

Hyperspectral Data Transformation and Vegetation Index Performance Based on the Universal Pattern Decomposition Method

Lifu Zhang

*Institute of Remote Sensing and GIS, Peking University, Beijing 100871, China
E-mail: Zhanglifu01@yahoo.com.cn*

Liangpei Zhang

*National Key Laboratory for Information Engineering in Surveying, Mapping, and Remote Sensing,
Wuhan University, Wuhan 430079, China*

Lei Yan

Institute of Remote Sensing and GIS, Peking University, Beijing 100871, China

Noboru Fujiwara

*Laboratory of Nature Information Science, Department of Information and Computer Sciences,
Nara Industrial University, Nara 630-8530, Japan*

Kanako Muramatsu

*Laboratory of Nature Information Science, Department of Information and Computer Sciences,
Nara Women's University, Nara 630-8506, Japan*

Motomasa Daigo

Department of Economics, Doshisha University, Kyoto 602-8580, Japan

Abstract. *Traditional vegetation indices are based on only a few spectral bands. However, hyperspectral spectrometers, such as the airborne visible infrared imaging spectrometer (AVIRIS), collect data with 224 contiguous spectral bands. Traditional vegetation index extraction methods lose much of the information contained in hyperspectral data. The universal pattern decomposition method (UPDM) is tailored for hyperspectral data analysis. In this article, we consider the UPDM as a type of multivariate analysis; standard patterns are interpreted as an oblique coordinate system and coefficients are thought of as the coordinates of a pixel's reflectance. This article describes UPDM hyperspectral data transformation of AVIRIS data, the performance of a vegetation index based on the universal pattern decomposition method (VIUPD), and the influences of a noise-to-vegetation index. The results demonstrate that the VIUPD is an effective vegetation information extraction approach for hyperspectral data. The VIUPD is more sensitive to vegetation conditions than the normalized difference vegetation index and enhanced vegetation index. Furthermore, noise influences can be neglected in VIUPD computations, with satisfactory accuracy. © 2007 Society for Imaging Science and Technology.*

[DOI: 10.2352/J.ImagingSci.Technol.(2007)51:2(141)]

INTRODUCTION

Vegetation indices (VIs) are spectral transformations of two or more bands designed to enhance vegetation properties and allow for reliable representations of photosynthetic ac-

tivity and structural canopy variations.¹ These indices are sensitive to a variety of biophysical vegetation canopy parameters, such as the leaf area index, fraction of vegetation cover, leaf angle distribution, and leaf chlorophyll concentration.² VIs derived from satellite remotely sensed data are primary sources of information for operational monitoring of the earth's vegetative cover. Most of these VIs are called broadband vegetation indices because they are based on algebraic combinations of reflectance in the red (*R*) and near infrared (NIR) spectral bands.³ The broadband indices use average spectral information over broad bandwidths, resulting in the loss of critical information available in specific narrow bands.⁴ In addition, broadband indices are heavily affected by the soil background in conditions of low vegetation cover.⁵

Hyperspectral sensors measure reflectance in a large number of narrow wavebands, generally with bandwidths of less than 10 nm. Reflectance and absorption features related to specific crop physical and chemical characteristics can be detected with these narrow bands.⁶ With many airborne imaging spectrometer systems in use today and the rise of spaceborne hyperspectral sensors, better understanding of this type of image data is increasingly needed.⁷ The airborne visible infrared imaging spectrometer (AVIRIS) has 224 contiguous spectral bands.⁸ However, application of a traditional vegetation index extraction method will lose much of the information contained in hyperspectral data. Thus, VIs

derived from the R and NIR bands are unsuitable for hyperspectral data (i.e., AVIRIS) analysis.

The universal pattern decomposition method (UPDM) is a sensor-independent method that is tailored for analysis of satellite-derived data.^{9,10} Sets of spectral reflectance measured by a sensor are transformed by the UPDM into three or four coefficients with three or four fixed spectral reflectance patterns. Spectral reflectance patterns are determined in the spectral region between 350 and 2500 nm and are called the “universal standard spectral patterns.” Sensor wavelength values are selected from the universal standard spectral patterns to analyze the spectral region of each sensor. These coefficients are “pattern decomposition coefficients.” The UPDM can be explained using two analysis methods: spectral mixing analysis and multivariate analysis. For the former, the UPDM expresses the spectrum of each pixel as the linear sum of three fixed, standard spectral patterns (i.e., the patterns of water, vegetation, and soil); each coefficient represents the ratio of the spectral patterns of the three components. If we think of the UPDM as multivariate analysis, standard patterns are interpreted as an oblique coordinate system, and coefficients are thought of as the coordinates of a pixel’s reflectance. If an additional supplementary pattern is needed, the reason for this addition is made clearer by the multivariate analysis explanation than that by the spectral mixing analysis. The UPDM method has been successfully applied to simulated data with wavelengths observed by Landsat/ETM+, Terra/MODIS, ADEOS-II/GLI, and the 92 bands-CONTINUE sensors⁹ and validated using MODIS and ETM+ satellite data from over the Three Gorges region in China.¹¹

