International Journal of Advanced Science and Research
ISSN: 2455-4227
Impact Factor: RJIF 5.12
www.allsciencejournal.com
Volume 2; Issue 6; November 2017; Page No. 156-160

Digital image processing algorithms for classification, accuracy assessment and change detection of earth's resources

Dr. Madan Mohan
Associate Professor of Geography, Centre for Study of Regional Development School of Social Sciences – III, Jawaharlal Nehru University, New Delhi, India

Abstract
Digital Image Processing (DIP) is performed using sophisticated computer algorithms on satellite's raw imagery. The main processes are as image rectification and restoration i.e. geometric and radiometric corrections; image enhancement i.e. effective display for interpretation; and geospatial information extraction using digital classification and thematic geovisualisation. In other words, satellite's digital imagery requires algorithmic processing to achieve better outcome in real time for real world. The usage of specific algorithms depends on individual's goals for specific purpose. So, digital processing of imagery is based on complex algorithms processed in high-end sophisticated processor for simple tangible tasks. As a result, processed remote sensing satellite imagery have been more popularity in earth's resources exploration and management for sustainable regional development.

Keywords: raw imagery**,** algorithms, feature selection, accuracy assessment, kappa coefficient, change detection, geovisualisation

Introduction: Satellites Imagery

IRS-1C and IRS-P6 imagery are used for many applications like urban mapping, mineral prospecting, forest resource survey, wasteland mapping etc. which have become an integral part of the resources management system. The popularity of satellite based remote sensing has created a need to providing data with better resolution, coverage and revisit. IRS-1C and IRS-P6 were conceived to meet these demands.

The IRS-1C is a three axis body stabilized remote sensing satellite. It is operational in a polar, sun synchronous orbit at a height of 817 kms. altitude with equatorial crossing time of 10:30 a.m., in descending node. The satellite repeativity is 341 orbits/ 24 days. The satellite has capability of revisit in 5 days (NRSA, 1998). Similar are the characteristics for the IRS-P6 satellite. The satellite payload consists of three sensors the Panchromatic Camera (PAN), Linear Imaging and Self Scanning Sensor (LISS) and Wide Field Sensor (WiFS).

The Linear Imaging and Self Scanning Sensor (LISS) sensor provides multi-spectral data collected in four bands as the visible (V), near infra-red (NIR) and short wave infra-red (SWIR) regions. While the spectral resolution and swath in the case of visible (two bands) and NIR (one band) regions is 23.5 m and 141 km respectively. And, there is 70.5 m and 148 km for the data collected in SWIR region. The LISS sensor operates between the 0.52 - 1.70 microns spectral bands. The details of the sensor and it characteristics are presented in the Table 1.

Digital Processing of Raw Imagery

A number of standard image processing techniques in ERDAS Imagine software, such as image sub-set, correction, rectification, restoration, and classification are available for the analysis of imagery. The radiometric correction applied to ensure that the imagery have common reference spectral characteristics. Imagery is further georeferenced and geometrically corrected corresponding to the polyconic projection system.

Satellite imagery and their respective range of reflectance, digital number (DN) values can also be studied thoroughly to ascertain the probable land use classes. Whereas, spectral profiles have been studied to ascertain the separability and relative difference in pixel values of various land use classes in several spectral bands (Goward *et al.*, 2006). For instance, seven separable land use classes have been identified, such as compact settlement, sparse settlement, open space (settlement), green cover vegetation, scattered vegetation, agricultural land and Yamuna River (water) for NOIDA area. The Maximum Likelihood Classification (MLC) based on supervised classification algorithm has been used in computation for the LULC classification of the imagery.

Table 1: Feature and Characteristics of LISS Sensor.

Sensor	Bands		Spectral Regions (in microns)	Spatial Resolution	Ground Swath (kms.)	Encoding	SNR
LISS	Band2	VIS	0.52 - 0.59 μ	23.5 m	141 km	7 bits	> 128
	Band3	VIS	0.62 - 0.68 μ	23.5 m	141 km	7 bits	
	Band4	NIR	0.77 - 0.86 μ	23.5 m	141 km	7 bits	
	Band5	SWIR	1.55 - 1.70 μ	70.5 m	148 km	7 bits	

Note: SNR stands for Signal to Noise Ratio (at saturation radiance).
Source: NRSA (1988) *IRS-1C Data Users Handbook*, National Remote Sensing Agency, Department of Space, Government of India, Hyderabad, India.

