LAND COVER CLASSIFICATION USING MULTITEMPORAL CHRIS/PROBA IMAGES AND MULTITEMPORAL TEXTURE

Huiran Jin, Peijun Li, Wenjie Fan

Institute of Remote Sensing and GIS, Peking University, Beijing 100871, P R China

ABSTRACT

Most existing multitemporal classification researches use spectral information alone. However, adding spatial structure and temporal correlation in the classification could improve the classification accuracy. This paper proposed a new method to extract multitemporal texture by the Pseudo Cross Variogram (PCV). The derived texture features were combined with the original spectral information for multitemporal classification. The performance of the proposed multitemporal texture was evaluated in land cover classification using bi-temporal hyperspectral CHRIS/PROBA images. The experiments showed that CHRIS/PROBA data is applicable in multitemporal classification, and including multitemporal texture in multitemporal classification could lead to a significant increase in overall classification accuracy, compared to the classification using spectral information alone.

Index Terms— CHRIS, pseudo-cross variogram, multitemporal texture, multitemporal classification

1. INTRODUCTION

Land use/land cover information is of great importance for scientific research and applications. Hyperspectral images have powerful capabilities of describing and discriminating land cover types. However, it is hard to obtain satisfying results with hyperspectral data alone, as many land cover classes, in particular vegetation classes, show similar spectral signatures. Therefore, other information, such as temporal correlation and spatial structure are incorporated in image classification in order to achieve a higher accuracy [1-3].

Most previous researches on multitemporal classification only use spectral information, without considering temporal and spatial correlation. In this study, a multitemporal texture was combined with the spectral information to improve the classification accuracy, which has not been applied in multitemporal classification study yet. Multitemporal texture features were extracted by pseudo-cross variogram, proposed by Clark et al. [4], numerically characterizing the spatial and temporal correlation between the bi-temporal images at the same time. The performance of the proposed multitemporal texture measure was evaluated in terms of classification accuracy.

Two CHRIS/PROBA images acquired on different dates were used, in order to find out whether CHRIS/PROBA data is applicable in multitemporal classification.

2. DATA AND STUDY AREA

The CHRIS, mounted on the PROBA satellite, is one of the first space-borne hyperspectral sensors offering high spatial resolution. It provides five spatial and spectral configurations (MODE), and acquires a set of images of the same scene at different angles during the same orbit. CHRIS/PROBA collects 18, 37 or 63 spectral bands at spatial resolution of 18 m or 36 m with a spectral coverage ranging from 405 nm until 1050 nm. The ground swath of each scene is 13.5 km. In this study, two scenes of Mode 5, acquired from the nadir direction with 37 spectral bands at 18m high spatial resolution, were selected for the test.

The study site is located in the Hengshui area of Hebei Province, China. It is a typical agricultural area, and main crops in the area are cotton, corn and wheat. Cotton grows from May to November. Corn grows from June to October. Wheat is cultivated in October and harvested in June of the next year.

In order to map the land use/land cover types in the area using CHRIS/PROBA data, it is ideal to use images acquired in July or August, since most crops are growing prosperously during the period. However, we are not able to obtain proper data in two years continuously, limited by weather conditions. Thus, the bi-temporal CHRIS/PROBA scenes, dated 3 May 2007 and 2 November 2007, were applied to multitemporal classification in this study.

A subset of 370×648 pixels was selected from each scene. Six land cover types were recognized in the area, i.e. wheat, cotton, orchard, water, residence and bare land. Based on the information from existing land cover maps and field observations, 5109 pixels were selected as training samples, including 1166 pixels for wheat, 1159 pixels for cotton, 816 pixels for orchard, 504 pixels for water, 1143 pixels for residence, and 231 pixels for bare land for image classification. The results were assessed through test samples generated by the stratified random strategy. One thousand reference pixels were selected at first. Some pixels located on the edge of different land cover types and hard to determine the attributes were omitted from the assessment. Finally, 983 pixels were used as test samples.