This article examines the UPDM hyperspectral data transformation method using AVIRIS hyperspectral data; the performance of a vegetation index based on the universal pattern decomposition method (VIUPD); a comparison of the VIUPD, normalized difference vegetation index (NDVI), and enhanced vegetation index (EVI); and the influences of noise on the above VIs.

METHOD

The Universal Pattern Decomposition Method

We developed a UPDM for hyperspectral data analysis.¹² The UPDM decomposes reflectance values at each pixel into a linear sum of standard spectral patterns for water, vegetation, soil, and any additional patterns using the following equation:^{11,12}

$$R(i) \rightarrow C_w \cdot P_w(i) + C_v \cdot P_v(i) + C_s \cdot P_s(i) + C_a \cdot P_a(i), \quad (1)$$

where $R(i)$ is the reflectance of band i measured by the satellite sensor; C_w , C_v , and C_s are the respective decomposition coefficients; C_a represents additional coefficients; and $P_w(i)$, $P_v(i)$, and $P_s(i)$ are the standard spectral patterns of water, vegetation, and soil for some typical sensor, which are intercepted from the same standard pattern normalized in the same wave region of 350–2500 nm for any sensor, with respect to the properties of the sensor. $P_a(i)$ is the additional standard pattern for i bands and is an optional component

that is also controlled according to the purpose of the study.

Equation (1) can be expressed using matrix notation as follows:

$$\mathbf{R} = \mathbf{PC} + \mathbf{r} \quad (2)$$

or

$$\begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} = \begin{pmatrix} P_{1w} & P_{1v} & P_{1s} & P_{14} \\ P_{2w} & P_{2v} & P_{2s} & P_{24} \\ \vdots & \vdots & \vdots & \vdots \\ P_{nw} & P_{nv} & P_{ns} & P_{n4} \end{pmatrix} \cdot \begin{pmatrix} C_w \\ C_v \\ C_s \\ C_a \end{pmatrix} + \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}, \quad (3)$$

where $\mathbf{R}=[R_1, R_2, \dots, R_n]^T$ is the column vector of observations, n is the number of spectral bands, $\mathbf{P}=[\mathbf{P}_w, \mathbf{P}_v, \mathbf{P}_s, \mathbf{P}_a]$ is the $n \times 4$ matrix of which the row vector is the standard spectral pattern for band number n , $\mathbf{C}=[C_w, C_v, C_s, C_a]^T$ is the column vector of UPDM coefficients, and \mathbf{r} is the residual column vector for band i .

The three standard spectral patterns (column vectors $\mathbf{P}_w, \mathbf{P}_v, \mathbf{P}_s$) are intercepted from $P_k(\lambda)$ ($k=w, v, s$), which are the normalized reflectance of $R_k(\lambda)$ in the wavelength region from 350 to 2500 nm according to Eq. (4),

$$P_k(\lambda) = \frac{\int d\lambda}{\int R_k(\lambda) d\lambda} R_k(\lambda), \quad (4)$$

where the $R_k(\lambda)$ values are the spectral reflectance patterns of standard objects, and $\int d\lambda$ refers to integration of the total wavelength range. The shapes and magnitudes of the standard patterns $P_k(\lambda)$ are fixed for any sensor.¹⁵

The fourth standard spectral pattern (column vector \mathbf{P}_a) is intercepted from $P_a(\lambda)$, which is defined using the normalized residual value of yellow-leaf, expressed by

$$P_a(\lambda) = \frac{r_a(\lambda) \int d\lambda}{\int |r_a(\lambda)| d\lambda}, \quad (5)$$

where $r_a(\lambda)$ is the residual value for a yellow leaf relative to i band, given by Eq. (6).

$$r_a(\lambda) = R_a(\lambda) - \{C_w P_w(\lambda) + C_v P_v(\lambda) + C_s P_s(\lambda)\}. \quad (6)$$

