Principal component-based algorithm on multispectral remote sensing data for spectral discrimination of tree cover from other vegetation types

Subhash Ashutosh
Forest Survey of India, Kaulagarh Road, Dehradun 248 195, India

Remote sensing technology offers an effective tool for mapping of forest resources in a cost-effective manner. The Forest Survey of India (FSI) carries out forest-cover mapping of the entire country once in two years, using remote sensing data. Delineation of forest-cover from contiguously occurring agriculture or bushy vegetation on remote sensing imagery is a problem, which is often faced in classification of remote sensing data by digital method. There are several indices, including NDVI for highlighting vegetation areas on the scene. However, these indices have not been found very effective in discriminating tree cover from other vegetation types. In this study, a new approach based on principal components has been tried for the above-stated purpose. A newly defined index has been found to yield good results in separating pixels having tree cover from the other types of smaller vegetation.

APPLICATION of remote sensing for sustainable management of natural resources has been widely demonstrated. Remote sensing technology offers a means for repetitive mapping of natural resources in a cost-effective manner. Monitoring of changes in the forest-cover by periodic mapping of the resource is an important application of the technology. The Forest Survey of India (FSI) carries out forest-cover mapping of the entire country in a cycle of two years, using remote sensing data. Area statistics along with analysis and forest-cover maps are published in the document known as 'State of Forests Report' (SFR), biennially. The latest report in the series, i.e. SFR–1999 has been released in September 2000.

A problem often encountered in forest-cover mapping is of delineation of tree cover from other types of vegetation like bushes, agriculture, etc., when they are occurring contiguously or in the vicinity. Though characteristic reflectance from tree cover and smaller vegetation like agriculture crops shows typical tones on the satellite imagery, yet often because of topographical and phenological changes and also due to occurrence of vegetation types of intermediate gradations, separation of tree vegetation from bushy or herbaceous vegetation becomes difficult. In the methodology followed by FSI, such areas are masked out by visual interpretation based on ground knowledge. This method of masking is not only time-consuming but often subjective.

e-mail: subhashashutosh@usa.net

Digital image processing of satellite data provides tools for analysing the image through different algorithms and mathematical indices. Based on reflectance characteristics, indices have been devised to highlight features of interest on the image. There are several indices for highlighting vegetation-bearing areas on a remote sensing scene. Normalized Difference Vegetation Index (NDVI) is a common and widely used index[1].

In the FSI's method of forest-cover mapping, NDVI is used for removing non-forest areas from the digital image and also to classify forest areas into different tree cover density classes, i.e. open forests and dense forests having crown density 0.1–0.4 and above 0.4, respectively. Ground truthing of classified scenes have shown that NDVI values for different types of vegetation are sometimes similar, which leads to error in classification, i.e. agriculture often get classified as forests and vice versa. However, NDVI gives good results when used for discriminating vegetation from non-vegetated areas in the scene.

With an objective of finding an approach which could provide better interpretability in terms of discriminating tree cover from other vegetation forms, principal component (PC) analysis of the IRS-1C LISS-III data was carried out (IRS-1C LISS-III sensor provides multispectral image in 4 bands, i.e. 0.52–0.59, 0.62–0.68, 0.97–0.86 and 1.55–1.75 μm. Spatial resolution of the first three bands is 23.5 m, whereas for the fourth band, it is 70.5 m). Pixel values of different PCs for different types of vegetation cover were analysed. The analysis of profile of pixel values shows distinct trends of PC1 and PC2 for different types of vegetation, viz. tree, agriculture, etc. and hence provides a basis for discriminating tree cover from agriculture crops and bushy vegetation cover.

PC analysis (often called PCA or Karhumen–Loeve analysis) has proved to be of value in the analysis of multispectral remotely-sensed data[2,3]. The transformation of the raw remote sensing data using PCA can result in new PC images that may be more interpretable than the original data[4].

To perform PC analysis, we apply a transformation to a correlated set of multispectral data. The application of the transformation to the correlated remote sensor data will result in another uncorrelated multispectral data set that has certain ordered variance properties[4].

Spread or variance of the distribution of points in the feature space gives an indication of the correlation and quality of information associated with the bands. If the data points are clustered in a small zone, it would probably provide very little information. The initial measurement coordinate axes may not be the best arrangement in multispectral feature space to analyse the remote sensing data. An approach could be to use PC analysis to translate and rotate the original axes, so that the original brightness values are redistributed (reprojected) onto a new set of axes[5]. Translation and rotation of axes in two-dimension (two bands), is shown in Figure 1.

