Photonirvachak

J. Indian Soc. Remote Sens. (September 2008) 36:235–241

RESEARCH ARTICLE

# Fractal-based Dimensionality Reduction of Hyperspectral Images

**Jayanta Kumar Ghosh · Ankur Somvanshi**

Received: 8 July 2006 / Accepted: 1 April 2008

**Keywords** Fractal dimension · Dimensionality reduction · Hyperspectral images · Spectral reflectance · Entropy

**Abstract** The spectral reflectance of any pixel in a remote sensing image depends on the characteristics of the particular land cover (LC) present in the Instantaneous Field of View (IFOV) of the sensor. The fractal dimension of the spectral reflectance curve (SRC) of any pixel can thus be visualized as a representation of the characteristics of the LC. Based on this, a fractal-based method for reduction of the dimensionality of Hyperspectral (HS) images has been investigated. The fractal dimension (FD) of SRC has been calculated by adopting a method based on Hausdorff metric that reduces the dimensionality from N HS bands to a single feature incorporating the characteristics associated with each of the bands. Further, it has been established that FD values can be used as a feature to identify anomaly in water cover.

J. K. Ghosh[1] (✉) · A. Somvanshi[2]
[1]Geomatics Engineering Group,
Civil Engineering Department,
Indian Institute of Technology Roorkee,
Roorkee – 247 667, India
[2] Project Engineer, WIPRO Technology,
Bangalore – 560 100, India

e-mail: gjkumfce@iitr.ernet.in

## Introduction

Hyperspectral remote sensing has become a potent source of environment monitoring because of its major success in the detection and identification of rocks and minerals, terrestrial vegetation, water optical properties, building materials etc. The hyperspectral sensors have very fast signal readout systems each of narrow bandwidth that could be rapidly sampled and reset so as to function as a working spectrometer from space. This has opened up a truly powerful new approach, which already has proven to be a superior identifier of materials found on the Earth's face. These objects can be identified more accurately and precisely in comparison to broad band multi-spectral data (Jensen, 2003). In essence, hyperspectral imaging yields continuous spectral signatures rather than the band histogram plots that result from systems like the Thematic Mapper which "lump" varying wavelengths into single-value intervals.

On the other hand, the large data volume generated by hyperspectral sensors invokes a challenge for traditional processing techniques. Some of the major challenges lie in removing redundant information and in assuring the continued relevance of vital information to the application at hand. Conventional methods may not be applicable for analysis of HS data due to the large data volumes used to characterize hyperspectral cubes. Thus, the analysis of HS data usually requires a preprocessing step, namely, dimensionality reduction. This is often used as a tool to speedup processing of hyperspectral images. It consists in selecting or generating a reduced set of bands that preserve the essential information content relevant for the application in mind.

Due to constraints both at the sensor and on the ground, dimension reduction is a common preprocessing step performed on many HS imaging datasets (Farrell *et al.*, 2005). Several methods have been developed and used to address the dimensionality reduction problem associated with HS imaging. Kaewpijit *et al.* (2001) have used a hybrid approach for data reduction using Principal Component Analysis (PCA) technique. Their approach not only reduces the intrinsic dimension of hyperspectral data but also generates the principal components that correspond to a user's specified information content. In this method, one can switch from one method to another method according to the information content needed and the best overall execution times. Robila (2004) has used spectral metrics for data reduction via spectral screening of hyperspectral imagery. This method uses various spectral metrics to characterize the similarity and tends to use the resulting subset in further computations instead of the full data. They compared the technique with PCA and reported better efficiency and performance, when the subset size is small. Kaewpijit *et al.* (2003) have proposed an automated spectral data reduction method based on wavelet decomposition. This method has shown great potential because it preserves the distinctions among spectral signatures. It also preserves high-frequency and low-frequency features, thus preserving peaks and valleys found in typical spectra. Huang *et al.* (2001) have used spectral reflectance curve to detect any differences in the leaf reflectance of weed-free soybean versus soybean in the presence of purple nuts edge. They used the Brushlet transform on a given HS reflectance curve and then extracted energy feature vectors. Finally, Fisher's linear discriminant analysis has been used to reduce the dimensionality of the feature vector. Farrell *et al.* (2005) have investigated the impact of dimension reduction on adaptive detection of difficult targets. They have shown that in many cases PCA has a minimal impact towards detection of a target that is spectrally similar to the background. According to them in many cases the four covariance-based adaptive detectors are robust for detecting difficult targets in datasets with only a handful of principal components. Xiuping *et al.* (1999) have used an approach in which dimensionality reduction requires initial partitioning of the complete set of bands into several highly correlated subgroups and then each subgroup gets transformed separately. The sub-groups are used as a guide to carry out feature selection and at last selected features are transformed again to achieve data reduction. This scheme has a computational advantage when compared to the PCT method. Several other methods are formulated based on calculation of fractal dimension of waveform such as, variance method, spectral method and morphological covering method etc. (Schepers *et al.*, 1992). These methods are computationally intensive and cumbersome to implement. Thus, in summary, methods used for dimensionality reduction of HS images are complex in nature and need high computational resources. Moreover, loss of information also occurs in many cases.

