DIGITAL IMAGE PROCESSING

Minakshi Kumar
Photogrammetry and Remote Sensing Division
Indian Institute of Remote Sensing, Dehra Dun

Abstract : This paper describes the basic technological aspects of Digital Image Processing with special reference to satellite image processing. Basically, all satellite image-processing operations can be grouped into three categories: Image Rectification and Restoration, Enhancement and Information Extraction. The former deals with initial processing of raw image data to correct for geometric distortion, to calibrate the data radiometrically and to eliminate noise present in the data. The enhancement procedures are applied to image data in order to effectively display the data for subsequent visual interpretation. It involves techniques for increasing the visual distinction between features in a scene. The objective of the information extraction operations is to replace visual analysis of the image data with quantitative techniques for automating the identification of features in a scene. This involves the analysis of multispectral image data and the application of statistically based decision rules for determining the land cover identity of each pixel in an image. The intent of classification process is to categorize all pixels in a digital image into one of several land cover classes or themes. This classified data may be used to produce thematic maps of the land cover present in an image.

INTRODUCTION

Pictures are the most common and convenient means of conveying or transmitting information. A picture is worth a thousand words. Pictures concisely convey information about positions, sizes and inter-relationships between objects. They portray spatial information that we can recognize as objects. Human beings are good at deriving information from such images, because of our innate visual and mental abilities. About 75% of the information received by human is in pictorial form.

In the present context, the analysis of pictures that employ an overhead perspective, including the radiation not visible to human eye are considered.

Satellite Remote Sensing and GIS Applications in Agricultural Meteorology
pp. 81-102

Thus our discussion will be focussing on analysis of remotely sensed images. These images are represented in digital form. When represented as numbers, brightness can be added, subtracted, multiplied, divided and, in general, subjected to statistical manipulations that are not possible if an image is presented only as a photograph. Although digital analysis of remotely sensed data dates from the early days of remote sensing, the launch of the first Landsat earth observation satellite in 1972 began an era of increasing interest in machine processing (Cambell, 1996 and Jensen, 1996). Previously, digital remote sensing data could be analyzed only at specialized remote sensing laboratories. Specialized equipment and trained personnel necessary to conduct routine machine analysis of data were not widely available, in part because of limited availability of digital remote sensing data and a lack of appreciation of their qualities.

DIGITAL IMAGE

A digital remotely sensed image is typically composed of picture elements (pixels) located at the intersection of each row i and column j in each K bands of imagery. Associated with each pixel is a number known as Digital Number (DN) or Brightness Value (BV), that depicts the average radiance of a relatively small area within a scene (Fig. 1). A smaller number indicates low average radiance from the area and the high number is an indicator of high radiant properties of the area.

The size of this area effects the reproduction of details within the scene. As pixel size is reduced more scene detail is presented in digital representation.

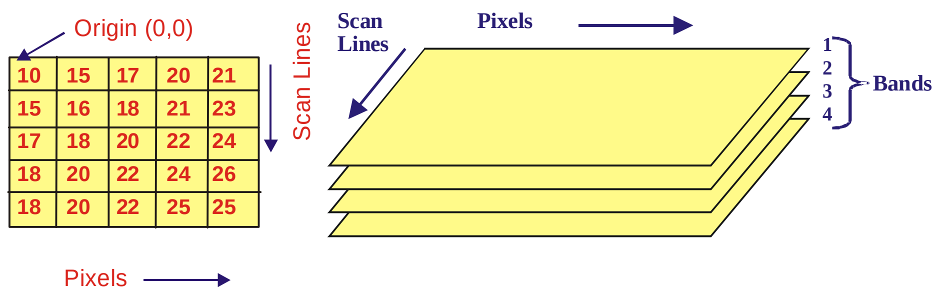

Figure 1 : Structure of a Digital Image and Multispectral Image

COLOR COMPOSITES

While displaying the different bands of a multispectral data set, images obtained in different bands are displayed in image planes (other than their own) the color composite is regarded as False Color Composite (FCC). High spectral resolution is important when producing color components. For a true color composite an image data used in red, green and blue spectral region must be assigned bits of red, green and blue image processor frame buffer memory. A color infrared composite 'standard false color composite' is displayed by placing the infrared, red, green in the red, green and blue frame buffer memory (Fig. 2). In this healthy vegetation shows up in shades of red because vegetation absorbs most of green and red energy but reflects approximately half of incident Infrared energy. Urban areas reflect equal portions of NIR, R & G, and therefore they appear as steel grey.

