

A NEW FEATURE SELECTION METHOD FOR EO-1/HYPERION IMAGE CLASSIFICATION—A CASE STUDY OF SUBEI REGION, CHINA

Xianbin LI^{1,2}, Xiaoguang JIANG¹, Xiaohuan XI¹, Liang LIU³, Lingli TANG⁴

1 Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, 100080, China

2 Graduate University of Chinese Academy of Sciences, Beijing, 100049, China

3 National Disaster Reduction Center of China, Beijing, 100053, China

4 China Remote Sensing Satellite Ground Station, CAS, Beijing, 100086, China

Information about presenter

Name: Xianbin LI

Title: Mr

Affiliation: Academy of Opto-Electronics, Chinese Academy of Sciences

Postal Address: Room 307, No.95 Zhongguancun East Road, Haidian District, Beijing, China

Tel: 86-10-62641267

Fax: 86-10-62643022

E-mail: lixianbin@aoe.ac.cn

KEY WORDS: Feature Selection, Spectrum Reconstruction, Basis Function, Spectral Interval

ABSTRACT

Developments in detector technology and microelectronics present new opportunities for remote sensing. For example, the spaceborne Hyperion, carried by EO-1 satellite, acquires image data in 220 spectral bands over the spectral range 0.4-2.5 μ m at approximately 0.01 μ m spectral resolution with 30m spatial resolution. Those hyperspectral data can provide abundant information, and make it possible to detect diagnostical spectral characteristics, such as red-edge drift, etc. However, there is relatively high correlation between different bands and much redundancy in hyperspectral data sets. Therefore, one of the most important procedures is to select optimal bands for extracting information from hyperspectral data effectively.

In this paper, we first apply several important pre-processing procedures to Hyperion L1R data, such as radiometric calibration, destriping, smile correction, and atmospheric correction, etc. Then we apply spectrum reconstruction approach, which use several basis functions and corresponding spectral intervals to describe the spectrum extracted from Hyperion hyperspectral data sets in Subei region, China. The feature selection method based on spectrum reconstruction is incrementally adding bands to the initial bands, followed by adjustment of band widths and locations. At last, we can aggregate several Hyperion bands into a new simulated band in each interval and apply MLC image classification methods to it. The overall accuracy and Kappa coefficient can be as high as 92% and 89% separately compared with in situ measurement, which support the validity of this feature selection method.

1 Introduction

The advent of hyperspectral remote sensing, which acquires reflectance values at tens of hundreds of different wavelengths, indicates a new era of remote sensing technology. Hyperspectral Images have unique advantages in many applications such as in resource management, agriculture, mineral exploration, and environmental monitoring,

etc, and its application extension is expected wider and wider while application depth will be deeper and deeper. However, there is relatively high correlation between different bands and much redundancy in hyperspectral data sets, which could bring us many problems for hyperspectral image analysis. For example, driven by classification or discrimination accuracy, one would expect that, as the number of hyperspectral bands increases, the accuracy of classification should also increase. Nonetheless, this is not the case in a model-based analysis. Redundancy in data can cause convergence instability of models. Furthermore, variations due to noise in redundant data propagate through a classification or discrimination model. The same is true of spectral information that has no relation to the feature to be discriminated in the underlying mathematical model. Such information is the same as noise to any statistical model, even if it is unique and accurate. Thus, processing a large number of hyperspectral bands can result in higher classification inaccuracy than processing a subset of relevant bands without redundancy. In addition, computational requirements for processing large hyperspectral data sets might be prohibitive, and for supervised classification methods the number of training samples requirement will increase greatly. Therefore, feature selection or dimension reduction is of great value for information extraction and also other applications.