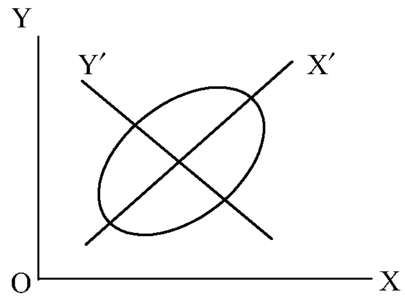

Figure 1. Translation and rotation of axis in 2D.

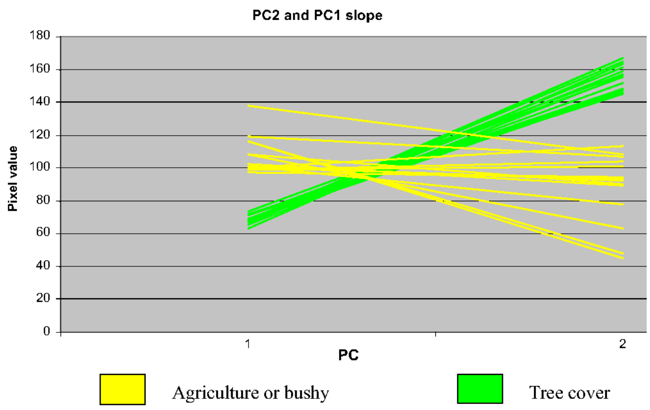

Figure 2. Profile of PC1 and PC2.

The X–Y coordinate system may be rotated about its new origin in the new coordinate system some φ degrees, so that the first axis X' is associated with the maximum variance in the scatter of points. The new axis is called first principal component (PC1). The second principal component is orthogonal to PC1 and contains lesser variance. The third, fourth, fifth and so on, components contain decreasing amounts of the variance found in the data set. Pixel values in the PC-transformed image represent transformed reflectance data set in which two channels (PC bands) are uncorrelated[6–8].

IRS LISS-III scene (path 96 and row 50) was geometrically rectified with the help of Survey of India (SOI) toposheets on 1 : 50,000 scale. A subset equivalent to SOI toposheet 53F/3, bounded by lat 77° and 77° 15′ and long 30° 15′ and 30° 30′ was extracted. This formed the study scene. Ground truth information on 28 points was collected for reference purpose. The points were selected in such a way that they represented tree cover, agriculture areas, bushy vegetation and area on fringes of forests and agriculture fields. Using K-means algorithm of unsupervised classification, non-vegetated area from the scene was masked out. This resulted in reduced feature space having pixels bearing vegetation cover only. PC transformation of the image was carried out to obtain the first two principal components, PC1 and PC2.

