High Dimensional Feature for Hyperspectral Image Classification

WANG Cailing1,*, WANG Hongwei2, Yinyong Zhang3, WEN Jia4, YANG Fan1
1School of computer science, Xi’an Shiyou University, Xi’an China
2engineering University of CAPF, Xi’an, China
3Department of electronic and electrical Engineering, University of Strathclyde, UK
4School of electronics Engineering, Tianjin Polytechnic University, Tianjin, China

Abstract. Making a high dimensional (e.g., 100k-dim) feature for hyperspectral image classification seems not a good idea because it will bring difficulties on consequent training, computation, and storage. In this paper, we study the performance of a high-dimensional feature by texture feature. The texture feature based on multi-local binary pattern descriptor, can achieve significant improvements over both its tradition version and the one we proposed in our previous work. We also make the high-dimensional feature practical, we employ the PCA method for dimension reduction and support vector machine for hyperspectral image classification. The two real hyperspectral image datasets are employed. Our experimental results with real hyperspectral images indicate that the high dimensional feature can enhance the classification accuracy than some low dimensional.

1 Introduction

Hyperspectral image (HSI) captures reflectance values from Visible to Infrared spectrum which cover the wide spectral range with hundreds of bands for each pixel in the image. This rich spectral information provides possibility to distinguish different materials spectrally. With the spatial resolution improves, many researches focus on spatial feature extraction for spectral-spatial hyperspectral image classification. For instance, morphological profiles [MPs][i] and attribute profiles [APs][ii] have been successfully employed to model structural information in hyperspectral image processing. Local binary pattern operator, which extract texture feature in spatial domain, is also used in HSI for classification[iii]. Meanwhile, multi-features[iv] and multi-hypothesis[v] pre-processing are present in literature.

A good low-level feature should be efficient to classification. However, there are several spatial information for one pixel such as shapes, texture and others. Different local size can also provide the different features. The high-dimensional spatial feature is the stack of different spatial features or same feature with different local size. Of course, high-dimensional feature leads to high cost while displays more information for classification. Even if we use a linear dimension reduction method like Principal component Analysis (PCA).

In this paper, we study the performance of a high-dimensional feature. We first extend texture feature based on multi-local binary pattern descriptor in spectrum domain. We also make the high-dimensional feature practical, we employ the PCA method for dimension reduction and support vector machine for hyperspectral image classification. Our experimental, conducted with a two real hyperspectral scenes indicate that the high dimensional feature can enhance the classification accuracy than some low dimensional.

2 Proposed approach

2.1 Texture feature extraction in spectrum domain

Local binary pattern (LBP) [vi] takes a local neighbourhood around each central pixel which thresholds the pixels of the neighbourhood at the value of central pixel, depending on whether the centre pixel has a larger intensity value or not. The neighbouring pixels generate a binary code for summarizing local grey-level structure. The basic LBP method considers a small circularly symmetric neighbourhood that has \( P \), along with selected neighbours \( \{ t_i \}_{i=0}^{P-1} \) and central pixel \( t_c \), the LBP is computed by Eq.(1).

\[
LBP_{P,R(i)} = \sum_{i=0}^{P-1} s(t_i - t_c)2^i
\]

\[
s(x) = \begin{cases} 
1, & x \geq 0 \\
0, & x < 0 
\end{cases}
\]

(1)

Where \( P, R \) denote the number of sampling points and radius of a circle based on a circularly symmetric
neighbour set of central pixel \( t \) and neighbouring pixels \( t_{i=1}^{P} \), and \( s(x) \) takes 1 if \( x \geq 0 \) and 0 otherwise.

We calculate LBP method in spectrum domain, which means that we compare the local neighbourhood with each central spectrum depending on whether their information divergence is larger than thresholds we give. We use Spectral information Spectral information divergence (SID)[8] to reveal information divergence.