The algorithm was trained by a supervised training process by taking a number of training samples. Each training sample consisted of at least 90 image pixels to satisfy the 10 *n* criteria, where *n* is the number of bands used for classification (Congalton, 1991) [9]. Signatures are further evaluated using three criteria to test whether these truly represent pixels to be classified for each class as the (i) histogram plots to examine various statistical parameters, like standard deviation and unimodality of the histogram; (ii) signatures separability using Average Divergence (AD) and Transformed Divergence (TD); and (iii) contingency matrix, which contains the number and percentage of pixels which are classified as expected. Signatures are refined, deleted, renamed and merged to ensure the unimodality of their histograms, statistical parameters, contingency matrix and AD and TD values.
After evaluating the separability, spectral band combinations with good separability (highest TD value) have been selected for final classification. Initially, the Okhla Barrage nearby area was classified as a separate class, however after field verification it has been merged with the River Yamuna. Ancillary information from various sources (topographic sheets and village boundary map) have been integrated with output which is further refined from the ground truth data collected from the field visits. Classified images have been validated using the ground truth data and available topographic sheets and maps from various agencies for NOIDA. Results have been found satisfactory for land use land cover classification for 2001 and 2006.
Classification accuracy of the results has been assessed using a reference data set about 256 randomly selected pixels. Land use for these pixels have been determined using an urban settlement map (prepared from the topographical survey carried out in the year 1991), and data collected from other maps (village boundary map, Census Atlas, Village and Town Directory of the Census of India and SOI topographic sheets). The original satellite images have also been used for accuracy assessment to avoid errors in the reference dataset for temporally sensitive classes such as vegetation. A settlement map of the NOIDA city and geographical locations of some important features, like type of vegetation at a particular location, important settlements, water bodies and river channels, flyover, toll bridge etc. collected during the field visits have also been used as ground truth data. Furthermore, an accuracy report and Kappa Coefficient have been generated using the ERDAS Imagine's accuracy assessment utility. Urban growth over a period of three decades (1976-2006) has been determined from the topographic sheets and classified images, and results are compared with the settlement maps prepared by NOIDA Development Authority.

Feature Selection
The feature selection process is used to derive useful information from multispectral imagery. The feature selection schemes play an important role to select those dimensions most suitable for processing without unnecessarily increasing computation time. This process is known as feature selection or separability analysis. There are various graphical and statistical methods of feature selection schemes which are used depending upon the classes and their statistics. The graphic feature selection method of feature space plots in two dimensions depicts the distribution of all the pixels in the scene using two bands at a time has been used to select the most appropriate bands for the present study (Schowengerdt, 1998). In addition to this, the statistical feature selection method the divergence analysis has been worked out by variance covariance methods which are mentioned as follows:

Average Divergence

$$Diver_{avg} = \frac{\sum_{c=1}^{m-1} \sum_{d=c+1}^{m} Diver_{cd}}{C} \quad (1)$$

Transformed Divergence

$$T\,Diver_{cd} = 2000\left[1 - \exp\left(\frac{-Diver_{cd}}{8}\right)\right] \quad (2)$$

These methods have been used to compute statistics and then calculated the relationship in form of the correlated and uncorrelated bands. The correlation of data among various bands has also been estimated graphically from the scatter plot of the data between two bands. In the process of feature selection, the highly correlated bands were rejected and those with lesser or no correlation selected for efficient land use land cover analysis for the study area.

Maximum likelihood classification algorithm

The maximum likelihood decision rule assigns each pixel having pattern measurement of features X to the class c whose units are most probable or likely to have given rise to feature vector X. It assumes that the training data statistics for each class in each band are normally distributed. In other words, training data with bi-modal or tri-modal histogram in a single band are not ideal. In such cases the individual mode probably represents individual classes that should be trained upon individually and labeled as separate classes. This would then produce unimodal, Gaussian training class statistics that would fulfill the normal distribution requirement. The decision rule applied to the unknown measurement vector X is computed using the following equation (Jenson, 1996):

Decide X is in class c if, and only if,

$$P_c \geq P_i \,, \tag{3}$$

Where:

$i = 1, 2, 3, \ldots, m$ possible classes

and

$$P_c = \{ -0.5 \log_e [\det (V_c)]\} - [0.5 (X - M_c)^T V_c^{-1} (X - M_c)\} \tag{4}$$

And det (V_c) is the determination of the covariance matrix V_c. Therefore, to classify the measurement vector X of an unknown pixel into a class, the maximum likelihood decision rule computes the value P_c for each class. Then it assigns the pixel to the class that has the largest (or maximum) value.

