Hyperspectral image classification based on composite kernel relevance vector machine

Cheng Sun 1*, Donghao Liu 1, JieGuo Han 2, Bei Yang 1 and Zhaoxiang Cheng 1

1 China Ordnance Industry Survey and Geotechnical Institute CO., LTD, Beijing 100053, China
2 Gansu Yinguang Chemical Industry Group Co., LTD, Baiyin 730000, China

*Corresponding author. Email: sunz0412@163.com

Abstract. This paper presents a composite kernel Relevance Vector Machine (RVM) algorithm, for enhanced classification accuracy of hyperspectral images. This paper constructs three forms of composite kernels based on properties of kernels. The spatial feature is extracted using multi-scale morphological method from the image after principal components transform. The final classification is achieved by composite kernel RVM classifier. The proposed approach is tested in experiments on AVIRIS data. Compared with spectral kernel RVM, the OA and Kappa coefficient of composite kernel RVM increased obviously. However, the training time dose not increased. Meanwhile, composite kernel RVM has ability to get high accuracy with relative small training set. The proposed method has practical use in hyperspectral imagery classification.

Keywords: image classification, relevance vector machine, composite kernel, multi-scale morphological profiles, hyperspectral imagery.

1. Introduction

Hyperspectral remote sensing data provide both detailed structural and spectral information. Therefore, the data should be useful for information and classification. However, the classification of hyperspectral data is a challenging problem[1]. First, the hyperspectral data contain a lot of information about the spectral properties in the pixel, but no spatial information is inherent in the spectral data. Therefore, many spectral/spatial classifier have been recently proposed for classification of hyperspectral data in the literature[2]. In this paper, we propose a composite kernel Relevance Vector Machine (RVM) classification method that is based on making use of both the spectral and spatial information for classification.

2. Algorithm

2.1. Relevance vector machine (RVM)

According to Relevance vector machine[3], We base our predictions upon some function \( y(x) \) defined over the input space:

\[
t_n = y(x_n; w) + \epsilon_n
\]  

(1)
where the output is a linearly-weighted sum of $M$, generally nonlinear and fixed, basis functions $K(x) = (K(x,x_1), K(x,x_2), K(x,x_3), \ldots, K(x,x_N))^T$, the adjustable parameters $w = (w_0, w_1, \ldots, w_n)^T$.

$$y(x;w) = \sum_{i=0}^n [w_iK(x,x_i) + w_0]$$

(2)

The function $y(x)$ is as defined in (2), which defined as a Gaussian distribution over $t_n$ with mean $y(x_n)$ and variance.

The likelihood of the complete data set with independence of $t_n$ can be written as:

$$p(t,w, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2}||k - \phi w||^2\right)$$

(3)

where, $\alpha = [\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_N]^T$, $\theta_n \in \{0,1\}$, $\phi(x_n) = [1, K(x_n, x_1), K(x_n, x_2), \ldots, K(x_n, x_N)]^T$, $w$ is a convolution of Gaussians:

$$p(w | \alpha) = \prod_{j=0}^N N(\theta_j | 0, \alpha_j^{-1})$$

(4)

2.2. Composite Kernel RVM

In this paper, we present three different kernel approaches for the joint consideration of spectral and textural information in a unified framework for hyperspectral image classification. In order to take advantage of spectral and textural character[4].

This paper presented a method of composite kernels for the combination of spectral and contextual information in this section.

First, we made pixel definition, A pixel is redefined both in the spectral domain using its spectral content $x_i^s$, and in the spatial domain by applying some feature extraction to its surrounding area $x_i^w$, which computated spatial features.

Then, we performed Kernel computation. According to Mercer theory kernel function that fulfills Mercer’s conditions can be used to constructed kernel matrices[5-6]. The following functions are valid kernels:

$$k(x,x') = k_1(x, x') + k_2(x, x')$$

(5)

$$k(x, x') = ak_1(x, x')$$

(6)

$$k(x, x') = k_1(x, x') \cdot k_2(x, x')$$

(7)

A. Summation kernel

A simple composite kernel combining spectral and textural information naturally comes from the concatenation of non-linear transformations $\phi_1(.)$ and $\phi_2(.)$ and into Hilbert spaces $H_1$ and $H_2$, respectively. Then, the following transformation was constructed:

$$K_{\text{sum}}(x_i, x_j) = \{\phi_1(x_i^s), \phi_2(x_i^w)\}$$

(8)

$$= \langle \{\phi_1(x_i^s), \phi_2(x_i^w)\}, \{\phi_1(x_j^s), \phi_2(x_j^w)\} \rangle$$

$$= K_s(x_i^s, x_j^s) + K_w(x_i^w, x_j^w)$$

The transformations is expressed as the sum of positive definite matrices accounting for the textural and spectral counterparts, independently

B. Weighted Summation Composite Kernel

A composite kernel that balances the spatial and spectral content can also be created. The corresponding weighted summation composite kernel can be computed as follows:
where, \( \mu \) is a positive real-valued free parameter.

