The Expanded PDCS Digital Image Processing System (EPDCS)

by
G. H. Xu
Chinese Academy of Forestry
Wan Shou Shan, Peking, China

Abstract

This paper is a research report describing the Computer Software System EPDCS, developed for processing remote sensing imagery data. The programme system has grown out of the PDCS programmes which the author developed and implemented in Sweden.

The system includes the following groups of programmes:

- -Tape handling and subscene extraction
- -Statistics and histogramme print
- -Image enhancement
- -Training area extraction, calculation and display
- -Supervised classification
- -Unsupervised classification
- -Cleaning-up of a classified map
- -Accuracy evaluation
- -Result output

In this paper, discussions are made about the techniques for the preprocessing, enhancement and classification of MSS data, with greater elaboration on the classification.

The system has been used in land use mapping and environmental monitoring. The results have shown some promise of usability in practical work and will be reported in another paper.

INTRODUCTION

One application of remote sensing is the use of Multi Spectral Scanner (MSS) data for land use inventory over larg areas. This report discusses the Expanded PDCS programme system developed for processing remotely sensed MSS data by the computer systems, both at the Department of Physical Geography, University of Stockholm and at the Chinese Academy of Forestry.

The programme system consists of a series of computer routines for image data handling, manipulation and classification, with the following two points considered in the course of the development of the programme package:

1. The programme system should be suitable for research work.
2. It should be possible for a user to run the programmes even without prior experience of advanced computer programmes.

Proceeding from these points, on the one hand, the EPDCS system was designed to derive different methods for dealing with the same subject. For instance, there are several methods for an image enhancement, training area extraction, supervised classification and unsepervised classification. That is useful either for selecting optimum method or for determining suitable condition of the different methods. On the other hand, the programme system was written

for conversational processing. The user only need to answer simple questions like yes or no or to type out some parameters like line and column numbers.

The PDCS package was tested by using data from the Bjärka-Säby area which is situated 200 km south of Linkoping, Sweden. The data were collected on three separate occasions. The system was also used in other research projects both at the Department of Physical Geography, University of Stockholm and at the Chinese Academy of Forestry.

I got very much help from my Swedish friends when I developed this system as Chinese guest research worker in Sweden. The programme should be considered a product of the friendship between China and Sweden. Hence the system is named PDCS, meaning Programmes Dedicated to China and Sweden.

THE CHARACTERISTICS

The whole system indudes the following groups of programmes:

-Tape handling and subscene extraction
-Statistics and histogramme print
-Image enhancement
-Training area extraction, calculation and display
-Supervised classification
-Unsupervised classification
-Cleaning-up of classified map
-Accuracy evaluation
-Result out put

1. Tape handling and subscene extraction

The programmes are used when handling original landsat computer compatible tapes Produced by EROS data center with different data formats. The landsat scene is divided into either four or two Strips. The programmes can read and dump any subscene using only one tape station and without temporary memory. They can also read the information including ID record. Header record. Ancillary record and Annotation record, and print it on the line-printer.

2. Statistics and histogramme print.

The progamme gives the user some initial information about the subscene under study.

The statistics calculated by this programme include mean and standard deviation of brightness levels of different bands, covariance and correlation matrices.

The histogramme shows absolute and relative frequency and accumulated frequency of brightness levels.

Both of statictics and histogramme are printed on the line-printer.

3. Image enhancement

Image enhancement is an important task of image processing. EPDCS includes the programmes used for a contrast enhancement, density slicing, combinations of images and edge enhancement.

One type of contrast enhancement is made by linear transformation of the grey scale. Another type of contrast enhancement is performed by using equal probability quantizer. In the EPDCS system the third method is used. It is mid- way between the first two methods I.e. on one hand, it is better than the first one to show

the most informative portion of the scale. On the other hand it is better than the second one to express pixels of which the brightness levels are far away from the most informative portion.

With this programme package, the user can also select different scales for every brightness level, based on the printed histogramme. Hence, the user can print out the pixels of specified brightness levels, showing the other pixels as blank. Density slicing is another type of contrast enhancement technique in the system. identification of objects in the multilevel density slicing picture is better than the original picture.

Another EPDCS programme used for an image enhancement is combinations of images, like additive, subtractive and ratio images.

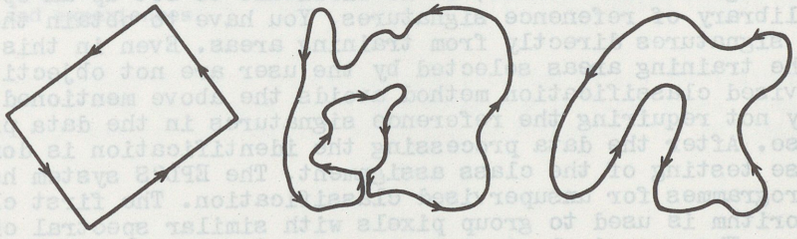
At last EPDCS programme can make edge enhancement by filter.

4. Training area extraction, calculation and display.

The computer analysis of remotely sensed data is often considered to consist of two parts: feature extraction and classification. The first process concerns the problem of selecting parts of the observed area so that classes are efficently described. These areas have to be extracted and analyzed so that they can be used as training areas. This part of programmes are used for this. They can perform the following tasks:

A. Training area extraction. The programme can deal with up to 30 training areas at the same time by means of three different methods:

1/ Line and column number of training area.
2/ Data from bounderies which have been digitized and transformed.
The digitization of training area boundaries is also performed by this programme. It possible to digitize the exact boundary of training areas. The data must be entered along the boundary in an anticlockwise direction. If you enter less than four points of one training area, its boundaries must be constructed of straight lines (cf figure).



Figure, Entering of boundary DATA of training areas. Note the anticlockwise direction.

