Land Cover Classification based on the Universal Pattern Decomposition Method

L. Zhang [a,*], L. Zhang [a], N. Fujiwara [b], K. Muramatsu [c], M. Daigo [d], S. Furumi [c]

[a] LIESMARS, Wuhan University, Wuhan, 430079, China
[b] Department of Information and Computer Sciences, Nara Industrial University, Nara, Japan
[c] Department of Information and Computer Sciences, Nara Women's University, Nara, Japan
[d] Department of Economics, Doshisha University, Kyoto, Japan

Abstract – The universal pattern decomposition method (UPDM) has been successfully applied to simulated data for Landsat/ETM+, Terra/MODIS, ADEOS-II/GLI and others using ground-measured data. The UPDM is tailored to decrease dimensions of hyper multi-spectral data that have sensor-independent characteristics and thus exploit hyper multi-spectral remotely sensed data classification. In this study, we classified Landsat/ETM+ data via a transformation of the original reflectance spectral space to the UPDM sub-space. Classification accuracy was compared using results from UPDM and the primary component transformation (PCT). Classification results for ETM+ data were also compared using several traditional classifiers. The UPDM and the PCT showed similar classification accuracy. The UPDM sub-space has definitive physical meanings. Classification results using UPDM are sensor-independent, which are very significant for comparison of results derived from different data.

Keywords: feature extraction; land cover classification; hyper-multi spectral data.

1. INTRODUCTION

Remotely sensed data supply a wealth of information to help monitoring changes in land cover. One approach to extract land-cover information from remotely sensed images is classification (Huang, *et al*. 2002), and various classification techniques using multi-hyper spectral data have been developed. These techniques including minimum distance classifier (MDC), maximum likelihood classifier (MLC), neural network classifier (NNC), and decision tree classifier (DTC). Classification accuracy is affected by training sample size and number of input variables (Huang, *et al*. 2002). As dataset dimensions increase, sufficient training sample data must be selected to yield reliable classification parameters. Feature extraction project data from the original feature space to a lower-dimensional subspace that has a more effective feature space (Hsu, *et al*., 1999). Principal components transformation (PCT) is a useful feature extraction method; however, components of the PCT have no physical meanings.

The universal pattern decomposition method (UPDM) is a sensor-independent method that is tailored for satellite data analysis (Zhang, *et al*. 2003, 2005). Sets of spectral reflectance measured by a sensor are transformed by the UPDM into three coefficients with three fixed spectral reflectance patterns. The spectral reflectance patterns are determined for a spectral region between 350 nm 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. The coefficients are "pattern decomposition coefficients."

Application of the UPDM to satellite reflectance datareduces the number of UPDM features from the original hyper-multi dimensional data. In contrast to PCT, UPDM components have definitive physical meaning. In the standard pattern, they are water, vegetation and soil. The classification operation performed using UPDM coefficients yields many benefits.

This paper reports on the application of UPDM to ETM+ data classification. Results are compared to classification accuracy from PCT using MDC (using minimum Euclidean distance and minimum Mahalanobis distance) and MLC algorithms.

2. METHODOLOGY

Fujiwara *et al*. developed a pattern decomposition method (PDM) for satellite data analysis (Fujiwara, et al., 1996), Muramatsu *et al*. studied the PDM for Landsat/MSS, TM data analysis (Muramatsu, et al., 2000), Daigo *et al*. applied the PDM for hyper-multi spectral data analysis (Daigo, et al., 2004), The UPDM decomposes reflectance values at each pixel into a linear sum of standard spectral patterns for water, vegetation, soil and any supplemental patterns using the following formula (Zhang, et al., 2003; 2005):

$$R_i \to C_w \cdot P_{iw} + C_v \cdot P_{iv} + C_s \cdot P_{is} \qquad (1)$$

where R_i is the reflectance of band i measured on the ground (or by satellite sensor)
C_w, C_v and C_s are the respective decomposition coefficients
P_{iw}, P_{iv}, and P_{is} are the standard spectral patterns of water, vegetation and soil normalized with respect to the properties of each sensor

The three standard spectral patterns as a continuous spectral function from 350 to 2500 nm are defined as follows (Zhang, et al., 2003; 2005):

$$\int \left|P_k(\lambda)\right| \mathrm{d}\lambda = \int \mathrm{d}\lambda \qquad (k = w, v, s) \qquad (2)$$

where $\int \mathrm{d}\lambda$ refers to integration of the total wavelength range, and the $P_k(\lambda)$ of the standard spectral pattern is defined as