Assuming that hyperspectral pixel \( X = (x_1, x_2, \cdots, x_L) \), each component \( x_i \) is the pixel of band \( B \) acquired at a particular wavelength \( \lambda_i \). \( \{\lambda_i\}_{i=1}^{L} \) is a set of \( L \) wavelengths. Suppose the \( X \) as a random variable by defining an appropriate probability space \( (\Omega, \Sigma, P) \) associated with it where \( \Omega \) is a sample space, \( \Sigma \) is an event space, and \( P \) is a probability measure. \( P = (p_1, p_2, \cdots, p_L) \) for \( X \) by Eq.(2).

\[
P(\{\lambda_i\}) = p_i = \frac{x_i}{\sum_i x_i}
\]  

(2)

Denoted another hyperspectral pixel \( Y = (y_1, y_2, \cdots, y_L) \) with probability \( Q = (q_1, q_2, \cdots, q_L) \), vector The SID is calculated by Eq. (3).

\[
SID(X, Y) = D(X\|Y) + D(Y\|X)
\]  

(3)

\[
D(X\|Y) = \sum_{i=1}^{L} p_i \log\left(\frac{p_i}{q_i}\right)
\]

Where

Supposed that the central hyperspectral pixel \( X_c \) and the neighborhood hyperspectral \( \{Y_1, Y_2, \cdots, Y_N\} \), \( N \) is the number of neighborhood pixels the threshold for binary is chosen by calculating the Mean of SID, shown in Eq. (4).

\[
\delta_c = \frac{1}{N} \sum_{i=1}^{N} SID(X_c, Y_i)
\]  

(4)

Then the SID-LBP is shown in Eq. (5).

\[
LBP_{X_c, Y_i} = \sum_{n=0}^{P} s(SID(X_c, Y_i) - \delta_c)2^{n-1}
\]  

(5)

\[
s(x) = \begin{cases} 
1, & x \geq 0 \\
0, & x \leq 0 
\end{cases}
\]

The texture feature extraction employing LBP in spectrum domain is generated for the centre spectrum by calculating the histograms in its corresponding local LBP image patch.

### 2.2 Constructing high-dimensional feature

As we all know that the texture feature is quite different for different image patch size, so we extract the high-dimensional texture feature by calculating histograms with different image patch size. In this paper, we choose the patch size from 5 to 25 with step is 2. The high-dimensional texture feature proposed is shown in Eq. (6).

\[
HTF = \sum_{size=5}^{size=25} \text{hist}_{size}(LBP) \quad \text{step} = 2
\]  

(6)

### 2.3 Dimension reduction

Dimension reduction can be viewed as making the ideals to practical. There are many dimensional reduction methods that are available in the literature [8, 9]. PCA disregards the target data and exploits correlations between the input variables to maximize the variance of the projects. In this paper, we employ the PCA method for dimension reduction

### 2.4 Classification approach

There are many methods for classification, One of the most popular classification is support vector machine (SVM)[10], which is often employed for HSI image classification. The key idea behind a kernel version of SVM is to map the data form its original input space into a high-dimensional kernel-induced feature space where classes may become more separable. SVM model is constructed on determining an optimal hyper-plane in the kernel- induced space by solving Eq. (7).

\[
\min_{\omega,\xi} \left\{ \frac{1}{2} \|\omega\|^2 + \xi \sum_{i=1}^{n} \xi_i \right\}
\]  

(7)

Subject to the constraints:

\[
y_i \left(\langle \omega, \phi(X) \rangle + p \right) \geq 1 - \xi_i
\]  

(8)

For \( \xi_i \geq 0 \) and \( i = 1, \cdots, n \), where \( \omega \) is normal to the optimal decision hyper-plane (i.e. \( \langle \omega, \phi(X) \rangle + p = 0 \)), \( n \) denotes the number of samples, \( P \) is the bias term, \( \xi \) is the regularization parameter which controls the generalization capacity of SVM, and \( \xi \) is the positive slack variable allowing us to accommodate permitted errors appropriately.

### 2.5 Proposed classification approach

In this paper, we propose a high dimensional feature based on multi-local binary pattern descriptor from spectrum domain by image patch size step is 2. In order to make the high-dimensional feature practical, the PCA method is employed for dimension reduction. SVM classifier is also requested to use here. The procedure is shown in Table 1.