3. METHODS

The procedure adopted in this paper can be summarized as follows. Stripings were first removed from the bi-temporal CHRIS/PROBA images. Then, multitemporal texture was extracted using pseudo-cross variogram. After feature extraction by Principal Component Analysis (PCA), the derived textural components were combined with the spectral components for image classification. The result was quantitatively compared to the classification using spectral information alone in terms of classification accuracy.

3.1. Pre-processing
CHRIS/PROBA images are mainly affected by two kinds of noise called horizontal and vertical noises. These noises should be first eliminated.

Horizontal noise consists in the random loss of partial data from some lines of the images. The lost lines always appear in different bands and positions. When that noise appears, there is not a complete loss of the whole line, maintaining a correct value followed by an incorrect one. Therefore, it is easy to detect and destripe horizontal noise. The errors are corrected using the average of the four nearest pixels with a correct value [5].

Vertical noise is due to two main factors: errors in the alignment of the sensors in the construction of the instrument, which can be considered as a constant, as well as thermal fluctuations during the orbit that causes small variations in the alignment of the optical elements, making that noise variable at all times. Recently, several destriping algorithms with focus on the specific characteristics of CHRIS have been developed [5-6]. In this study, the method developed by Barducci et al. [6] was used to correct the vertical striping. The corrected image was obtained by dividing the noisy signal by the correcting profile, which represents the shape of noise pattern as the ratio between original and smoothed profile. Since the correcting profile oscillates around the value 1, the newly defined image has on average the intensity distribution the same as the original signal. However, vertical striping effect is well eliminated.

After removing horizontal and vertical noises, the bi-temporal images also need to be co-registered exactly for further applications like multitemporal classification of land cover types.

### 3.2. Multitemporal texture by Pseudo Cross Variogram

Generally, texture features characterize the spatial variability of digital numbers (DNs) of a satellite image. Traditional image texture is usually extracted individually from a single spectral band, describing the spatial variability of land cover types within the single-band image. The problem is that it only uses the texture information from a selected band of multispectral images. Recently, some two-band texture measures have been proposed to express the joint spatial variability, i.e., cross correlation, which is a positive complement to the single-band texture. In this study, a special kind of two-band texture, bi-temporal texture, was extracted from the same band of bi-temporal images. The bi-temporal texture was measured by pseudo-cross variogram [4], a geostatistical tool.

Geostatistical methods were used in image texture extraction and land cover classification in recent years [2-3,7-8]. Traditionally, univariate variogram function has been widely accepted as a measure of texture features, characterizing the spatial autocorrelation of radiometric data within a simple band. Under the intrinsic hypothesis, the variogram is defined as follows:

\[
\gamma(h) = \frac{1}{2} E \left[ DN(x) - DN(x + h) \right]^2
\]  

where \( h \) is a distance vector, and \( \gamma(h) \) represents half of the mathematical expectation of the quadratic increments of pixel pair values at a distance \( h \) between the pixels \( x \) and \( x + h \), i.e., semivariance; \( \gamma(h) \) is a vectorial function depending on the modulus and the angle of \( h \).

The Pseudo Cross Variogram (PCV) quantifies the joint spatial variability (cross correlation) between two regionalized variables [9], defined as:

\[
\gamma_{ij}(h) = \frac{1}{2} E \left[ DN_i(x) - DN_j(x + h) \right]^2
\]  

where \( h \) is a distance vector. The PCV measures the variance of the cross differences of two spectral bands. In remote sensing applications, the PCV was used as a spatial texture measure which describes the spatial correlation between two different bands [2].

It can also express the cross correlation between the values of the same variable at different times, e.g., the values of the same band acquired on different dates [10-11]. In this study, the PCV is considered as the temporal texture, which describes the temporal and spatial correlation between radiometric values of the same band acquired on different dates.