* Corresponding author. E-mail: zhanglifu01@yahoo.com.cn

$$P_k(\lambda) = R_k(\lambda)\frac{\int \mathrm{d}\lambda}{\int \left|\mathrm{R}_k(\lambda)\right|\mathrm{d}\lambda} \tag{3}$$

where $R_k(\lambda)$ represents the spectral reflectance patterns of standard water, vegetation and soil. The shapes and magnitudes of the standard patterns $P_k(\lambda)$ are fixed for all sensors.

For each sensor band, we intercepted P_{ik} values. Thus, the standard patterns for each sensor are defined by

$$P_{ik} = \frac{\int_{si}^{ei} P_k(\lambda)\mathrm{d}\lambda}{\int_{si}^{ei}\mathrm{d}\lambda} \qquad (k = w, v, s) \tag{6}$$

where si and ei are the start and end wavelengths for band i, respectively, and $\int_{si}^{ei}\mathrm{d}\lambda$ is the wavelength width of band i. The decomposition coefficients C_k were obtained for each sensor by the least squares method using equation (1). In principle, nearly equal values should result for the same object; furthermore, coefficient precision is expected to improve as the number of bands increases.

Spectral reconstruction precision was evaluated using reduced Chi-square values that satisfied the expression (Zhang, *et al.*, 2003; 2005)

$$\chi^2 = \sum_{i=1}^{n} r(i)^2/(n-3) \tag{7}$$

Here, n is the number of bands, and r is the error of band i.

3. DATA PROCESSING

3.1 Data preprocessing

We used Landsat/ETM+ standard product data acquired over the Yangtze River in China on 24 December 1999 by the Beijing Remote Sensing Ground Station. Data were geo-referenced and had a spatial resolution of 28.5 m. Six spectral bands (bands 1–5, and 7) in the ETM image were converted to top-of-atmosphere reflectance using the common method of Joachim *et al*. (1991).

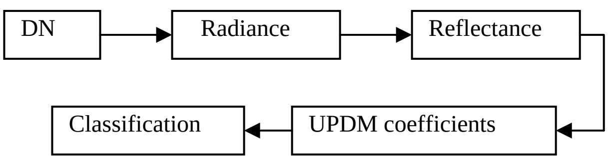

Figure 1. Flowchart of data processing

Only the Rayleigh scattering was corrected because the image was clear over the research area. Fig. 1 schematically shows the data processing method. Table 1 shows the Rayleigh scattering values for ETM+ data used in this study.

Table 1 Rayleigh scattering correction values for ETM+

Band	Wavelength nm	E_0 W/m^2/µm/sr	Rayleigh scattering
1	450.0~515.0	1969.0	0.055
2	525.0~605.0	1840.0	0.031
3	630.0~690.0	1551.0	0.017
4	775.0~900.0	1044.0	0.007
5	1550.0~1750.0	225.7	0.000
7	2090.0~2350.0	82.1	0.000

3.2 Feature extraction

A subspace projection was performed on the original reflectance multi-spectral space after ETM+ data were converted to reflectance data. A new UPDM feature was thus obtained using Eq. (1). Three normalized standard spectral patterns of water, vegetation, and soil in the 350nm-2500nm wavelength (excluding the vapor absorption wave portion) were computed. The intercepted standard spectral pattern for ETM+ sensor was (Zhang, *et al*. 2005):

$$\begin{bmatrix} 3.277077 & 0.175195 & 0.545911 \\ 2.672011 & 0.384025 & 0.786754 \\ 1.449789 & 0.171269 & 0.925836 \\ 0.817368 & 2.311455 & 0.979686 \\ 0.219794 & 0.961035 & 1.251477 \\ 0.205009 & 0.332513 & 1.164075 \end{bmatrix}$$

New UPDM coefficients (three UPDM features) can be computed using the least squared method on Eq. (1).