**Table 1.** The procedure of proposed method

| Input: Training set with N samples for each class, initial texture patch size sets W1 and Window size step: | Output: Classification maps and accuracy |
|---|---|
| 1. Calculate the texture feature histograms by SID-LBP with different window size. | 2. Dimension reduction by PCA |


3 Experimental results

In this section, we evaluate the proposed approach using two real hyperspectral datasets. These datasets include different contexts, different spatial resolutions and different bands in order to assess the performance of proposed approach.

The first data was collected by Airborne Visible/Infrared Image Spectrometer (AVIRIS) over Northwest Indiana, Indiana, USA, in June 1992. The second data was collected by the Reflective Optics System Imaging Spectrometer sensor covering the University of Pavia, Italy. Two dataset are shown in Fig.1.

![Figure 1](image1.png)

(a) The 10th band of Indian Pine image; (b) Indian Pine ground survey; (c) The 10th band of the University of Pavia; (d) the University of Pavia ground survey.

To investigate the effects of image patch size, the image patch size is set from 5 to 25 with step is 2. The training samples are selected as 5% of total samples. The classification accuracy is assessed by overall accuracy (OA) and Kappa statistic. The experiment results are shown in Table 2 and Table 3.

**Table 2.** The classification accuracy obtained by the proposed method with training samples ratio is 5% in Indian Pines dataset and the University of Pavia dataset.

| Dataset name         | OA    | Kap.  |
|----------------------|-------|-------|
| Indiana Pines        | 0.9799| 0.9607|
| University of Pavia  | 0.9757| 0.9517|

From the table 2, we can get important information that the image patch size is one of main factors for image classification accuracy. Therefore, we investigate the high dimensional feature is very necessary.

From the table 3. We can conclude that combined the texture feature from different image patch size together could improve for 2%~20% compared to single patch size texture feature.

4 Conclusion and future research lines

In this paper, we have investigate that the high dimensional feature is good for HSI classification. We first make high dimension feature by LBP texture feature extraction method expanded in spectrum domain and reduced dimension by PCA method. The SVM classifier is employed for classification. After experiments, we get information that the image patch size is one of more important factors for classification. Therefore, we combine the texture features with different image patch size together as a high dimensional feature for classification. We can conclude that combining the texture feature from different image patch size together could improve for 1%~3% compared to single patch size texture feature.

Although the results obtained are very encouraging, further experiments with additional scenes and some others feature extraction methods should be conducted. We also need to choose some other classifiers for experiments.

Acknowledgments

This work was supported in part by National Natural Science foundations of China (Grant Nos. 41301382, 61401439, 41604113, and 41711530128) and foundation...
of Key lab of spectral imaging, Xi’an Institute of Optics and Precision Mechanics of CAS.

References

1. Mura, Mauro Dalla, et al. INT J REMOTE SENS., 31, 22(2010)
2. J A Benediktsson. INT J REMOTE SENS., 31, 22(2010).
3. Li W, Chen C, Su H, et al. IEEE Trans. on Geo. & REMOTE SENS., 53, 7(2015)
4. Huang X, Zhang L. IEEE Trans. on Geo. & REMOTE SENS, 51, 1(2013)
5. Chen C, Li W, Tramel E W, et al. IEEE J -STARS, 7, 4(2014)
6. Ojala T, Pietikäinen M, Mäenpää T. IEEE T PATTERN ANAL, 24, 7, 2002.
7. C I Chang. IEEE T INFORM THEORY, 46, 6(2000)
8. J A Benediktsson, J A Palmason, J R Sveinsson. IEEE Trans. on Geo. & REMOTE SENS., 43, 3 (2005).
9. J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, and S. Marshall, et al. ISPRS J PHOTOGRAMM, 93, 7(2014).
10. V. Tatyana, L. Bruzzone, and G. Camps-Valls. IEEE Trans. on Geo. & REMOTE SENS., 47.3(2009)