C. Dot Product Composite kernel

The corresponding dot product composite kernel can be computed as follows:

\[
K_{\text{product}}(x_i, x_j) = \{\phi_1(x'_i), \phi_2(x''_i)\} \\
= \mu K_s(x'_i, x'_j) + (1 - \mu) K_w(x''_i, x''_j)
\]

(9)

This paper use Radial basis function to calculate the spectral kernel \( K_s \) and the spatial kernel \( K_w \). For example, the spectral kernel \( K_s \) can be computed as follows:

\[
K_s = \exp(-\gamma\|x'_i - x'_j\|^2)
\]

(10)

Where, \( x'_i \) is spectral feature of the image.

Any kernel function, which fullfils Mercer’s condition can be used in RVM. In this paper, we used the kernel function that fullfils Mercer’s condition.

2.3. Multiscale mathematical morphology

Mathematical morphology theory[7] is widely used in non-linear image analysis and pattern recognition problem. In this paper, Morphological Profiles (MPs) are defined using the granulometry. An MP is composed of the opening profile (OP) and the closing profile (CP). The OP at the pixel \( x \) of the image \( I \) is defined as an \( n \)-dimensional vector.

\[
\Delta(x) = \left\{ \Delta_c = \Delta \phi_{\alpha_c = c}, \forall c \in [1, n] \right\} \\
\Delta(x) = \left\{ \Delta_c = \Delta \gamma_{\beta_c = c}, \forall c \in [n + 1, 2n] \right\}
\]

(12)

In which, \( n \) is the number of the computation, \( c = 1, 2, ..., 2n \).

In order to apply the morphological approach to hyperspectral data, principal components of the hyperspectral imagery are computed. The most significant principal components are used as base images for an extended morphological profile. Following the previous notation, the EMP is an \( m(2n + 1) \)-dimensional vector. Where, \( m \) is the number of retaining principal components.

3. Experiment and results

3.1. Data

Experiments were carried out using the familiar AVIRIS image taken over northwest Indiana’s Indian Pine test site in June 1992. We used a part of the 145×145 scene, which contains 220 band, and 9 classic labeled classes (the background pixels were not considered for classification purposes). We removed several noisy bands covering the region of water absorption, and finally worked with 169 spectral bands.

3.2. Experiment process

We performed principal component analysis (PCA) transform to this image. Then, we chose the first three principal components to perform multi-scale mathematical morphology and the square shapes structure is used.

Image classification based on Composite kernel Relevance Vector Machine method is performed. We used half of the sample to perform classification and the other half to perform test. The classification results are shown as Fig.3. In which, \( K_s \) and \( K_w \) is spectral kernel and spatial kernel, \( K_{\text{sum}} \) is the summation kernel, \( K_{\text{product}} \) is the product kernel, and \( K_{\text{weight}} \) is the weighted summation kernel.

\[
K_{\text{weight}}(x_i, x_j) = \{\phi_1(x'_i), \phi_2(x''_i)\} \\
= \mu K_s(x'_i, x'_j) + (1 - \mu) K_w(x''_i, x''_j)
\]

(9)
4. Conclusions
This paper presents a composite kernel Relevance Vector Machine(RVM) algorithm, for enhanced classification accuracy of hyperspectral images. This paper constructs three forms of composite kernels based on properties of kernels. The spatial feature is extracted using multi-scale morphological method from the image after principal components transform. The final classification is achieved by composite kernel RVM classifier. The proposed approach is tested in experiments on AVIRIS data. The Overall Accuracy(OA), Kappa Coefficient(KC), training time were used as the evaluation criteria in order to test the performance of the algorithm. According to the experiment, the Overall Accuracy(OA), Kappa Coefficient(KC) of composite kernel Relevance Vector Machine(RVM) algorithm is higher than the spectral RVM algorithm. And training time of composite kernel RVM algorithm is shorter than the spectral RVM algorithm. Overall, the composite kernel RVM algorithm can get higher accuracy with a more fewer training set. The proposed method has practical use in
hyperspectral imagery classification. This approach is successful in the task of integrating spatial and spectral information in the analysis process.

Reference

[1] F. Melgani, L. Bruzzone, Classification of hyperspectral remote sensing images with support vector machines[J]. IEEE Trans Geosci Remote Sens, 8,42,(2004)

[2] G M Foody, RVM-based multi-class classification of remotely sensed data[J]. Int J Remote Sens, 6,29,(2008)

[3] M E Tipping, A. Faul, Fast marginal likelihood maximisation for sparse Bayesian models [C] // Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics. Key West,Florida:AI & Statistics,(2003).

[4] Jimenez, L.O., Rivera-Medina, J.L., Rodriguez-Diaz, E. et al.. Integration of spatial and spectral information by means of unsupervised extraction and classification for homogenous objects applied to multispectral and hyperspectral data[J]. IEEE Trans Geosci Remote Sens, 4,43,(2005)

[5] Benediktsson, J.A., Palmason J.A., and Sveinsson J.R.. Classification of hyperspectral data from urban areas based on extended morphological profiles[J]. IEEE Trans Geosci Remote Sens,3,43,(2005)

[6] Camps-Valls G, Gomez-Chova L, Munoz-Mari J et al.. Composite Kernels for Hyperspectral Image Classification[J]. IEEE Geosci. Remote. Sens. Lett., 1,3,(2006)

[7] Serra, J., Image Analysis and Mathematical Morphology. Academic Press, London, 1982.