4. RESULTS AND DISCUSSION

4.1 Classification using PCT and UPDM components

The comparison of the classification efficiency and accuracy considers data derived from three PCT components and three

Figure 2. ETM+ reflectance image expressed by bands 6, 4, 1 in red, green, and blue

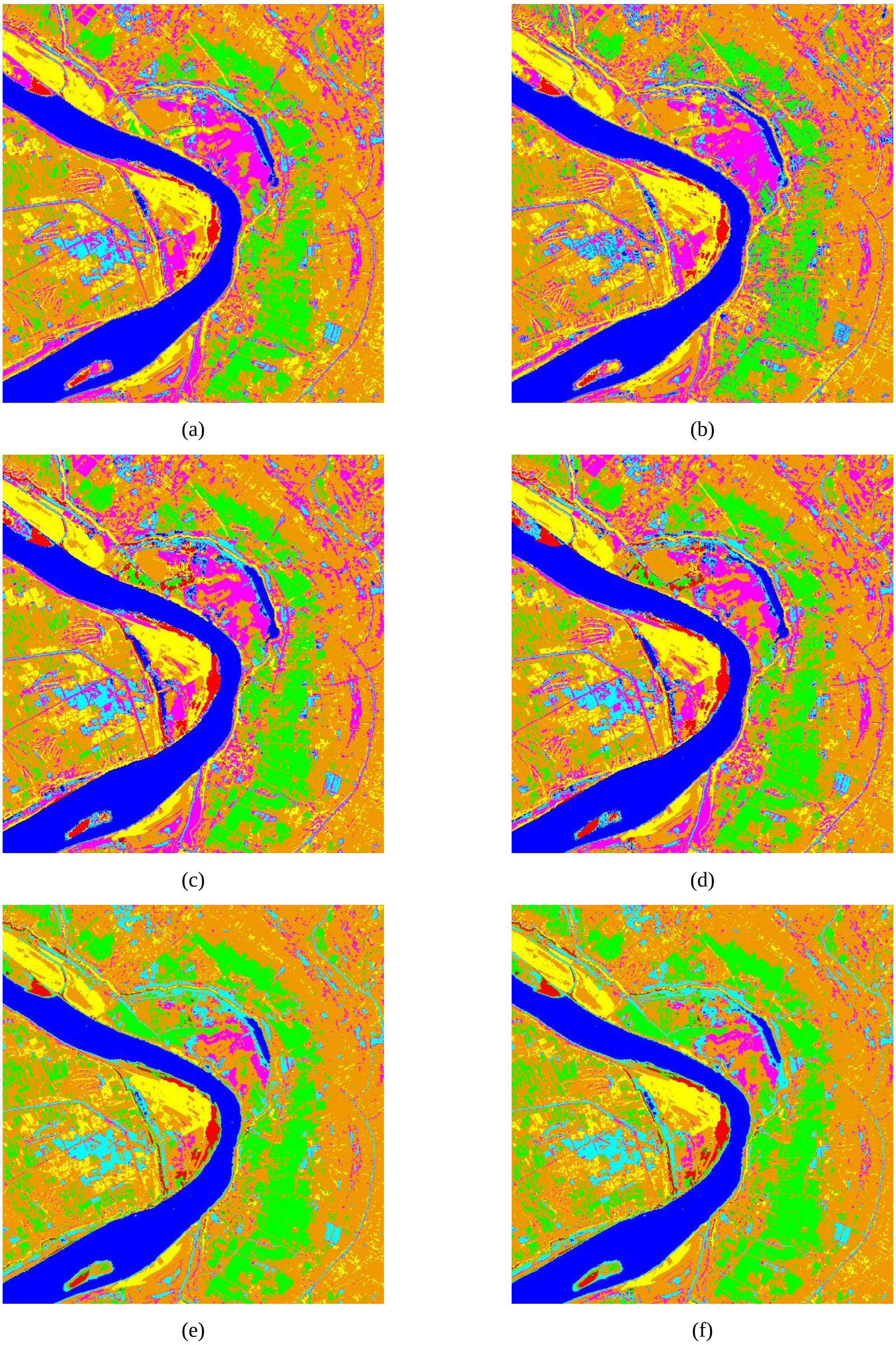

Figure 3. Classification results from different classifiers. Top row, (a) and (b): Minimum distance classifier; Middle row, (c) and (d): minimum Mahalanobis distance classifier; Bottom row, (e) and (f): maximum likelihood classifier. Results from three PCA components (left column) and from three UPDM components(right column).