Here, $R_a(\lambda)$ is the measured value for the yellow-leaf sample and $r_a(\lambda)$ is the residual value. For any sensor, the standard patterns for each sensor are defined by

$$P_k(i) = \frac{\int_{\lambda_s(i)}^{\lambda_e(i)} P_k(\lambda) d\lambda}{\int_{\lambda_s(i)}^{\lambda_e(i)} d\lambda} \quad (k = w, v, s, a), \quad (7)$$

where $\lambda_s(i)$ and $\lambda_e(i)$ are the start and end wavelengths for band i , respectively, and $\int_{\lambda_s(i)}^{\lambda_e(i)} d\lambda$ is the wavelength width of

band i . The values of $P_w(i)$, $P_v(i)$, $P_s(i)$, and $P_a(i)$ differ with the sensor, but they are all intercepted from the same normalized standard spectral pattern.

The multiple value of $\int d\lambda$ used in numerator of Eqs. (4) and (5) is only for the purpose of expands the normalized result value, respectively.

Inverting Eq. (2) and minimizing the sum-of-squared-error criterion function yields a unique solution of \mathbf{C} as follows:

$$\mathbf{C} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{R}, \quad (8)$$

where \mathbf{R} is a vector known from satellite data and \mathbf{P} is a standard spectral pattern matrix as described above. The spectral pattern matrix is derived from normalized standard spectral patterns of water, vegetation, soil, and supplementary data.¹¹

Vegetation Index Algorithm

A commonly used vegetation index is the NDVI, given as follows:

$$\text{NDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + \rho_{\text{red}}}. \quad (9)$$

However, this index uses only red and near infrared reflectance data.¹³ The EVI uses the red and near infrared bands but also includes blue-band reflectance data to correct for aerosol influences in the red band, and some other aerosol resistance coefficients.¹ The EVI is given as follows:

$$\text{EVI} = G \frac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + C_1 \times \rho_{\text{red}} - C_2 \times \rho_{\text{blue}} + L}, \quad (10)$$

where ρ values are atmospherically corrected or partially atmospherically corrected (e.g., for Rayleigh and ozone absorption) surface reflectances; L is the canopy with background adjustment addressing nonlinear, differential near infrared- and red-band radiant transfer through the canopy; and C_1 , C_2 are aerosol resistance coefficients that use the blue band to correct for red-band aerosol influences. The coefficients of the EVI algorithm are $L=1$, $C_1=6$, $C_2=7.5$, and $G=2.5$, where G is the gain factor.¹⁴

The above methods, namely broadband indices, use either two or three satellite-observed wavelength bands or require some additional coefficient inputs. We developed a new VIUPD.¹⁵ The VIUPD is defined as the linear sum of the pattern decomposition coefficients but is sensor independent. The new vegetation index is normalized by total reflectance or total brightness to minimize shadow effects and obtain stable values. The index is a function of the linear combination of the pattern decomposition coefficients. The formula is given as follows:

$$\text{VIUPD} = \frac{(C_v - a \times C_s - C_a)}{C_w + C_v + C_s}, \quad (11)$$

where $(C_w + C_v + C_s)$ represents the sum of total reflectance, even if a additional pattern is included, because the inte-

Table I. Wavelength regions and the AVIRIS bands used in this analysis.

Selected wavelength (nm)	AVIRIS band	AVIRIS bands actually used
371.0–900.0	2–57	5–57
991.0–1100.0	68–78	68–78
1191.0–1300.0	88–101	88–101
1521.0–1750.0	124–147	124–147
2081.0–2360.0	182–209	182–199

grated value of $\int r_a(\lambda) d\lambda$ equals zero, and coefficient a has a value of 0.10 and represents the standard soil pattern coefficient.¹⁵ The value of coefficient a is experimentally determined so that for the dead leaf, the average value of VIUPD is about zero, and for standard vegetation, it is about 1.0.

DATA AND ALGORITHM PERFORMANCE

The AVIRIS instrument contains 224 different detectors, each with a wavelength sensitive range (also known as a spectral bandwidth) of approximately 10 nm, allowing it to cover the entire range between 380 and 2500 nm.⁸ AVIRIS standard data products were downloaded from Jet Propulsion Laboratory Web site.¹⁶ These data were acquired over Moffett Field, CA, an area that included vegetation, urban structures, and water. The measurement data were converted to ground reflectance data with proper calibration and correction for atmospheric effects.