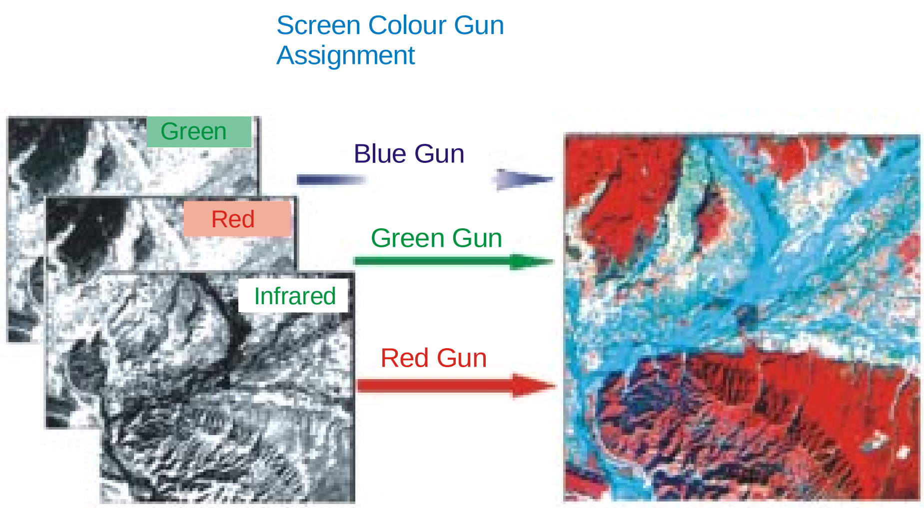

Figure 2: False Color Composite (FCC) of IRS : LISS II Poanta area

IMAGE RECTIFICATION AND REGISTRATION

Geometric distortions manifest themselves as errors in the position of a pixel relative to other pixels in the scene and with respect to their absolute position within some defined map projection. If left uncorrected, these geometric distortions render any data extracted from the image useless. This is particularly so if the information is to be compared to other data sets, be it from another image or a GIS data set. Distortions occur for many reasons.

For instance distortions occur due to changes in platform attitude (roll, pitch and yaw), altitude, earth rotation, earth curvature, panoramic distortion and detector delay. Most of these distortions can be modelled mathematically and are removed before you buy an image. Changes in attitude however can be difficult to account for mathematically and so a procedure called image rectification is performed. Satellite systems are however geometrically quite stable and geometric rectification is a simple procedure based on a mapping transformation relating real ground coordinates, say in easting and northing, to image line and pixel coordinates.

Rectification is a process of geometrically correcting an image so that it can be represented on a planar surface , conform to other images or conform to a map (Fig. 3). That is, it is the process by which geometry of an image is made planimetric. It is necessary when accurate area, distance and direction measurements are required to be made from the imagery. It is achieved by transforming the data from one grid system into another grid system using a geometric transformation.

Rectification is not necessary if there is no distortion in the image. For example, if an image file is produced by scanning or digitizing a paper map that is in the desired projection system, then that image is already planar and does not require rectification unless there is some skew or rotation of the image. Scanning and digitizing produce images that are planar, but do not contain any map coordinate information. These images need only to be geo-referenced, which is a much simpler process than rectification. In many cases, the image header can simply be updated with new map coordinate information. This involves redefining the map coordinate of the upper left corner of the image and the cell size (the area represented by each pixel).

Ground Control Points (GCP) are the specific pixels in the input image for which the output map coordinates are known. By using more points than necessary to solve the transformation equations a least squares solution may be found that minimises the sum of the squares of the errors. Care should be exercised when selecting ground control points as their number, quality and distribution affect the result of the rectification.