Feature selection is defined as “the research for a subset of the original measurement features that provide an optimal trade-off between probability of error and cost of classification”(Swain, etc, 1978). Ben-Bassat (1982) grouped the evaluation into three categories: distance measures, information or uncertainty measures, and dependence measures. Distance measures, also known as separability or divergence, are most widely used for feature selection, such as the Bhattacharya distance, Euclidean distance and Jeffries-Matusita (J-M) distance. Spectrum reconstruction for feature selection used in this paper is actually a method of information content evaluation. J.C Price (1975) proposed this method and evaluated the spectral variability in the spectral domain 6.25-25um with nine wide spectral bands. He (1990) also reconstructed the spectrum of soils with a few selected channels. Z.L Li (1999) reconstructed soil and vegetation spectral emissivity using six channels in the 8-13um spectral region. In this paper, we, according to spectrum reconstruction theory, incrementally add bands to the initial bands, followed by adjustment of band widths and locations. As we all known, the spectral width and central wavelength determination is very important and difficult for special application. The feature selection method proposed in this paper can be very effective for sensor design, especially channel selection.

2 Preprocessing EO-1/Hyperion data to spectrum

The Hyperion sensor, carried by the National Aeronautics and Space Administration (NASA) Earth Observing 1 (EO-1) satellite, is the first spaceborne hyperspectral instrument to acquire both visible near-infrared (VNIR, 400–1000 nm) and shortwave infrared (SWIR, 900–2500 nm) spectra. Hyperion is a pushbroom sensor with two spectrometers and a single telescope. Hyperion product has three levels: L1A, L1B, L1R. This paper used L1R image in Subei region, whose processing steps include (Pamela Barry, 2001): smear correction, echo removal, background removal, calibration, abnormal pixels removal and image quality checking, etc. In this paper, we apply destriping, smile correction and atmospheric correction to gain basic data for spectrum reconstruction.

2.1 Abnormal pixels correction

Each distributed Hyperion L1R image contains a log file that reports the bad or corrupted pixels but they are not fixed. In addition, there are dark vertical stripes in the image that are not reported in the log file. One possible reason for the abnormal pixels could be that the calibration of the detectors in the detector array becomes unbalanced. The Hyperion system acquires data in pushbroom mode, in which there is a separate detector to gather data for each column. If the detectors are not calibrated properly striping artifacts could easily be generated. Therefore, we should deal with those artifacts before we can use it.

At first, we check the log file to find columns whose DN values are zero, and replace them with the average DN values of their immediate left and right neighboring pixels. A general approach to remove vertical stripes is similar to methods used in the past to balance horizontal stripes in mirror scanner images by histogram equalization (Berthold K. P, 1979), or to balance detectors in airborne pushbroom sensors. The pixel balancing introduced here is globally. The statistical moments of each column are modified to match those for the whole image for each band. The abnormal pixel correction algorithm works well for most bands, except in areas that are intensively striped and noisy (water absorption bands). The visual effect can be seen in Fig.1.

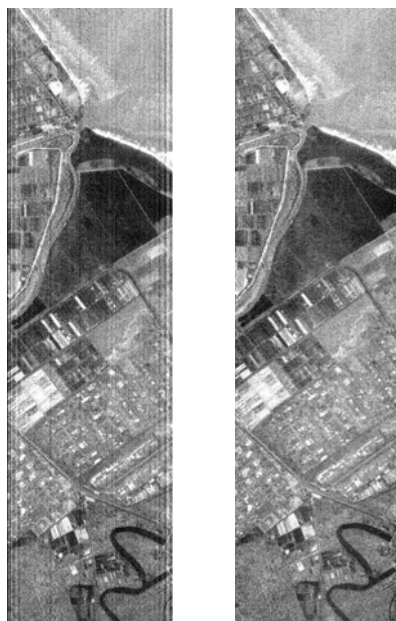


Fig.1 A visual effect of destriping



Fig.2. A visual effect of smile correction (MNF2)

2.2 Smile correction

Smile, which exists in all Hyperion datasets, refers to an across-track wavelength shift from center wavelength, which is due to the change of dispersion angle with field position. The shifts are dependent on pixel position in across-track direction. For VNIR bands, the shifts range between 2.6–3.5nm while for SWIR bands, the shifts are less than 1 nm and are not significant for most applications. Considering the high spectral resolution of the Hyperion data, the 2.6–3.6-nm shifts of VNIR bands cannot be ignored, for they alter the pixel spectra and reduce classification accuracy.