In this study, we used the same water, vegetation, soil, and additional standard spectral patterns as Refs. 9–12 and 15, i.e., the fixed standard spectral pattern for any sensor. For AVIRIS data, bands that fell into regions with strong atmospheric absorption and those with poor quality or zero values were removed. Table I shows wavelength regions and the AVIRIS bands used in this analysis.

In the UPDM, however, we converted an uninterrupted spectral wavelength range from 350 to 2500 nm, excluding regions with strong atmospheric absorption. The excluded spectral regions are 901–990, 1101–1190, 1301–1520, and 1751–2080 nm, respectively (Table I). Therefore, the total number of bands equaled 1260 with 1 nm bandwidth. When applying the UPDM approach to AVIRIS data, we intercepted the AVIRIS standard matrix \mathbf{P} using Eq. (7). Because some of AVIRIS bands had zero signal values, the final band number was 120; thus $\mathbf{P} = [\mathbf{P}_w, \mathbf{P}_v, \mathbf{P}_s, \mathbf{P}_a]$ is the 120×4 matrix vector. Figure 1(a) shows the universal 1260-band normalized standard patterns of soil, water, vegetation, and the additional pattern; Fig. 1(b) shows the 120-band AVIRIS normalized standard patterns.

Figure 2 shows the steps for AVIRIS data transformation and vegetation index computation. The 120-band AVIRIS data were obtained from the original 224-band data recorded by the AVIRIS sensor; VIUPD was computed using UPDM coefficients after conducting AVIRIS data transformation. The NDVI was computed using Eq. (9) and the

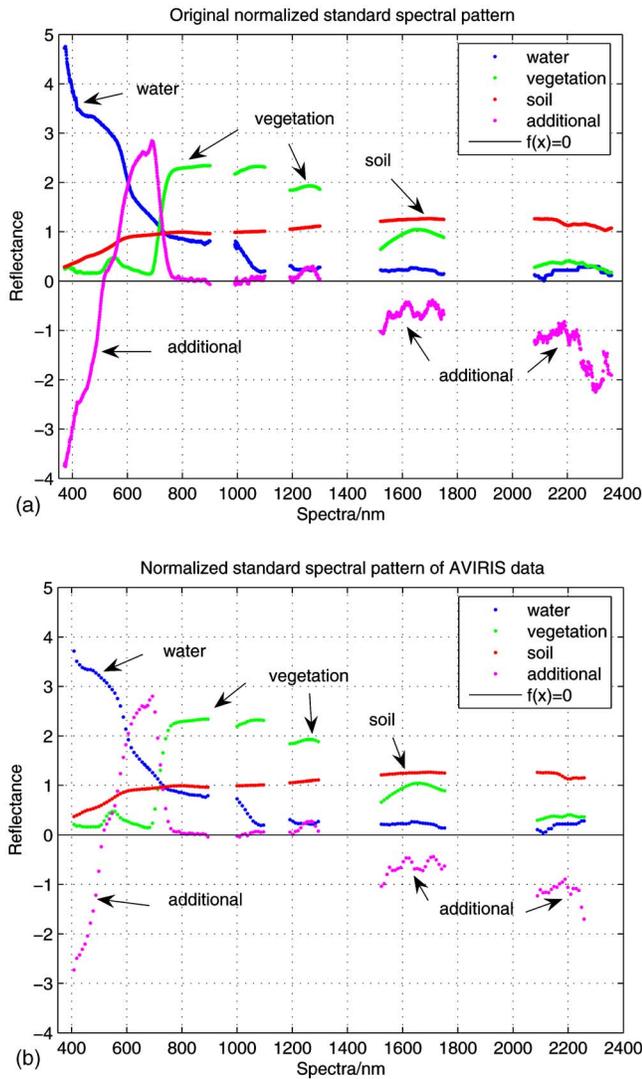


Figure 1. The universal 1260-band normalized standard patterns of soil, water, vegetation, and the additional pattern (a) and the 120-band AVIRIS standard pattern (b). (b) was intercepted from (a) using Eq. (7).

directly selected near-infrared band 47 (837 nm) and red band 25 (646 nm). The EVI was created using Eq. (10) and the same NIR and R bands as the NDVI, as well as the blue band 6 (458 nm).