Generally, in most remote sensing applications the thumb rule is that there is high probability of encountering some classes more often than others. It is possible to include this value a priori information in the classification decision. So, it can be done by weighing each class c by its appropriate a priori probability. The equation then becomes as given below:

Decide X is in class c, if and only if.

$$P_c (a_c) \geq P_i (a_i), \tag{5}$$

Where:

$i = 1, 2, 3, \ldots, m$ possible classes

and

$$P_c (a_c) = \log_e (a_c) - \{0.5 \log_e [\det (V_c)]\} - [\, 0.5 (X - M_c)^T (V_c^{-1})(X - M_c)] \tag{6}$$

Therefore, the maximum likelihood classification of remotely sensed data involves considerable computational effort - because it calculates a large amount of information on the class membership characterization of each pixel.

Accuracy Assessment

The accuracy assessments for land use land cover classification require to prepare the classification error matrix which is also known as confusion or contingency table. It helps to compare the relationship between known reference data (ground truth) and the corresponding results of an automated classification. It is a square matrices with number of rows and columns. So, the image analyst has to prepare an error matrix to determine accuracy of classification that has categories for representative subset of pixels used in training process of a supervised classification. The training set pixels are located along major diagonal of the error matrix. Omission errors correspond to non-diagonal column elements. The commission errors are represented by non-diagonal row elements. The overall accuracy of classification is determined by dividing the total number correctly classified pixels by the total number of reference pixels. The producer's accuracy indicates that how well the training set pixels of a given cover types are classified. It is determined by dividing the number of correctly classified pixels in each category by number of training sets used for that category (column total). On the other hand, the user's accuracy is computed by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified in that category (row total) (Congalton, 1991) [9].

Accuracy assessment at the pixel level is important to compare every pixel in an image with reference source. Such comparison has value in research but it is expensive. Another consideration is making certain that the randomly selected test pixels of areas are geographically representative of the data set under analysis. Stratified random sample is presently used in such cases. Consideration must also be given to the sample unit employed in accuracy assessment. Broadly, it is suggested that a minimum of 50 samples of each land use category, can be included in the error matrix. However, based on the importance of particular, the number of samples in that class can be adjusted accordingly. By and large for quantative assessment of classification accuracy requires the collection of the prior knowledge of some parts of terrain which is compared with the remote sensing derived classification land use map. So, to correctly perform classification accuracy assessment, it is necessary to compare two sources of information such as (i) remote-sensing-derived classification map; and (ii) reference test information (which may in fact contain error). The relationship between these two sets of information is commonly summarized in an error matrix. These procedures allow quantative evaluation of the classification accuracy. Their proper use enhances the credibility, of using remote sensing derived land use land cover information (Foody, 2002) [14].

Kappa Coefficient

The statistical approach of the accuracy assessment consists of different multivariate statistical analyses. A commonly used measure is Kappa coefficient. In other words, the KAPPA coefficient analysis is a discrete multivariate technique used for accuracy assessment. KAPPA analysis yields a *K*hat statistic (an estimate of KAPPA) that is a measure of agreement or accuracy. The *K*hat statistic is computed using the given equation (Jension, 1996):

$$K_{\text{hat}} = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} * x_{+i})} \qquad (7)$$

Where:

r	=	is the number of rows in the matrix,
x_{ii}	=	is the number of observation in row i and column i, and
x_{i+} and x_{+i}	=	are the marginal totals for row i and column i, respectively'
N	=	is the total number of observations

Digital change detection algorithm

Remote Sensing techniques are effective in monitoring manmade features on the surface of earth. Manmade features are dynamic and constantly changing over time and space. So, it is noteworthy to monitor accurately the physical and human processes by using change detection methods based on remotely sensed data. It is found that the digital change detection is affected by spatial, spectral, temporal and thematic constraints (Howarth and Wickware, 1981) [19]. Therefore, the selection of appropriate method is having considerable significance. By and large, the most digital change detection methods are based on per-pixel classifiers and change information contained in the spectral radiometric domain of the images (Foody and Dan, 2007; Fung and Le Drew, 1987) [15, 16]. However, there have been developed a large number of digital change detection algorithms during the last quarter of the 20th Century. The most commonly used and empirically validated digital change detection algorithms are mentioned as follows:

1. Change Detection Using Write Function Memory Insertion

   It is analog method for qualitatively assessing the changes in a region and do not provide quantities information of the changes took place in a region.