The experimental PCV can be computed as follows:

\[
\gamma_{ij\_exp}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[ DN_i(x) - DN_j(x + h) \right]^2
\]  

where \( N(h) \) is the number of pixel pairs with a certain distance vector \( h \), \( DN_i(.) \) and \( DN_j(.) \) denote digital values of pixels \( x \) and \( x + h \), located in the two bands \( j \) and \( k \), respectively. Thus, the experimental PCV is simply computed by the semivariance of the cross increments. When \( j \) and \( k \) denote the same band acquired on different dates, the function above is considered to be a powerful bi-temporal texture measure. Compared with the single-temporal texture, bi-temporal texture is capable to numerically describe the temporal correlation as well as the spatial structure of land cover types at the same time.

A geostatistical image texture is obtained by computing the PCV within a neighborhood using a moving window. For a specific lag \( h \), the average of PCV, considered as the multitemporal texture, is assigned to the central pixel of the moving window. Therefore, window size and lag distance (including size and direction) are two crucial parameters. The window should not be too large in order to avoid the influence of adjacent land cover types, which would reduce the classification accuracy; on the other hand, the window should not be too small either, as otherwise the texture extracted would not reflect the spatial structure correctly. In this study, the most proper pair of parameters was confirmed after several experiments to best quantify the spatial variability at the local level. It is showed that the derived multitemporal texture by the proposed method is effective in multitemporal image classification.

### 3.3. Feature extraction by PCA

Since every CHRIS/PROBA image contains 37 spectral bands, there are a total of 74 spectral bands in the bi-temporal images, from which 37 texture features were extracted by PCV. Supposing the derived textural information is simply combined with the spectral information, which means a total of 111 bands are used in image classification, it might cause Hughes phenomenon owing to too much feature bands. Therefore, feature extraction is desirable.

In this study, Principal Component Analysis (PCA) was conducted on the original bi-temporal multispectral images, respectively, as well as the derived texture images. Several principle components separately selected from spectral and texture
bands were combined for image classification, in order to improve the overall accuracy.

3.4. Multitemporal classification by SVMs

An appropriate classifier is required to effectively carry out the land cover classification using combined multitemporal spectral and texture information. Support Vector Machines (SVMs), a recently developed statistical learning method, were used as the classifier in this study. Compared with the traditional supervised classifiers, such as Maximum Likelihood classifier, the SVMs do not require features to obey Gaussian distribution, but also it is less sensitive to Hughes phenomenon. Consequently, the SVM approach can effectively handle high-dimensional data with a limited number of training samples, providing higher classification accuracy. In the present paper, the SVMs based on a ‘one-against-one’ strategy are employed to realize multi-class classification for multitemporal images and multitemporal texture features.

3.5. Accuracy assessment

Accuracy assessment is very important in remote sensing classification. Compared the classification generated from the image with the reference data obtained from ground observations for each test sample, we can set out the confusion matrix. As a starting point, further analysis techniques of overall classification accuracy, Kappa coefficient, as well as producer’s and user’s accuracies would be possible.

The Kappa analysis is a discrete multivariate technique used in accuracy assessment for statistically determining if the selected classification approach is significantly better than a random classification, and it is a more objective method than overall accuracy. Kappa coefficient can range from 0 to 1, and a higher value means a better classification result obtained. Lands and Koch [12] pointed out that a value greater than 0.80 represents strong agreement between the remotely-sensed classification and the reference data.

In order to statistically compare the difference between two classification results, Z-test is conducted to determine if one is significantly better than another. The Z-test statistic is expressed by:

\[ Z = \frac{\hat{K}_1 - \hat{K}_2}{\sqrt{\text{var}(\hat{K}_1) + \text{var}(\hat{K}_2)}} \]  

where \( \hat{K}_1, \hat{K}_2 \) denote the estimates of the Kappa coefficient for confusion matrices set out from classification results using pure spectral information and plus multitemporal texture, respectively, and \( \text{var}(\hat{K}_1), \text{var}(\hat{K}_2) \) denote the corresponding estimates of the sample variances [13]. At the 95% confidence level, the critical value is 1.96, which means if the value of Z statistic is greater than 1.96, we can conclude that the difference is significant.