RESULTS AND DISCUSSION

Sensitivity Comparison of Vegetation Indices

In this study, three vegetation indices (the VIUPD, NDVI, and EVI) were computed from AVIRIS data. Figure 3 shows the VIUPD, EVI, and NDVI images (400 × 400 pixels) derived from the AVIRIS data. From Fig. 3 we can see that the VIUPD image shows more detailed information than the NDVI and EVI images. The theoretic interpretation of which is that the VIUPD was computed using four UPDM coefficients, i.e., the VIUPD was a linear function of C_w , C_v , C_s , C_a , and was normalized with the total reflectance value. The traditional broadband vegetation indices are usually constructed with NIR and R bands, while the VIUPD is computed using all the observed wavelengths, excluding

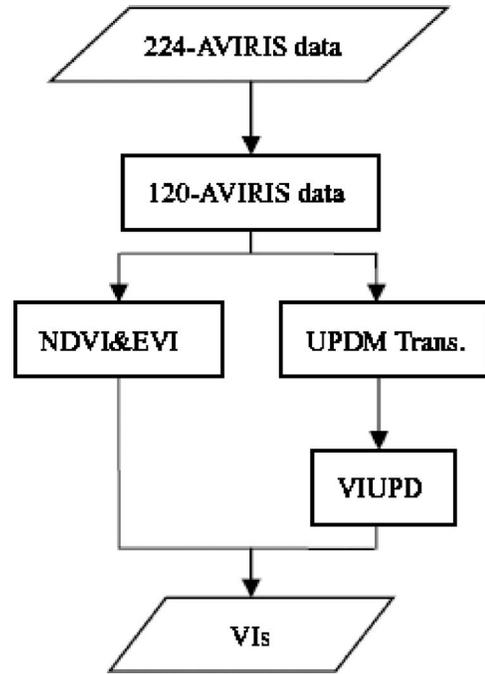


Figure 2. Flowchart of vegetation index computation.

wavelengths with strong atmospheric absorption and bands with poor quality data. Therefore, the VIUPD is more sensitive than the EVI and NDVI as shown in Fig. 3. Previous studies have reported that the VIUPD reflects vegetation concentrations, the amount of CO₂ absorption, and the degree of terrestrial vegetation vigor more sensitively than the NDVI and EVI, or is especially sensitive to CO₂ absorption,¹⁵ however, the data used are ground-measured reflectance data (i.e., measured using ASD field spectra instrument) rather than data acquired via aircraft-mounted or satellite sensor.

This paper assess the performance of the VIUPD using the AVIRIS image collected in 1997 over the Moffett Field (CA), However, without the knowledge of ground truth data to be used, the experiments using the AVIRIS data are only qualitative with a reasonable theoretic explanation.

For proving the better performance of the VIUPD over NDVI and EVI, we also compared the relationship between the VIUPD and the traditional EVI and NDVI (top left corner of Fig. 3; 200 × 200 pixels), as shown in Fig. 4. Figure 4(a) shows the relationship between the VIUPD and EVI; the horizontal axis is the VIUPD, and the vertical axis represents the EVI. Compared to the EVI, the VIUPD expands the ranges of the vegetation index. For example, the values for water range from -0.2 to 0.4 for the VIUPD but only from -0.1 to 0.1 for the EVI, as shown by Fig. 4(a). The VIUPD can also detect small vegetation changes caused by the phytoplankton concentration in the water. For the soil and vegetation samples, the VIUPD and EVI range from 0.2 to 1.2 and 0.2 to 0.8, respectively. The linear correlation coefficient is 0.9425 between the EVI and VIUPD for all samples. This means that the VIUPD and EVI have a high correlation but the regression lines have a lower slopes (Fig. 4(a)), thus the VIUPD is more sensitive than the EVI. The