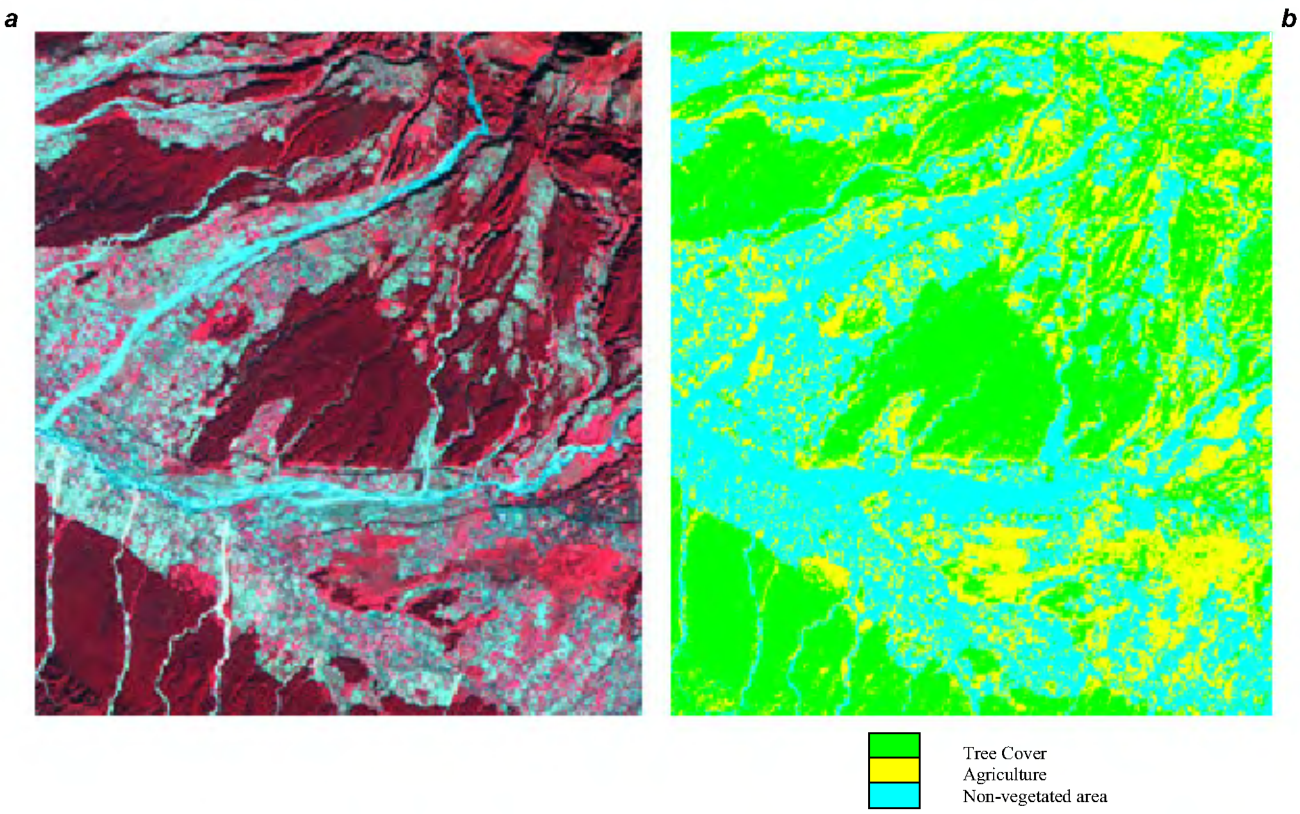

Figure 3. ***a***, FCC of part of the study scene, and ***b***, Classified image.

Table 1. Error matrix

Image classification	Ground truth			
	Tree cover	Agriculture	Non-vegetated areas	Total
Tree cover	14	1	0	15
Agriculture	0	10	0	10
Non-vegetated areas	0	0	5	5
Total	14	11	5	30

An image comprising a stack of 5 layers; 3 bands of LISS-III image and the 2 PCs was formed. Spectral profile of the 5-layer image was analysed on the points for which ground truth information was gathered. On the basis of trend, an index was defined and the masked image was classified into tree cover and agriculture areas. Ground validation was done on 30 points for accuracy estimation.

The profile of the pixel values of 5-layer image was analysed on the reference points. It was observed that the slope between PC2 and PC1 layers is quite distinct for the two vegetation types, i.e. tree cover and agriculture. Figure 2 shows the profile of the two PCs. The trend of the slope for the two vegetation types suggests that the ratio of PC2 and PC1 may be an effective index to separate the two vegetation types. An algorithm to compute the ratio of the two PCs for the image was written in 'spatial modeler' and applied on the scene. The output was stretched on 0 to 255 scale. Analysis of the ratio on the ground truth points revealed that the pixels for which the ratio was greater than a certain threshold (1.2 in the present case) pertained to tree cover. The image was classified into tree cover, agriculture bushy area and non-vegetated areas using the above algorithm. The classified image was taken to the field for ground validation. Out of 30 points on which the classification was checked, 29 were found in accordance with the classification, showing an accuracy of over 96%. The false colour composite and the classified scenes are shown as Figure 3.

The index may be defined as PC2/PC1, and termed 'Principal Component Vegetation Index (PCVI)'.

The index shows higher values for the tree cover. In other words, the slope between PC2 and PC1 is higher in case of pixels having tree cover than those having other types of vegetation. Trial of the index on other scenes has shown consistency in the results, though threshold value of the index separating tree cover from agriculture varies from scene to scene.

Ground validation result based on a small sample as shown in Table 1 indicates high accuracy in delineating tree cover from other forms of vegetation. The study is being continued further to examine and analyse applicability of the index, to diverse scenes and with greater intensity of ground validation.

1. Rouse *et al.*, Proceedings of the Eighth International Symposium on Computer-Assisted Cartography, Baltimore, MD, 1973.
2. Press *et al.*, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 1992.
3. Wang, L., *Photogramm. Engg. Remote Sensing*, 1990, **56**.
4. Singh, N. and Harrison, A., *Int. J. Remote Sensing*, 1985, **6**, 883–896.
5. Jensen, J. R., *Digital Image Processing: A Remote Sensing Perspective*, 1986.
6. Campbell, J. B., *Introduction to Remote Sensing*, 1987.
7. Fung, T. and Le Drew, E., *PE&RS*, 1987, **53**, 1649–1658.
8. Howard, J. A., *Remote Sensing of Forest Resources*, 1991.

ACKNOWLEDGEMENTS. I acknowledge support and encouragement given by the Director and the Joint Director, Forest Survey of India and the Dean, Indian Institute of Remote Sensing.

Received 11 June 2001; revised accepted 1 September 2001