The objective of this paper is to propose a computationally less intensive fractal-based method for reduction of dimensionality of HS images.

## Background Theory

It is well established that fractal dimension provides a means for describing and analyzing the geometry of linear figures (Mandelbrot, 1999). In the case of HS remote sensing images, the characteristics of the LC present in the IFOV of the sensor is defined by its spectral response. This is meaningfully represented by SRC. Thus, fractal dimension of SRC at each pixel location in HS images will provide the characteristics of the object present at that location. It further results in reduction in dimensionality of N HS bands to a single dimension. In this study, the fractal dimension (FD) of SRC (a waveform) has been obtained by making use of the Hausdorff metric (Sevcik, 1998).

### Fractal-based Dimensionality Reduction

The Hausdorff Dimension ($D_h$) of a set in a metric space is given by:

$$D_h = \lim_{\varepsilon \to 0} \frac{-\ln[N(\varepsilon)]}{\ln(\varepsilon)} \quad (1)$$

where, N (ε) is the number of open balls of radius ε needed to cover the set. In the case of a line of length L consisting of segments of length 2× ε each, there will be N (ε) [=L/ (2× ε)] segment in the line i.e. covered by N (ε) open balls of radius ε. Thus, Equation (1) may be rewritten as :

$$D_h = \lim_{\varepsilon \to 0}\left[\frac{-\ln(L) + \ln(2 * \varepsilon)}{\ln(\varepsilon n}\right] = \lim_{\varepsilon \to 0}\left[1 - \frac{\ln(L) - \ln(2)}{\ln(\varepsilon n}\right] = \lim_{\varepsilon \to 0}\left[1 - \frac{\ln(L)}{\ln(\varepsilon n}\right] \quad (2)$$

In this study, SRC of a pixel in a HS image, a planar curve in a space, is being considered as the metric space. Here, SRC gets scaled linearly into a normalized space, to map the original SRC into an equivalent metric space. The first scaling normalizes every point in the abscissa as:

$$x_i^* = \frac{x_i}{x_{max}} \quad (3)$$

where $x_i$ is the spectral band number, $i$ and $x_{max}$ is the maximum value of spectral band number.

The second scaling normalizes the ordinate as:

$$y_i^* = \frac{y_i - y_{min}}{y_{max} - y_{min}} \quad (4)$$

where $y_i$ is the spectral reflectance value at the pixel location in band, $i$ and $y_{min}$ and $y_{max}$ are respectively the minimum and maximum value of the spectral reflectance of the considered pixel location over the entire bands. These two linear scaling map the N values of the SRC into another space that belongs to a unit square. This unit square may be visualized as covered by a grid of N (bands) × N (spectral reflectance) cells. Each of them contains one point of the scaled SRC. Calculating L of the scaled SRC and taking ε=1/ (2×N') [where N' (number of segments) = N-1] Equation (2) becomes :

$$D_h = \lim_{N' \to \infty} \left[ 1 - \frac{\ln(L)}{\ln(1/(2\times N'))} \right]$$

(as ε = 1/ (2×N'), lim ε → 0 leads to N' → ∞)

$$= \lim_{N' \to \infty} \left[ 1 - \frac{\ln(L)}{\ln(1) - \ln(2\times N')} \right]$$

$$= \lim_{N' \to \infty} \left[ 1 + \frac{\ln(L)}{\ln(2\times N')} \right]$$

$$\therefore \text{Fractral feature}, D \approx 1 + \frac{\ln(L)}{\ln(2 * N')} \quad (5)$$

where, $L = \sum_{i=1}^{N} \text{dist}(i, i+1)$

The approximation to D expressed in Equation (5), improves as N'→∞.

## Data Used

In this study, data from AVIRIS (Airborne Visible-Infrared Imaging Spectrometer) hyper spectral sensor, collected by Jet Propulsion Laboratory USA, is being used. It covers the wavelength region from 0.4 to 2.5 micro-meters (μm) in 224 spectral channels at a nominal spectral resolution of 10 nano-meters (nm).