The effect of “smile” is not obvious in the individual band. But smile becomes observable when the image is transformed into Minimum Noise Fraction (MNF) space. For Hyperion images with significant smile, there is a brightness gradient in the first or second eigenvalue image. There is no brightness gradient in MNF for images without significant smile.

There are three typical smile correction approaches (David G. etc, 2003): Moving Linear Fitting and Interpolation, Column Mean Adjusted in Radiance Space, Column Mean Adjusted in MNF Space. Goodenough (2003) proved that three smile correction methods maintain the basic spectral absorption features. However, only the linear interpolated spectra closely match the original spectra. The column means adjusting methods create significant departures from these original spectra. The moving linear fitting and interpolation retains spectral fidelity and was therefore selected for smile correction in this paper. The result can be seen in Fig.2.

2.3 Atmospheric correction

The apparent reflectance should be derived for spectral analysis because the data released from USGS is only radiometrically corrected. This process is termed atmospheric correction. There are currently several packages commercially available to perform this function, such as ACORN (ImSpec, 2004), ATCOR (Leica, 2004), and FLAASH (Spectral Sciences, Inc.). This paper uses FLAASH in ENVI (RSI, 2001). This software is a MODTRAN based package with several defined atmosphere models and aerosol models and a simplified user interface. Compared with ACORN, FLAASH is less sophisticated and has fewer input requirements. We set the parameters according Users' guide (Richard Beck, etc, 2003), and the result can be seen in Fig.3.

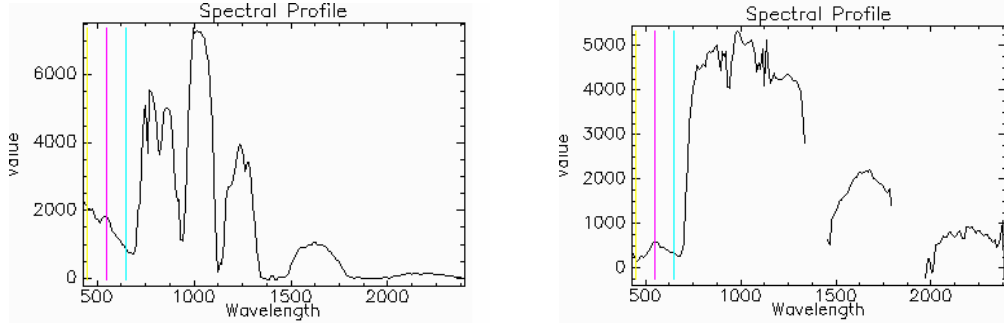


Fig.3.The radiance (left) and spectral reflectance (right) profile of vegetation

3 Feature selection based on spectrum reconstruction

3.1 Spectrum reconstruction theory

Since spectra generally vary in a smooth continuous fashion, a measurement in one spectral band is correlated with and thus provides information about reflectance values over a wider spectral range. Considering that all spectral bands are not independent, we can select a set of uncorrelated spectral intervals from which all other can be derived. Such a method has been developed by Price (1975) and applied to spectra from Iris and to laboratory and field reflectance spectra in VNIR (1990). Let $X^\alpha(\lambda) = (x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)$ represent a measured spectrum for the set of n wavelength values $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$, with superscript denoting the individual sample. We shall describe spectrum by a set of spectral basis functions (Price, 1997):