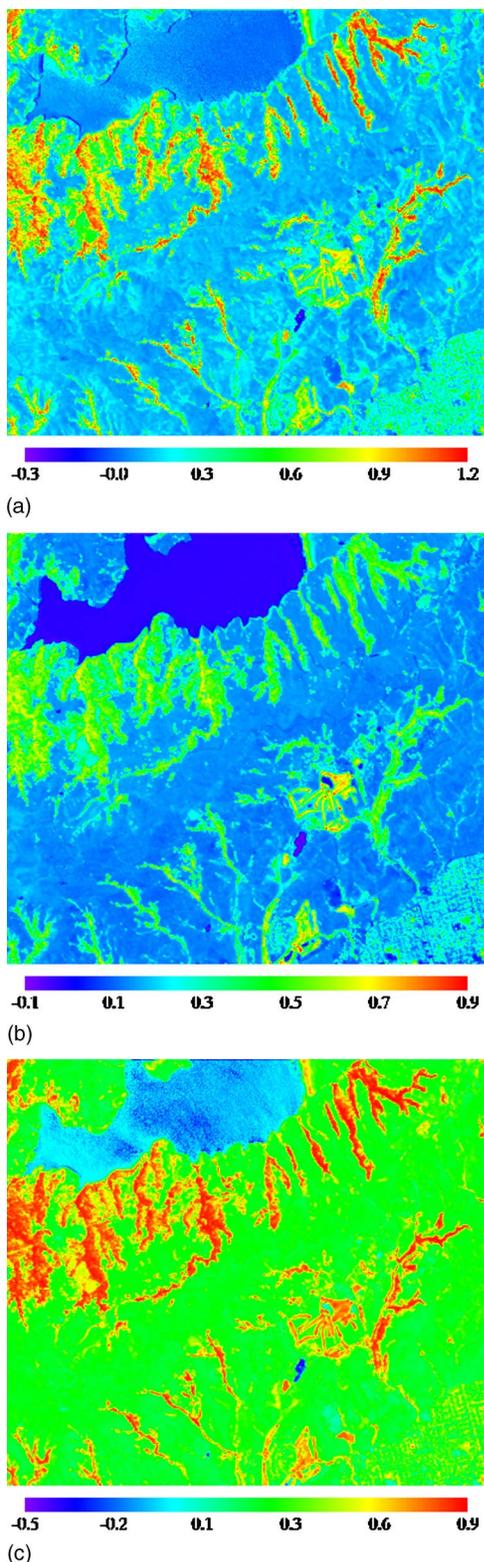


Figure 3. Image of vegetation indices derived from AVIRIS data (400 × 400 pixels).

lower and the upper limits for the VIUPD are about -0.3 and 1.2, which are determined by the factor a in the VIUPD formula.

Figure 4(b) shows the relationship between the VIUPD

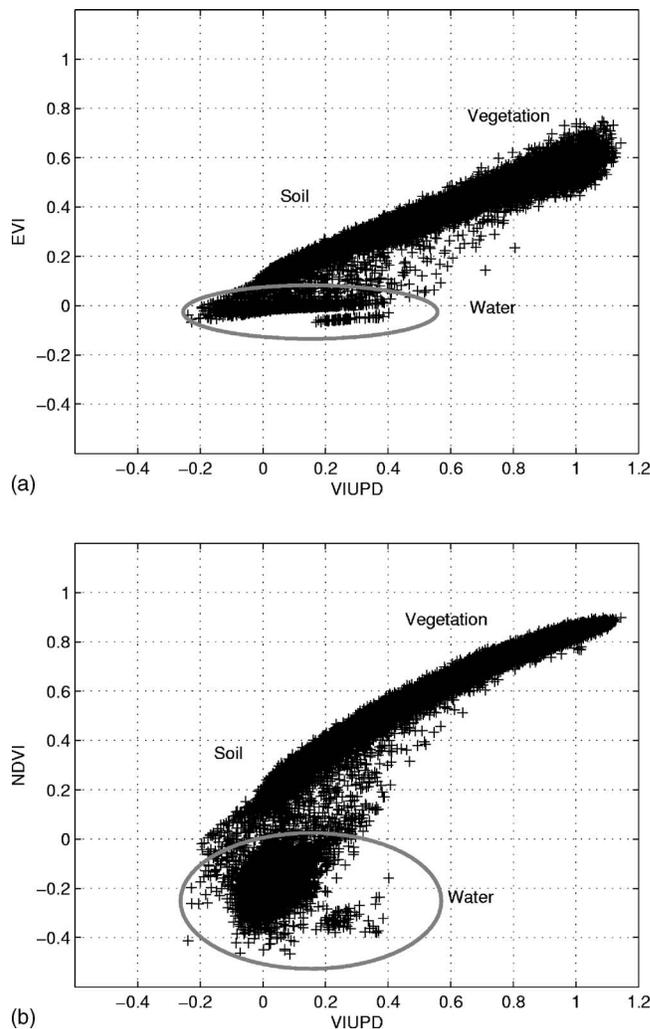


Figure 4. Relationship between the VIUPD and traditional EVI and NDVI.

and NDVI. The horizontal and vertical axes represent the VIUPD and NDVI, respectively. For water, the VIUPD and NDVI have approximately the same change in scope. However, for other samples, they show an increasingly depressed trend for the NDVI value, with the NDVI finally reaching its saturation point. The linear correlation coefficient is 0.8960 between the NDVI and VIUPD for all samples.

