Vegetation Classification from Time Series NOAA/AVHRR Data

Yoshifumi Yasuoka, Ai Nakagawa, Keiko Kokubu, Krishna Pahari
Institute of Industrial Science, University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505 JAPAN
E-mail: yyasuoka@cc.iis.u-tokyo.ac.jp

Mikio Sugita
Yamanashi Institute of Environmental Sciences
5597-1 Azakenmaruo, Kamiyoshida, Fujiyoshida-shi, Yamanashi 403-0005 Japan

Masayuki Tamura
National Institute for Environmental Studies
16-2 Onogawa, Tsukuba, Ibaraki 305-0053 Japan

Abstract

Vegetation cover classification is examined based on a time series NOAA/AVHRR data. Time series data analysis methods including Fourier transform, Auto-Regressive (AR) model and temporal signature similarity matching are developed to extract phenological features of vegetation from a time series NDVI data from NOAA/AVHRR and to classify vegetation types. In the Fourier transform method, typical three spectral components expressing the phenological features of vegetation are selected for classification, and also in the AR model method AR coefficients are selected. In the temporal signature similarity matching method a new index evaluating the similarity of temporal pattern of the NDVI is introduced for classification.

Key words: Time series NDVI, NOAA/AVHRR, Vegetation classification

1. Introduction

Vegetation distribution and its change are essential parameters in assessing the impact of human activities to environmental and climatic changes. And the importance of monitoring vegetation cover conditions has been pointed out in the management of environment in both of a regional and global scale. It is, however, not easy to survey vegetation cover conditions over extensive areas only with the ground observation. Remote sensing from satellite is expected to provide an efficient tool to monitor vegetation distribution and its changes.

In this study vegetation classification methods are investigated by using multi-temporal satellite data. A time series NDVI data from NOAA/AVHRR is used to extract phenological features of vegetation and to classify vegetation types. Time series data analysis methods including Fourier transform (Andres, L., 1994, Sugita, M., 1997), Auto-Regressive model (Box, G.E.P., 1970) and temporal signature similarity matching are applied to the monthly composite NDVI data covering East Asia region. In the Fourier transform method, typical three spectral components expressing the phenological features of vegetation are selected for classification, and also in the AR model method AR coefficients are selected by solving the Yule-Walker equation. In the temporal signature similarity matching method a new index evaluating the similarity of temporal pattern of the NDVI is introduced for classification. The effectiveness of the classification based on the phenology of vegetation is validated with so called "Green Census" data which is a set of actual vegetation map with the scale of 1/50,000 covering whole areas of Japan archipelago.

2. Time series NDVI data and "Green Census" data

2.1 Monthly NDVI data

A time series NDVI images was produced from monthly composite NOAA/AVHRR data set from 1996 and 1997 by National Institute for Environmental Studies. The images cover most of the East Asian region with a spatial resolution of 1.1km. Geometric correction was performed with the ground control points (GCPs).

2.2 "Green Census" data

The "Green Census" data has been produced by Japan Environment Agency around once five years since 1973, and it includes actual vegetation maps with a scale of 1/50,000, and their digital format data (vector data). It covers whole Japan archipelago prefecture by prefecture, and the number of vegetation categories in the actual vegetation map is as large as 766.

3. Vegetation classification with time series analysis methods

3.1 Fourier Transform

Fourier spectrum of NDVI time series data

Monthly NDVI time series data V(t) at a certain NDVI image coordinate of (x, y) is a function of time t ,which is in discrete unit of month and has a range from 1 to 24 (2 years). With Fourier Transform. V(t) is written in a form of discrete Fourier series such as

$$V(t)=\sum_{n=0}^{m-1} f(n)\exp\left(-\frac{2\pi nt}{m}i\right), \quad (1)$$

where n-th imaginary Fourier coefficient f(n) is given by

$$f(t)=\sum_{t=1}^{m} V(t)\exp\left(-\frac{2\pi nt}{m}i\right). \quad (2)$$

The power density spectrum P(n) of V(t) is given by

$$P(n)=\frac{1}{m}|f(n)|^2. \quad (3)$$

Here, Fourier power density spectra was calculated for each pixel in the NDVI time series images over Japan area. Figure 1 shows typical Fourier power density spectra calculated for five different land cover categories. In this Figure it is indicated that the spectral components corresponding to seasonal cycle in vegetation phenology are dominant. Each Fourier coefficient is used to remove the noise of the high frequency, and to extract periodicity at the same time.

Classification

In this study the number of classified categories which can be distinguished by satellite data is supposed to be beyond 30, therefore training area is set automatically from actual vegetation data by random sampling. As phenology of vegetation is the key characteristics of land cover classification, 0-th, 1-th, 2-th and 4-th power spectrum, corresponding to averaged value of two years period, 12 months period component and 6 months period component are used. Classification is done with maximum likelihood classification. Figure 2. shows an example of land cover classification with Fourier spectrum.

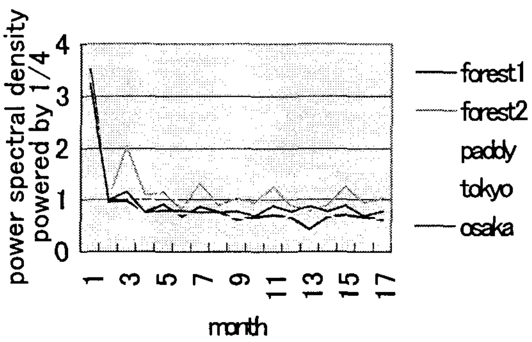

Fig. 1 Examples of Fourier power spectra of 24 months NDVI data.