The study area is in US known as Moffit field. It lies between the latitudes 37.45016 - 37.44982 N and longitudes 121.80560 - 122.21638 W, covering an area approximately about 480 km$^2$. The data were acquired on 9-July-98 at an altitude of near about 20,000 meters. The data are available in the form of a scene stored in band interleaved by pixel format with dimensions $224 \times 614 \times 512$ where 224 is number of bands, 614 is number of samples and 512 is number of lines in the scene. Pixel values are stored as 16-bit integers in inverted reflectance data format multiplied by 10000. A part of the scene, in true colour composite (Bands 63, 24, 8), is shown in Fig. 1. Further, fractal dimension images (FDI) have been created using the computed values of the fractal dimension of the spectral reflectance curve at each pixel location in HS images. FDI gets displayed by rescaling the FD values to 0-255. The FDI of the considered part of the scene is given in Fig. 2.

## Results and Discussion

In order to evaluate the quality of output from reduced dimension, the characteristics of the FDI are studied and compared with that of composite images (3 bands). A comparison and evaluation of both the images on the criteria of information content and visual assessment have been taken into consideration in carrying out the evaluation.

Entropy, a measure of information (Klir *et al.*, 2000) of an image has been used to find the inherent information content in the FDI and to compare it with that of original HS images. It is known that higher the entropy of an image, greater is its information content. Observing results from Table 1, it can be found that the entropy value for the original image is as low as 0.16010 (Image 1 band 8) and maximum as 6.05019 (Image 1 band 186). For FDI of Image 1 entropy value has been found to be 6.68657. Similar are the cases

**Fig. 1** True colour composite of a part of the scene (Bands 36, 24, 8 for RGB).

**Fig. 2** FDI image of hyperspectral data of the considered part of the scene.

for other FDIs. The entropy values for FDI of Image 2, Image 3, and Image 4 are 7.07133, 6.67140, and 6.687066 respectively. For the original images maximum entropy values are 5.34707 (band 100), 4.77166 (band 8), 5.60969 (band 100) respectively. These results suggest that the FDIs have a higher information content than HS images.

**Table 1** Entropy values of different images

| Image | Images | Entropy |
|---|---|---|
| Image 1 | HS-8 | 0.16010 |
| | HS-100 | 6.05019 |
| | HS-186 | 5.79360 |
| | FDI | 6.68657 |
| Image 2 | HS-8 | 5.32656 |
| | HS-100 | 5.34707 |
| | HS-186 | 1.81542 |
| | FDI | 7.07133 |
| Image 3 | HS-8 | 4.77166 |
| | HS-100 | 4.56157 |
| | HS-186 | 1.13029 |
| | FDI | 6.7140 |
| Image 4 | HS-8 | 5.55340 |
| | HS-100 | 5.60969 |
| | HS-186 | 1.53279 |
| | FDI | 6.87066 |

Observing carefully Fig. 1 and Fig. 2, it can be noted that the FDI have greater information content than the original true colour composite images.

## Identification of anomaly in water cover

In order to establish the potential of FD, it has been used as a feature for identification of anomaly in water cover. In this anomaly identification study, pure training samples from four dominant LCs, present in the part of the scene (Fig. 1) have been collected. The samples are extracted interactively from the AVIRIS image, using the image processing software ENVI 4.1. The Google image of the study area has been used as a reference data to identify the land cover. The training samples of pure water are then used to calculate the fractal dimension of SRC using (5). The statistics of the FD of pure water pixel is as given in Table 2. It has been found that the mean and standard deviation of FD for water are 1.492 and 0.027. These show that dispersion of FD as an indicator for water is low and thus it is a good feature for identification of water cover.

**Table 2** FD statistics for pure water cover

| Statistical criteria | FD values |
|---|---|
| Minimum | 1.4353 |
| Maximum | 1.5472 |
| Mean | 1.4922 |
| Median | 1.4873 |
| Standard deviation | 0.0273 |

Further, to classify the image using maximum likelihood classifier, three images of bands (8, 100, and 186) are taken to prepare a composite image of the scene. These bands are chosen depending on minimum correlation among these three bands and are widely separated in the Hyperspectral Electromagnetic Spectrum. The training samples collected have been used to classify the scene and generate an MLC classified output as shown in Fig. 3.

The region classified as water (red in colour) in Fig. 3 can also be identified as water (dark blue) in Fig. 1. The overall classification accuracy of the image is 98.66% and Kappa coefficient has been found to be 0.9784. These indicate that the classification is quite satisfactory. But, Google Image (Fig. 4) shows that the water region is not pure water cover (tone varies from dark to light blue). It is associated with sediment and other polluting materials. In order to test the potential of FD as a feature to identify this anomaly, two pixels having identification 79_45 (pure water pixel) and 79_53 (anomalous pixel as appear from visual identification) are considered. Their FD have been found [using (5)] to be 1.4867 and 1.6066 respectively. This indicates that the FD value of anomalous pixel is much different from the mean value of the training samples and even from the maximum value (1.547). These clearly indicate FD can be used as a feature to identify anomaly in water cover.