$$X^\alpha(\lambda) \approx \sum_{i=1}^M S_i^\alpha \varphi_i(\lambda)$$

Where the basis functions $\varphi(\lambda_1, \lambda_2, \dots, \lambda_n)$ are spectral shapes, as defined from statistical analysis of a collection of spectra, and the coefficients S_i^α are wavelength integrals relating to the original measurements. Each basis function has an associated spectral interval $[\lambda_{i(\min)}, \lambda_{i(\max)}]$, which represents the domain of integration in specifying the coefficients S_i^α . Each φ has a mean value of 1.0 in its spectral interval and then decreases according to the degree of wavelength correlation in the ensemble of measured spectra. The expansion represents successive approximations to the original spectra. M is the number of basis functions required to describe $X^\alpha(\lambda)$ to within very small residuals.

3.2 Feature selection method

It has been proved that spectrum reconstruction is self-contained and can be used in feature selection. We first discuss the spectrum reconstruction of vegetation spectrum. Consequently, ten basis functions can represent 97.8% of spectrum variance. When additional spectral data sets become available, and we should incorporate these spectra into an existing description, with a consequent increase in the number of basis functions and corresponding spectral

intervals. The approach for adding additional spectra to the description is straightforward (Price, 1994). First, describe the new spectra in terms of the existing basis functions; compute each residual δx^α and the mean square residual $(\delta x^2)^{1/2}$ for the new collection; and save for reference some, say 10, of the largest residuals. Second, following these computations, examine the mean square residual as a function of wavelength, and the individual 10 worst-case residuals, to identify spectral regions that have large residuals. In these regions additional basis functions are needed to describe this remaining variability. If some spectra are not well described by the existing basis functions, then there are two cases: (a) The residual is large in a spectral region that has not yet been identified; i.e. it lies between two previously selected wavelength intervals. This interval should now be added as an additional spectral channel. (b) The residual is large at the border between existing spectral intervals, in which case the existing wavelength bounds must be adjusted, and the new interval must be inserted between them; i.e., two spectral intervals become three. After selection of the new additional wavelength ranges, we obtain the revised basis functions for the entire spectral collection. The resultant spectral intervals are given in Table 1.

Table 1 The result of feature selection based on spectrum reconstruction

Hyperion data sets	Spectral intervals(um)									
Vegetation	0.46-0.54	0.55-0.59	0.61-0.69	0.73-0.81	0.83-0.93	0.99-1.09	1.12-1.16	1.52-1.61	1.97-1.99	2.29-2.39
Vegetation	0.46-0.51	0.53-0.58	0.59-0.65	0.66-0.69	0.70-0.74	0.75-0.81	0.83-0.91	0.93-0.99	1.00-1.09	1.12-1.16
water	1.35-1.41	1.52-1.61	1.73-1.80	1.97-1.99	2.17-2.23	2.29-2.39				
Vegetation, Water,	0.46-0.51	0.53-0.58	0.59-0.65	0.66-0.69	0.70-0.74	0.75-0.81	0.83-0.91	0.93-0.99	1.00-1.09	1.12-1.16
and Bare land	1.20-1.31	1.35-1.41	1.44-1.49	1.52-1.61	1.62-1.71	1.73-1.80	1.82-1.88	1.97-1.99	2.17-2.23	2.29-2.39

4 Image aggregation and classification

In last selection, we introduced the basic theory and result of feature selection. What we get from feature selection is several relatively wider spectral intervals, so we can aggregate these hyperspectral images in each interval into a simulated image.

The general process of aggregation is to calculate a weighted sum of the Hyperion bands that covered each spectral interval. The weights used in the sum were derived by comparing the spectral response of the hyperspectral bands with the broad band. Mathematically, the process involved convolving the Hyperion Gaussian spectral response with the broad band spectral response and normalizing carefully to ensure consistent units between instruments. The method can be represented by (P.S. Barry, etc.):

$$W_{zb} = \sum_{k=z-15}^{z+15} G_{zk} \times F_{kb} \quad P_{zb} = \frac{W_{zb}}{fwhm * \sum_z W_{zb}} \quad HB_b = \sum_z P_{zb} \times H_z$$

Where z is Hyperion spectral band, k is spectral range bounding the Hyperion band, b is spectral interval, F is simulated image spectral response, G is Hyperion Gaussian spectral response, W is unnormalized spectral weight for each Hyperion pixel, P is normalized spectral weight for each Hyperion pixel, H is radiance value for each Hyperion pixel and HB is aggregated radiance for simulated pixel.