UPDM components. Data arise from the same classifier, namely the minimum distance classifier, minimum Mahalanobis distance classifier, and maximum likelihood classifier. Fig. 2 shows the original reflectance image derived from bands 6, 4, 1(red, green, and blue, respectively). Fig. 3 shows results from three conventional classifiers using PCA and UPDM components. Classification accuracy from the maximum likelihood classifier is higher than the other two classifiers in this example.

4.2 Classification accuracy comparison

We used two measures of accuracy: overall accuracy and the kappa coefficients. As noted above, the maximum likelihood classifier was more accurate than the others in this study, which used only three components to define the classification. Reliable parameter can therefore be estimated from the training samples. We considered seven classes training sample data and 584 test sample data in the original reflectance imageusing the method discussed by Murat (2002). Table 2 compares classification accuracy. Results show that UPDM is suitable for feature extraction before classification and that UPDM and PCT have similar classification accuracy. More importantly, UPDM components have definite physical meanings.

Table 2. Classification accuracy

Data used	Overall accuracy	Kappa coefficients
Original reflectance	88.9%	0.87
PCT components	89.2%	0.87
UPDM components	88.5%	0.86

5. CONCLUSIONS

This paper considered ETM+ data for a classification study. The satellite digital signal number (DN) was first converted to reflectance value after adjusting for the influence of Rayleigh scattering. Subsequently, a PCT and UPDM transformation were performed. The classification used PCT and UPDM were similar. Unlike PCT components, UPDM components have physical meanings. Classification results using UPDM are sensor-independent (Zhang, *et al*. 2005), which are very significant for comparison of results derived from different data.

Previous studies using simulated ground-measured data have demonstrated the suitability of UPDM for hyper spectral data. Further research will use hyper spectral data, for example, 224-bands AVIRIS data.

6. REFERENCES

C. Huang, L. S. Davis, and J. R. G. Townshend, “An assessment of support vector machines for land cover classification”, International Journal of Remote Sensing, vol. 23, No. 4, 2002, pp.725-749.

D. M. Murat, and D. Landgrebe, “A model-based mixture-supervised classification approach in hyperspectral data analysis,” IEEE Transactions on Geoscience and Remote Sensing. 2002, vol. 40, No. 12, pp.2692–2699

K. Muramatsu, S. Furumi, N. Fujiwara, A. Hayashi, M. Daigo, and F. Ochiai, “Pattern decomposition method in the albedo space for Landsat TM and MSS data analysis”, International Journal of Remote Sensing, vol. 21, No. 1, 2000, pp. 99-119.

Lillesand, T. M., and R. W. Kiefer, Remote Sensing and Image Interpretation, 4nd edition, John Wiley & Sons Inc., New York, 2000.

L. 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 Publications, Wuhan University, China, 2003.

L. Zhang, L. Zhang, N. Fujiwara, and K. Murumatsu, “The universal pattern decomposition method based on the hyper spectral satellite remotely sensed data”, (in Chinese), Journal of Geomatics and information science of Wuhan University, vol.30, No.3, 2005, pp.264-268.

Pai-Hui Hsu, Yi-Hsing Tseng. Feature extraction for hyper spectral image. Asia remote sensing conference, 1999, Hongkong.

M. Daigo, A. Ono, N. Fujiwara, and R. Urabe, “Pattern decomposition method for hyper-multi-spectral data analysis”, International Journal of Remote Sensing, vol. 25, No. 6, 2004, pp.1153-1166.

N. Fujiwara, K. Muramatsu, S. Awa, A. Hazumi, and F. Ochiai, “Pattern expansion method for satellite data analysis”, (in Japanese), Journal of Remote Sensing Society of Japan, vol. 17, No. 3, 1996, pp.17-37.

P. M. Teillet, K. Staenz, and D. J. Williams, “Effects of Spectral, Spatial, and Radiometric Characteristics on Remote Sensing Vegetation Indices of Forested Regions,” Remote Sensing of Environment, vol.61, 1997, pp. 139-149.

ACKNOWLEDGMENTS

This work was supported under the ADEOS-II/GLI Project of the Japan Aerospace Exploration Agency (JAXA), the Academic Frontier Promotion Project of the Ministry of Education, Science, Sports, and Culture of Japan, the 973 Project of the People’s Republic of China (Project Number 2003CB415205), and the National Natural Science Foundation of China (Project Number 40471088).