2. Multi-Date Composite Image Change Detection

   The multiple data sets or imagery are registered to a single database. Such composite data sets are used to extract information using unsupervised classification technique resulting into a class of change and no change.

3. Image Algebra Change Detection

It is possible to simply identify the amount of change between two images by band rationing or image differencing the same band in two images that have previously been rectified to a common base map. Image differencing involves subtracting the imagery of one date from that of another. The subtraction results in positive and negative values in areas of radiance change and zero values in areas of no change in a new change image. In an 8-bit analysis with pixel values ranging from 0 to 255, the potential range of difference values is -255 to 255. The results are normally transformed into positive values by adding a constant, *c* (e.g., 127) (Jenson, 1996). The computation operation is expressed mathematically as follows:

$$D_{ijk} = BV_{ijk}(1) - BV_{ijk}(2) + c \qquad (8)$$

Where,

D_{ijk}	= change pixel value
$BV_{ijk}(1)$	= brightness value at time 1
$BV_{ijk}(2)$	= brightness value at time 2
c	= a constat (e.g., 127)
i	= line number
j	= column number
k	= a single band (e.g. LISS-III band 4)

However, the change image produced using image differencing usually yields a brightness value (BV) distribution approximately in Gaussian nature. Otherwise, the pixels of no BV change are distributed around the mean and pixels of change are found in the tails of the distribution. Band rationing involves exactly the same logic, except a ratio is computed and the pixels that did not change have a ratio value of 1 in the change image (Jensen, 1996). So, in the present study a specific attention has been given to review different methodologies for digital change detection as has been mentioned and discussed above in detail. However, among the above mentioned algorithms the image algebra change detection has been empirically applied and largely validated for digital change detection. Therefore, this algorithm has been used in the present study for LULC change detection analysis.

Conclusions

Digital Image processing is based on some computer based algorithm operations on raw satellite imagery for enhanced of imagery and to extract useful geospatial information. In other words, it is a type of digital signal processing of raw imagery as an input and output imagery provides enhanced characteristics/features available on processed. At present, image processing methods and techniques are rapidly growing for better and precise results. So, digital image processing essentially includes three main steps as importing raw imagery using image acquisition tools; secondly, it is used for analyses and manipulating of information; and finally, the result or output which is based on processed imagery for specific purpose of real world.

References

1. Aronoff S. Classification Accuracy: A User Approach, Photogrammetric Engineering and Remote Sensing, 1982; 48:1299-1307.
2. Bernstein R, Ferneyhough DG. Digital Image