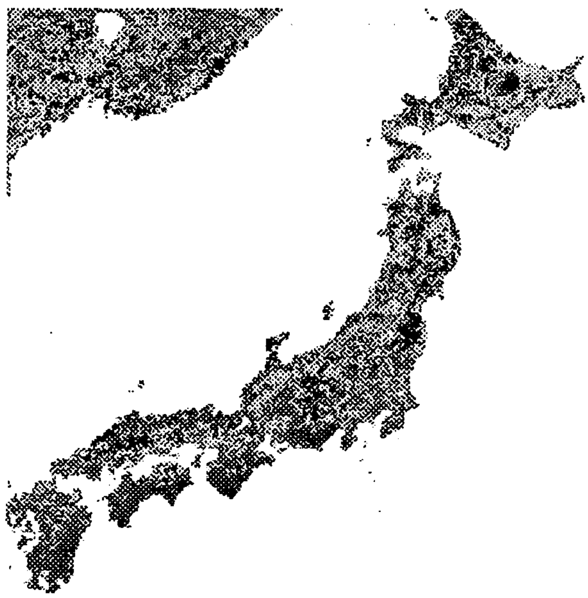

Fig. 2 An Example of vegetation classification with Fourier spectrum.

3.2 Auto-Regressive model

Auto-Regressive model for a time series NDVI data is described as follows;

$$V(t) + \sum_{k=1}^{p} a_k \cdot V(t-k) = \varepsilon(t) \qquad (4)$$

where $\varepsilon(t)$ is white Gaussian noise. Time series data is characterized by the coefficient a_k, and the coefficients are calculated by solving the Yule-Waker equation,

$$\sum_{l=1}^{p} a_l \cdot R_{VV}(k-l) = -R_{VV}(k) \qquad (5)$$

$$k=1,2,\cdots,p$$

where R_{VV} is an auto-correlation function of $V(t)$. Here for each time series NDVI the coefficients are calculated, and based on these coefficients,

each pixel is classified into vegetation categories with maximum likelihood classification scheme.

3.3 Temporal signature similarity matching

A new index of similarity for temporal signature of NDVI, TSS (Temporal Signature Similarity), is devised to make vegetation clusters depending on temporal signatures (phenological features) and classify each signature to a vegetation clusters. The $TSS_{i,j}$ is defined for two temporal signatures of NDVI, i and j, and described as follows;

$$Si = \{S_1^i, S_2^i, \cdots, S_{12}^i\}$$

$$Sj = \{S_1^j, S_2^j, \cdots, S_{12}^j\}$$

$$|d| = \sum_{k}^{12} \{abs(S_k^i - S_k^j)\}$$

$$TSS\,ij = |d| / Rij. \qquad (6)$$

Where, S_i and S_j are the signatures of clusters i and j as represented in the 12 monthly NDVI plots, |d| is the average absolute distance between the signaures and R_{ij} is the coefficient of correlation between the plots for i and j respectively.

The process of classification with this method involves first, clustering the original image into a large number of clusters using simple isodata clustering, calculating the TSSij of each cluster with all other clusters and thus making a matrix of TSSij between clusters, and then merging clusters with the least value of TSSij between them until a condition is achieved whereby the the total number of classes is equal to the desired number of meaningful classes of interest. Each class is then assigned a specific class value based on the ground truth data or any existing source of land cover information.

3.4 Comparison of Classified Vegetation Distribution and "Green Census" Data

Digital actual vegetation data from "Green Census" is registered with the classified results. There are 766 vegetation categories in the actual vegetation data set which is used as the precise validation data. It is impossible to categorize all the vegetation classes in the validation data directly from the satellite data, therefore 766 categories are stratified according to phenology. Also original vegetation vector data whose precision is about 100m is converted into 50m mesh data. It is integrated to the 1km mesh data to be compared with the result of classification. Figure 3 shows an example of actual vegetation data for Osaka Prefecture. The result of validation is presented in the Symposium.

4. Conclusions

Vegetation classification based on the time series NDVI data from NOAA/AVHRR is investigated to observe the vegetation distribution over the extensive area. The result is compared with the so-called "Green Census" data, and the feasibility of time series NDVI data for vegetation classification was evaluated. Three methods including Fourier transform, Auto-Regressive model and temporal signature similarity matching were applied to the time series NDVI data. The results demonstrated the possibility of using time series satellite data for vegetation classification.

Authors would like to express our thanks to the Japan Environment Agency for providing us with the "Green Census" data, and also thanks to *National Institute for Environmental Studies for* providing us with the NDVI data set.

References

(1) Andres, L., et.al. (1994): Fourier analysis of multi-temporal AVHRR data applied to a land classification, Int. J. Remote Sens., Vol.15 (5), pp.1115-1121.

(2) Box,G.E.P. and G.M. Jenkins (1970): *Time Series Analysis Forcasting and Control,* Holden-Day.

(3) Sugita, M. and Y. Yasuoka (1997): Land cover classification of East Asia using Fourier spectra of monthly NDVI data, Proc. of IGARSS'97.