The simulated image is given in Fig.4, and we can see its texture is clear and contrast is relatively high. Then we apply Maximum Likelihood Classification to it, and the result is shown in Fig.5. The overall accuracy and Kappa coefficient can be as high as 91.625% and 89.003% compared with in situ measurement, which support the validity of this feature selection method.



Fig.8 Aggregated image using selected channels



Fig.9 Classification result of simulated image

5 Conclusion

Hyperspectral remote sensing represents the orientation of the development of remote sensing. It can help us to detect the diagnostical spectral characteristics which are 10 nm level and can't be detected in multispectral remote sensing. Furthermore, it can provide us more accurate spectral information to choice. Ideally, we expected the discrimination accuracy will increase with the number of input channels. Unfortunately, the relatively high correlation between different bands and much redundancy in hyperspectral data sets lead to instability of models and unbearable computation. Therefore, feature selection or dimension reduction is one of the most important steps in many applications.

Traditional feature selection approaches include distance measures (such as Bhattacharya distance, J-M distance, etc) and image transform (such as PCA, K-T transform etc). But the former can't utilize abundant information effectively and the latter's resultant images can't directly be related to physical significance. Feature selection based on spectrum reconstruction can utilize the spectral information sufficiently, and we can reconstruct spectrum with basis functions and spectral intervals. The selected channels are the optimal band for objects discrimination. Because the resultant spectral intervals are board channels, we can aggregated hyperspectral images into multispectral image, which and direct us to select channel width and central wavelength for sensor design. The result shows that for many applications 10nm level spectral resolution is unnecessary and the channels needn't cover the whole spectrum. Therefore, we can research the spectral resolution requirement for special application, and gain the best trade-off among spectral resolution, spatial resolution, signal to noise ratio and cost.

References

- [1] <http://eo1.gsfc.nasa.gov>
- [2] Bisun Datt, etc. *Preprocessing EO-1 Hyperion Hyperspectral Data to Support the Application of Agricultural Indexes*. *IEEE Trans. Geosci. Remote Sensing*, Vol. 41, No. 6, 2003(7)
- [3] David G. etc. *Processing Hyperion and ALI for Forest Classification*. *IEEE Trans. Geosci. Remote Sensing*, Vol. 41, No. 6, 2003(7)
- [4] John C. Price. *Spectral Band Selection for Visible-Near Infrared Remote Sensing: Spectral-Spatial Resolution Tradeoffs*.

IEEE Trans. Geosci. Remote Sensing, 1997(5).

[5] <http://www.geo.wvu.edu/~warner/research/feature/paper/Price.htm>

[6] Zhao-Liang Li, F.Becker, etc. Channel selection for soil spectrum reconstruction in 8-13um region. *Journal of Geophysical research*, Vol.104, No.D18, P22271-22285, 1999(9).

[7] John C. Price. Band selection for procedure for multispectral scanners. *Apr.Opti.* Vol.33, No15,1994(5)

[8] Pamela Barry. *EO-1/ Hyperion Science Data User's Guide*,2001(11)

[9] Berthold K. P, etc *Dest-ripping Landsat MSS images by histogram modification. Computer graphics and image processing* (1979)

[10] Richard Beck, etc. *EO-1 User Guide (Version 2.3)*. 2003(7)

[11]P.S. Barry, etc. *EO-1 Hyperion hyperspectral aggregation and comparison with EO-1 Advanced land imager and Landsat 7 ETM+*