Noise Influences on the Vegetation Indices

To estimate the influences of noise on the vegetation index, we compared the vegetation index derived from original data and those derived from noise-removed data using the principal component transform (PCT) method illustrated in Fig. 5.¹⁷ Applying the PCT algorithm to AVIRIS data, the first 16 components of the PCT contained 99.99% of the data variances with the noise removed. The horizontal axes of (a), (b), and (c) show the original data derived VIUPD, EVI, and NDVI, respectively. The vertical axis shows the noise-removed vegetation index. Figure 5 shows that the VIUPD and EVI have neglectable noise influence, while the NDVI has more influences for NDVI values below zero.

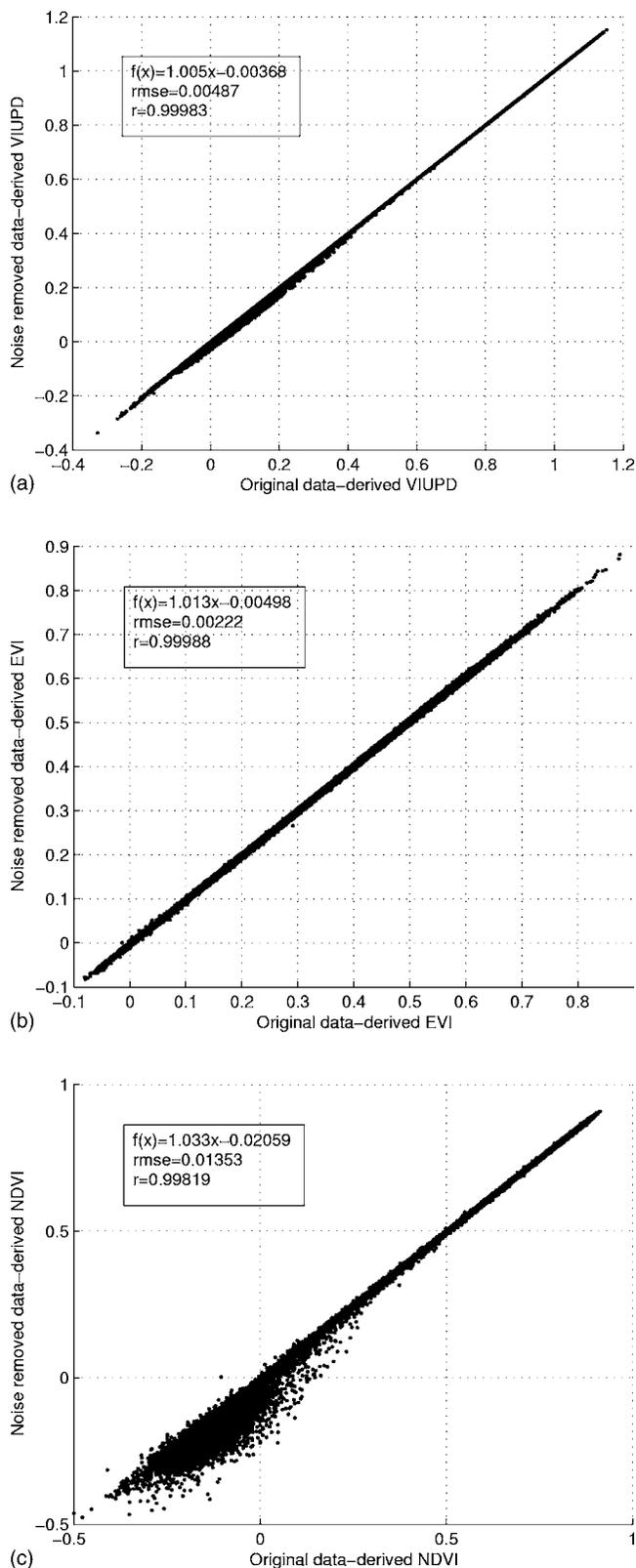


Figure 5. Vegetation indices derived from original data vs those derived from noise-removed data using the PCT method.

CONCLUSIONS

This study examined UPDM hyperspectral data transformation of AVIRIS hyperspectral data, the performance of a VIUPD and the influences of noise on vegetation indices.

The VIUPD reflects the linear sum of the four pattern decomposition coefficients. Two or three reflectance bands are used to calculate the EVI and NDVI, while the VIUPD is computed using four UPDM coefficients. Thus the VIUPD is more suitable for hyperspectral data analysis than the EVI and NDVI.