The boundary transformation from the local coordinate system to the corresponding line and column number of the landsat subsence, is performed by the programme which will make an affine transformation.

3/ Data from a mask which represents training areas by means of different thematic codes.

B. The histogramme and the symbol coded picture of the training area. The programme can display them either on the terminal screen or on the line-printer.

C. When given thresholds, the pixels, with digital levels exceeding the thresholds, are deleted from the symbolic picture and will not be included in the calculation of the statistics of the training area.

D. Calculating statistical values as averages, standard deviations, covarience matrices and correlation matrices of the training areas.

At last, the programme will give the user the possibility to combine the statistical values of the training areas extracted from either optional files or the terminal, and to delete some statistical values which the user does not require.

the user can store all the statistical values of the training areas onto a disk file.

5. Supervised Classification

Classification is based on various supervised classification techniques that require refrence signatures of targets represented by training areas on the ground.

The EPDCS system provides the user with three different supervised classificution methods-maximum likelihood classification, minimum distance classification and table look-up classification. A user can select different training areas, channels and thresholds for the classification, using data stored on the same file. The classified image on the paper with a symbolic code representing different classes are produced by line-printer. Classified colour or black-white image on the film will be produced by rotating drum recorder.

6. Un-Supervised classification

The supervised technique is associated with high variability of spectral signatures. When so, it is difficult to set up an operational library of refenence signatures. You have to obtain the reference signatures directly from training areas. Even in this practice, the training areas selected by the user are not objective. The unsupervised classification method avoids the above mentioned difficulty by not requiring the reference signatures in the data procesing phase. After the data processing the identification is done by pixelwise testing of the class assignment. The EPDCS system has three programmes for unsupervised classification. The first clustering algorithm is used to group pixels with similar spectral characteristics. The second clustering algorithm is used to group pixels by means of the local textural parameter which was proposed by H. BEGUIN, H.Do TU and J. WILMET. The third algorithm also consists in combining spectral information with spatial information. Firstly the image is separated into the "small regions", pixels of which must be neighbours and must have very similar spectral values. Then, the unsupervised classification based on the "small regions" is performed according to spectral mean value and deviation of the regions

7. Cleaning-up classified images

The result from a supervised classification and some unsupervised classifications are pixelwise classified images using multispectral information only. This can result in a classified image having single or isolated pixels with a class assignment that differs from those of the surrounding. In many cases, such detailed information is not required, instead a generalised classified image is more useful. Such images may be obtained by comparing the class assignment of one pixel to its neighbours. The EPDCS system for the processing of a digital thematic map will eliminate all regions within an area less than a predefined minimum. Regions which represent different features can have different minima.

8. Accuracy evaluation

In order to compare the landsat classified image with ground truth, we developed this programme which transfer data between image and map by ground control points, then we can calculate the following parameters:

A. Difference map.

The map will show the pixels, which have been correctly classified, by means of an unique symbol. It will also show the pixels, which have been incorrectly classified, by means of different symbols representing the change of class.

B. Confusion matrix.

The table shows the amount of changed pixels of different classes.

C. Over all classification accuracy.

D. Mapping accuracy of every class.

E. Over all mapping accuracy.

ACKNOWLEDGE

I am very grateful to professor Leit Wastenson, Dr. Walter Arnberg, Mr. Göran Alm and Miss Laine Boresjo at the Department of Physical Geography, University of Stockholm, professor Juri Talts at LMV for their very useful ideas and experiences.

References

1. Valeric. L. Thomas, Generation and Physical Characterristics of the Landsat 1 and 2 MSS Computer Compatible Tapes, Goddard Spa-[illegible]cen Flight Center.
2. Patrick F. Holkenbrink, Manual on Characteristics of Landsat Computer-Compatible Tapes Produced by the EROS Data Center Digital Image Processing System, United States Geological Survey 1978
3. Dieter Steiner and Michael E. Kirby, Geometrical Referencing of Landsat Images by Affine Transformation and Overlaying of Map Data, Photogrammetria 33 (1977) 41-75.
4. Francis J. Ahern and Jennifer Murphy, Radiometric Calibration and Correction of Landsat 1, 2, and 3 MSS Data, Canada Center for Remote Sensing Energy Mines and Resources, Ottawa.
5. K. S. Fu, Digital Pattern Recognition, Springer-verlay Berlin Heicklberg New York, 1980.
6. David Landgrebe, Computer-Based Remote Sensing Technology-A Look to the Future, Remote Sensing of Environment 5,229-246 (1976).
7. Jack Bryant, On the Clustering of Multidimensional Pictorial Data, Pattern Recognition Vol.11 PP 115-125.
8. Z. Kalensky, ERTS Thematic Map from Multidate Digital Lmages, Prceeding Symposium on Remote Sensing and Photo Interpretation Vol.2.
9. L. Borriello and F. Capozza, A Clustering Algorithm for Unsupervised Crop Classification, CCATA, Bari, Italy.
10. Torleiv Orhaug and S Ingvar Akersten, Digital Processing of Multispectral Data, FOA Repport C30075-E1.
11. Beguin H. Do Tu H, and Wilmet. J, Comparison of Classification Methods for Urban Images Interpretation,XIV Congress of the International Society of Photogrammetry, Commission VII.
12. Torleiv Orhaug, Leif Wastenson and S Ingvar Akersten, Forest Inventory Using Landsat Digital Imagery, FOA Repport A30008-E1.
13. W. A. Davis and F. G. Pett, The Identification and Reclassification of Small Regions on Digital Thematic Maps, Forest Management Institute Information Report FMR-X-90.
14. Z. Kalensky and L. R. Scherk, Accuracy of Forest Mapping from Landsat Computer Compatible Tapes.