuracy and dimensions on Indian Pines dataset in the first 100 dimensions under the different algorithms. Tab.1 shows the best result and the corresponding dimensions obtained by each algorithm on the Indian Pines dataset.

From Fig.5, we can know that the classification results obtained by WSSPCA algorithm are superior to PCA, LPP algorithms. The maximum classification accuracy gotten by WSSPCA is far beyond the baseline, but the results obtained by PCA, and LPP are slightly lower than the baseline. The reason is that WSSPCA effectively used the spatial and spectral information, which can improve the classification accuracy.

From Table 1, we can see that WSSPCA, as a new spatial-spectral algorithm of HSI, can get the best result under the same conditions. The maximum value of OA which is obtained by WSSPCA algorithm is 90.90%, exceeding the baseline 17.93%, and kappa coefficient is 89.61%, beyond the baseline 20.50%. Fig.5 shows that WSSPCA proposed in this paper performs best.

Fig.6 shows the classification maps of the ground truth, train samples, test samples, the results of no DR, PCA, LPP, WSSPCA when the kappa coefficient is maximum.
Average Classification Accuracy (%) | Overall Classification Accuracy (%) | | \( \gamma \) (%)
--- | --- | ---
Baseline | | |
PCA | 72.97(1) | 69.11(1) |
LPP | 72.94(20) | 69.08(20) |
WSSPCA | 90.90(21) | 89.61(21) |

**HSI Classification on PaviU Data Set.** In this experiment, we choose randomly 5% of samples per class for training, and the remaining samples are used for testing. The same as Indian Pines data set, the parameters are chosen by experiments. The optimal parameters are \( w = 15, \gamma = 2.0 \).

We repeated 10 times classification experiments, and computed the mean of AA, OA, and kappa coefficient as the final result. Fig.7 shows the relationship between classification accuracy and dimensions on PaviU dataset in different dimensions under the different algorithms. Tab.2 shows the best result and the corresponding dimensions obtained by each algorithm on PaviU dataset.

From Fig.7 and Table 2, we can know that the classification results obtained by WSSPCA algorithm are superior to PCA, LPP algorithms. The maximum value of OA which is obtained by WSSPCA algorithm is 96.69%, exceeding the baseline 14.36%, and kappa coefficient is 95.59%, beyond the baseline 19.38%. However, the results obtained by PCA and LPP are always lower than the baseline.

Fig.8 shows the classification maps of the ground truth, train samples, test samples, the results of no DR, PCA, LPP, WSSPCA when the kappa coefficient is maximum. From Fig.8, we can see the classification map obtained by WSSPCA performs best. The error or missed phenomenon obtained by WSSPCA is the least.
The relationship between AA(%) and dimensions (b) The relationship between OA(%) and dimensions (c) The relationship between kappa(%) and dimensions

Figure 7. The relationship between classification accuracy (%) and dimensions on PaviU dataset.

(a) Ground truth (b) Train samples (c) Test samples (d) Baseline

(e) PCA (f) LPP (g) WSSPCA

Figure 8. Classification results of different algorithms in PaviU data set.

Table 2. The best results (%) with optimal dimensions.

<table>
<thead>
<tr>
<th>method</th>
<th>OA(dimension)</th>
<th>kappa(dimension)</th>
</tr>
</thead>
<tbody>
<tr>
<td>baseline</td>
<td>82.33(1)</td>
<td>76.21(1)</td>
</tr>
<tr>
<td>PCA</td>
<td>82.35(20)</td>
<td>76.25(16)</td>
</tr>
<tr>
<td>LPP</td>
<td>82.42(5)</td>
<td>76.48(5)</td>
</tr>
<tr>
<td>WSSPCA</td>
<td>96.69(20)</td>
<td>95.59(21)</td>
</tr>
</tbody>
</table>

Conclusion

This paper proposed a new weighted spatial and spectral principle component analysis (WSSPCA) algorithm. This algorithm considers the physical characteristics to reconstruct the HSI by using the nearest window scale and spectral factor, which can effectively eliminating the influence of singular point. PCA is used to reduce the dimension of the reconstructed data, which can lower the redundancy among the bands. On the Indian Pines and PaviU data sets, the experimental results show that the new method proposed in this paper is superior to PCA and LPP algorithms. On Indian Pines data set, the maximum OA is 90.90% and kappa coefficient is 89.61% when 10% samples are randomly selected as training samples in each class. On PaviU data set, when 5% samples are randomly selected as training samples in each class, the maximum OA is 96.69% and kappa coefficient is 95.59%. The WSSPCA algorithm effectively improved the classification accuracy.
References