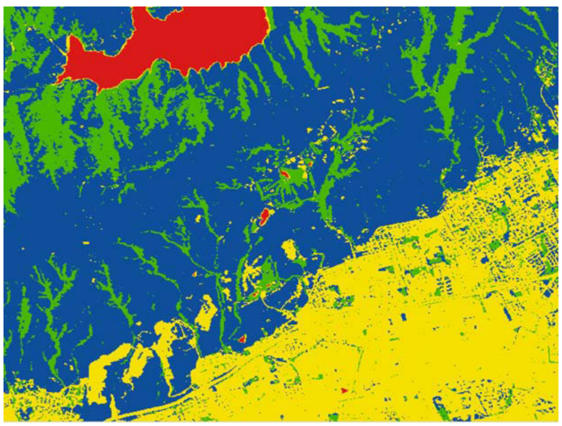

**Fig. 3** MLC Classified image of Fig. 1 Image (Bands 8, 100, 186) (Red: Water, Green: Vegetation, Blue: Rock, Yellow: Urban).


**Fig. 4** Google image of a portion of water region in Fig. 1.

## Conclusions

In this study, a fractal-based method for reduction of dimensionality of HS data has been proposed. The method is based on computation of fractal dimension of SRC at each pixel location of the HS image. The algorithm used for calculation of fractal dimension is linear in nature, thereby reducing computational overhead dramatically. The versatility of the proposed method lies in reduction of the dimensionality of data from N bands to a single band where N can be any number of bands. Further, the information content of the new dimensionally reduced single image is higher than any one of the bands of HS images of the same scene. On visual examination, many variations and anomalies present in the pixels can be clearly deduced from FDI.

The salient conclusions that may be arrived at from this study are that fractal technique provides very high reduction in dimensionality. The entropy or information content of an FDI is higher than any of the original HS images. Visual quality of the FDI is more informative than original true colour composite images.

This study has proved the use of the fractal method in automated feature extraction and anomaly recognition from HS images for the future.

**Acknowledgement** The data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the U.S. Geological Survey (USGS) Center for Earth Resources Observation and Science (EROS) http://LPDAAC.usgs.gov." Authors are thankful to anonymous reviewers for their constructive comments that improved the manuscript.

## References

Farrell MD and Mersereau RM (2005) On the impact of PCA dimension reduction for hyperspectral detection of difficult targets. *IEEE Geoscience and Remote Sensing Letters* 2(2):192–195

Huang Y, Lori Mann Bruce and Jiang Li (2001) Brushlet transform for hyperspectral feature extraction in automated detection of nut sedge presence in soybean. *IEEE Transactions on Geosciences and Remote Sensing* 39(3):527–529

Jensen JR (2003) Remote sensing of environment, Pearson Education, India

Kaewpijit S, Jacqueline Le Moigne and Tarek El-Ghazawi (2001) A hybrid algorithm for automatic detection of hyperspectral dimensionality. *Proceedings of IEEE Geosciences and Remote Sensing Symposium*, Vol.2, 649–651

Kaewpijit S, Jacqueline Le Moigne and Tarek El-Ghazawi (2003) Automatic reduction of hyperspectral imagery using wavelet spectral analysis. *IEEE Transactions on Geoscience and Remote Sensing* 41(4): 863–871.

Klaus Strodl, Ursula Bens, Frank Bliiser, Thomas Eiting and Alber Moreira DLR (1994) A comparison of several algorithms for on-board SAR raw data reduction. *German Aerospace Research Establishment (DLR) Institute for Radio Frequency Technology*, 2197–2199

Klir GJ, Folger and B Yuan (2000) Fuzzy sets and fuzzy logic, *Prentice-Hall India Pvt Ltd, New Delhi*, 574

Mandelbrot BB (1982) The fractal geometry of nature. *W.H. Freeman and Company, New York*, 1–468

Robila SA (2004) An analysis of spectral metrics for hyperspectral image processing. *Proceedings of IEEE Geosciencs and Remote Sensing Symposium, Vol.5,* 3233–3236

Schepers HE and HGM Johannes (1992) Four methods to estimate the fractal dimension from Self-Affine signals. *IEEE Engineering in Medicine and Biology* 11(2):57–64

Sevcik C (1998) A procedure to estimate the fractal dimension of waveforms, Complexity *International*, Vol. 5

Xiuping Jia and Richards JA (1999) Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification. *IEEE Transactions on Geoscience and Remote Sensing* 37(1): 538–542