The results demonstrate that the VIUPD is an effective vegetation information extraction approach for hyperspectral data. The VIUPD is more sensitive to vegetation conditions than the NDVI and EVI. Furthermore, noise influences can be neglected in VIUPD computations with satisfactory accuracy.

The experiments using the AVIRIS data are only qualitative for the lack of ground real reference data. The lower and the upper limits for the VIUPD are about -0.3 and 1.2 , respectively, which are determined by the factor a in the VIUPD formula.

ACKNOWLEDGMENTS

We wish to thank Jet Propulsion Laboratory for providing the AVIRIS data on the web site. This study was supported by the National Natural Science Foundation of China (Project No. 40471088), Open Research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (Project No. WK(06)0102), and the China Postdoctoral Science Foundation (Project No. 2005038002).

REFERENCES

- ¹A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices", *Remote Sens. Environ.* **83**, 195 (2002).
- ²A. Huete, H. Q. Liu, K. Batchily, and W. J. D. Leeuwen, "A comparison of vegetation indices over a global set of TM images for EOS-MODIS", *Remote Sens. Environ.* **59**, 440 (1997).
- ³A. Bannari, D. Morin, F. Bonn, and A. Huete, "A review of vegetation indices", *Remote Sens. Rev.* **13**, 95 (1995).
- ⁴G. A. Blackburn, "Quantifying chlorophylls and carotenoids at leaf and canopy scales: An evaluation of some hyperspectral approaches", *Remote Sens. Environ.* **66**, 273 (1998).
- ⁵F. Baret and G. Guyot, "Potentials and limits of vegetation indices for LAI and APAR assessment", *Remote Sens. Environ.* **35**, 161 (1991).
- ⁶I. B. Strachana, E. Pattey, and J. B. Boisvert, "Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance", *Remote Sens. Environ.* **80**, 213 (2002).
- ⁷K. Lee, W. B. Cohen, R. E. Kennedy, T. K. Maierperger, and S. T. Gower, "Hyperspectral versus multispectral data for estimating leaf area", *Remote Sens. Environ.* **91**, 508 (2004).
- ⁸G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter, "The airborne visible/infrared imaging spectrometer (AVIRIS)", *Remote Sens. Environ.* **44**, 127 (1993).
- ⁹L. Zhang, Y. Mitsushita, S. Furumi, K. Muramatsu, N. Fujiwara, M. Daigo, and L. Zhang, "Universality of modified pattern decomposition method for satellite sensors", Asia GIS Conference, Wuhan, 8–11 October 2003. Available online at: www.hku.hk/cupem/asiagis/publications.htm (accessed April 8 2005).
- ¹⁰L. Zhang, S. Furumi, K. Muramatsu, N. Fujiwara, M. Daigo, and L.

- Zhang, "Sensor-independent analysis method for hyperspectral data based on the pattern decomposition method", *Int. J. Remote Sens.* **27**, 4899 (2006).
- ¹¹ L. Zhang, N. Fujiwara, S. Furumi, K. Murumatsu, M. Daigo, and L. Zhang, "Assessment of the universal pattern decomposition method using MODIS and ETM+ data", *Int. J. Remote Sens.* **28**, 125 (2007).
- ¹² L. Zhang, L. Zhang, N. Fujiwara, and K. Murumatsu, "The universal pattern decomposition method based on the hyperspectral satellite remotely sensed data", (in Chinese) *Journal of Geomatics and Information Science of Wuhan University* **27**, 264 (2005).
- ¹³ R. Nemani, L. Pierce, S. Running, and L. Band, "Forest ecosystem processes at the watershed scale: Sensitivity to remotely-sensed leaf area index estimates", *Int. J. Remote Sens.* **14**, 2519 (1993).
- ¹⁴ A. Huete, H. Q. Liu, K. Batchily, and W. J. D. van Leeuwen, "A comparison of vegetation indices over a global set of TM images for EOS-MODIS", *Remote Sens. Environ.* **59**, 440 (1997).
- ¹⁵ L. Zhang, S. Furumi, K. Murumatsu, N. Fujiwara, M. Daigo, and L. Zhang, "A new vegetation index based on the universal pattern decomposition method", *Int. J. Remote Sens.* **28**, 107 (2007).
- ¹⁶ <http://aviris.jpl.nasa.gov/html/aviris.freedata.html>.
- ¹⁷ J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis*, 3rd ed. (Springer-Verlag, Berlin, 1999), pp. 133–148.