4. EXPERIMENTS AND DISCUSSIONS

Fig.1 shows an original image (left) and the corrected image using the method proposed in the present study (right). From the figure, most of the noises have been removed.

A total of 74 spectral bands from the bi-temporal images were used to extract texture features by PCV. Window size and lag distance are two important parameters to be considered. In order to choose an appropriate window size for texture computation, several window sizes, 3x3, 5x5 and 7x7 pixels were experimentally tested, and we finally chose the size of 5x5 pixels. In this case, only three lags less than half of the window size, i.e. h=0, 1 and 2 are available. It was found that the overall accuracy is higher with a lag distance of 0, revealing an apparent temporal correlation between two images. Thus, h=0 was chosen to extract texture information within a 5x5 moving window. Another factor that should be taken into account is the direction of lag distance. Since anisotropy is not noticeable for the images, the omnidirectional variogram was used in this study.

After feature extraction by PCA, three spectral PCs of each CHRIS/PROBA image, representing 99.4% and 99.0% of original information respectively, as well as four textural PCs (99.6%) were selected for this study. The derived texture components were combined with the spectral components as additional bands for image classification. The result was compared to the classification using spectral information from the selected PCs of bi-temporal images alone, which was considered as the benchmark.

The classification results by SVMs are summarized in Table 1. We can see that the overall accuracy using spectral information alone is 85.66%, and Kappa coefficient is 0.7856. This result is considerably satisfying, which means CHRIS/PROBA data is quite effective in multitemporal classification. On the other hand, the joint use of spectral and multitemporal texture information leads to a significant increase in overall accuracy and Kappa coefficient, with a improvement of over 3% and 5%, respectively. As mentioned above, this represents a stronger agreement with the ground distribution of land cover types, for the reason that Kappa coefficient is greater than 0.80.

Further analysis shows that value of Z-test statistic is 2.206, greater than 1.96, which means the difference between two classification results is significant at the 95% confidence level. Therefore, adding texture information extracted by the proposed method is helpful to improvement of multitemporal classification accuracy.

<table>
<thead>
<tr>
<th>Table 1 Classification Results by SVMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral Classification</td>
</tr>
<tr>
<td>PA</td>
</tr>
<tr>
<td>Wheat</td>
</tr>
<tr>
<td>Cotton</td>
</tr>
<tr>
<td>Orchard</td>
</tr>
<tr>
<td>Water</td>
</tr>
<tr>
<td>Residence</td>
</tr>
<tr>
<td>Bare land</td>
</tr>
<tr>
<td>OA</td>
</tr>
<tr>
<td>Kappa</td>
</tr>
</tbody>
</table>

PA: Producer’s Accuracy (%); UA: User’s Accuracy (%); OA: Overall Accuracy (%)

By comparing the classification results using spectral information alone and incorporated with multitemporal texture information (Table 1), it is obvious that both producer’s accuracies and user’s accuracies of most land cover types have been increased to a varying extent. Some are considerably notable, for example, producer’s accuracy of wheat and user’s accuracy of bare land is improved by 14.75% and 27.76%, respectively, indicating less...
Multitemporal CHRIS/PROBA images, incorporated with multitemporal texture features were used for classification. The experimental results demonstrated that CHRIS/PROBA images are applicable in multitemporal classification. The study also showed that adding PCV texture as additional bands in multitemporal classification leads to a significantly higher accuracy. The proposed method can be used in relevant applications. Further study will be focused on the evaluation of the performance of multitemporal texture in multitemporal classification using more datasets.

6. REFERENCES