Processing, Photogrammetric Engineering, 1975; 41:1465-1476.
3. Blom RG, Daily M. Radar Image Processing for Rock Type Discrimination, IEEE Transactions on Geoscience and Remote Sensing, 1982; 20:343-351.
4. Buiten HJ, Clevers JGPW. Land Observation by Remote Sensing: Theory and Applications, in Gorden and Breach, Current Topics in Remote Sensing, 1993; 3:32-39.
5. Cakir HI, *et al*. Correspondence Analysis for Detecting Land Cover Change, Remote Sensing of Environment, 2006; 102(3-4):306-317.
6. Campbell JB. Introduction to Remote Sensing, London, Taylor and Francis, 2007; 280-281.
7. Camps-Valls G, Tuia D, Gómez-Chova L, Jiménez S, Malo J. Remote Sensing Image Processing, San Rafael, CA, Morgan and Claypool, 2011.
8. Castleman KR. Digital Image Processing, New Jersey, Prentice-Hall, Englewood Cliffs, 1977.
9. Congalton RG. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data, Remote Sensing of Environment, 1991; 37:35-46.
10. Cracknell Arthur P, Ladson Hayes. Introduction to Remote Sensing, London and New York, CRC Press, Boca Raton, 2009.
11. Defries RS, Townshend JRG. NDVI-derived Land Cover Classification at a Global Scale, International Journal of Remote Sensing, 1994; 15(17):3567-3586.
12. Fitzgerald RW, Lees BG. Assessing the Classification Accuracy of Multisource Remote Sensing Data, Remote Sensing of Environment, 1994; 47:362-368.
13. Foody GM, Campbell NA, Trodd NM, Wood TF. Derivation and Applications of Probabilistic Measures of Class Membership From The Maximum Likelihood Classification, Photogrammetric Engineering and Remote Sensing, 1992; 58:1335-1341.
14. Foody GM. Status of Land Cover Classification Accuracy Assessment, Remote Sensing of Environment, 2002; 80(1):185-201.
15. Foody Giles M, Dan HTX. Variability in Soft Classification Prediction and its Implication for Sub-Pixel Scale Change Detection and Super Resolution Mapping, Photogrammetric Engineering & Remote Sensing, 2007; 73(8):923-933.
16. Fung T, LeDrew E. Application of Principal Component Analysis for Change Detection, Photogrammetric Engineering & Remote Sensing, 1987; 53(12):1649-1658.
17. Gonzalez RC, Wintz P. Digital Image Processing, Addison-Wesley Publishing Co., Reading, Mass, 1977.
18. Green K, Kempka D, Lackey L. Using Remote Sensing to Detect and Monitor Land-Cover and Land-Use Change, Photogrammetric Engineering and Remote Sensing, 1994; 60:331-337.
19. Howarth PJ, Wickware GM. Procedures for Change Detection using Landsat Digital Data, International Journal of Remote Sensing, 1981; 2(3):277-291.
20. Jager G, Benz U. Measures of Classification Accuracy Based on Fuzzy Similarity, IEEE Transactions on Geoscience and Remote Sensing, 2000; 38:1462-1467.
21. Jensen JR. Introductory Digital Image Processing: A Remote Sensing Perspective, Upper Saddle River, NY, Prentice Hall, 2005.
22. Jensen JR. Introductory Digital Image Processing: A Remote Sensing Perspective, Upper Saddle River, NY, Prentice Hall, 2005.
23. Jensen JR. Remote Sensing of the Environment: An Earth Resource Perspective, Delhi, Dorling Kindersley India Pvt. Ltd., Pearson Education, Inc., IInd Edition, 2009.
24. Joshi PK, *et al*. Vegetation Cover Mapping in India using Multi-Temporal IRS Wide Field Sensor WiFS Data, Remote Sensing of Environment, 2006a; 103(2):190-202.
25. Lillisand TM, Kiefer RW, Chipman JW. Remote Sensing and Image Interpretation, New Delhi, Wiley-INDIA John Wiley & Sons Asia Pte, Ltd., Fifth Edition, 2004.
26. Liu Jian Guo, Philippa J Mason. Essential Image Processing and GIS for Remote Sensing, United Kingdom, Wiley-Blackwell A John Wiley & Sons, Ltd., Publication, 2009.
27. Lu D, Weng Q. Urban Classification Using Full Spectral Information of Landsat ETM+ Imagery in Marion County, Indiana, Photogrammetric Engineering and Remote Sensing, 2005; 71(11):1275-1284.
28. Pratt WK. Digital Image Processing, New York, John Wiley & Sons, 1977.
29. Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins Digital Image Processing Using Matlab, USA, Pearson Education, 2007.
30. Reddy M, Anji, Hari Shankar Y. Textbook of Digital Image Processing, Hyderabad, BS Publications, 2009.
31. Richards JA, Jia X. Remote Sensing Digital Image Analysis: An Introduction, New York, Berlin, Heidelberg, Germany, Springer-Verlag, 2006.
32. Sabins Floyd F. Remote Sensing: Principles and Interpretation, Kolkata, Levant Books, 2013.
33. Schowengerdt Robert A. Techniques for Image Processing and Classification in Remote Sensing, USA, Academic Press, 1998.
34. Schowengerdt Robert A. Remote Sensing: Models and Methods of Image Processing, USA, Academic Press An Imprint of Elsevier, 2009.
35. Schowengerdt Robert A. Remote Sensing: Models and Methods of Image Processing, USA, Academic Press An Imprint of Elsevier, 2009.
36. Stefanov WL, *et al*. Monitoring Urban Land Cover Change: An Expert System Approach to Land Cover Classification of Semiarid to Arid Urban Centres, Remote Sensing of Environment, 2001; 77:173-185.
37. Tucker CJ. Maximum Normalized Difference Vegetation Index Images for Sub-Saharan Africa for 1983-1985, International Journal of Remote Sensing, 1986; 7(15):1383-1384.
38. Welch R. Spatial Resolution Requirements for Urban Studies, International Journal of Remote Sensing, 1982; 3(2):139-146.