Once the mapping transformation has been determined a procedure called resampling is employed. Resampling matches the coordinates of image pixels to their real world coordinates and writes a new image on a pixel by pixel basis. Since the grid of pixels in the source image rarely matches the grid for

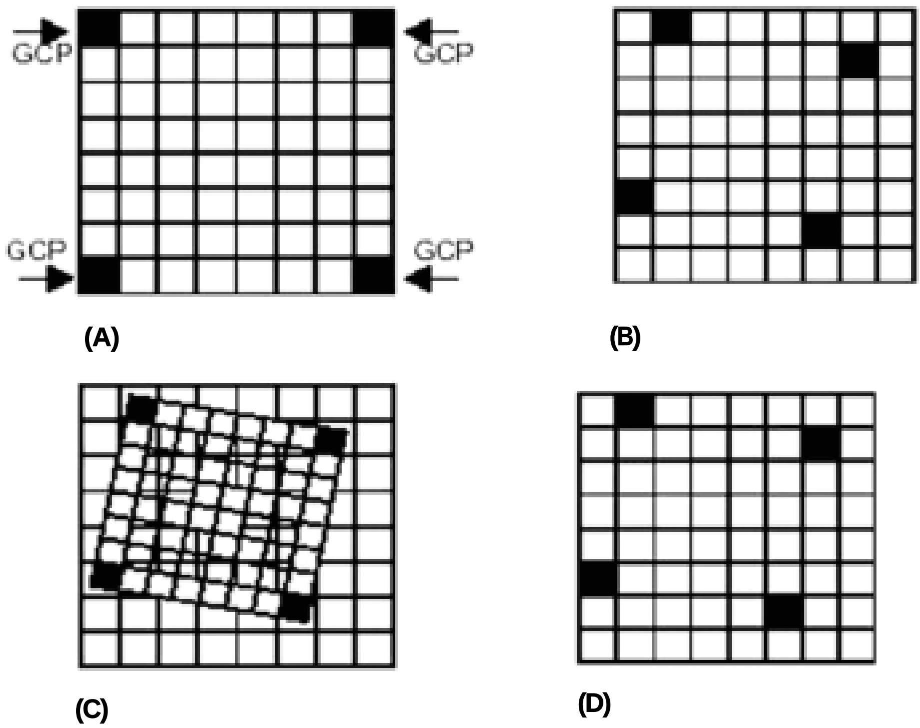

Figure 3 : Image Rectification (a & b) Input and reference image with GCP locations, (c) using polynomial equations the grids are fitted together, (d) using resampling method the output grid pixel values are assigned (source modified from ERDAS Field guide)

the reference image, the pixels are resampled so that new data file values for the output file can be calculated.

IMAGE ENHANCEMENT TECHNIQUES

Image enhancement techniques improve the quality of an image as perceived by a human. These techniques are most useful because many satellite images when examined on a colour display give inadequate information for image interpretation. There is no conscious effort to improve the fidelity of the image with regard to some ideal form of the image. There exists a wide variety of techniques for improving image quality. The contrast stretch, density slicing, edge enhancement, and spatial filtering are the more commonly used techniques. Image enhancement is attempted after the image is corrected for geometric and radiometric distortions. Image enhancement methods are applied separately to each band of a multispectral image. Digital techniques have been found to be most satisfactory than the photographic technique for image enhancement, because of the precision and wide variety of digital processes.

Contrast

Contrast generally refers to the difference in luminance or grey level values in an image and is an important characteristic. It can be defined as the ratio of the maximum intensity to the minimum intensity over an image.

Contrast ratio has a strong bearing on the resolving power and detectability of an image. Larger this ratio, more easy it is to interpret the image. Satellite images lack adequate contrast and require contrast improvement.

Contrast Enhancement

Contrast enhancement techniques expand the range of brightness values in an image so that the image can be efficiently displayed in a manner desired by the analyst. The density values in a scene are literally pulled farther apart, that is, expanded over a greater range. The effect is to increase the visual contrast between two areas of different uniform densities. This enables the analyst to discriminate easily between areas initially having a small difference in density.

Linear Contrast Stretch

This is the simplest contrast stretch algorithm. The grey values in the original image and the modified image follow a linear relation in this algorithm. A density number in the low range of the original histogram is assigned to extremely black and a value at the high end is assigned to extremely white. The remaining pixel values are distributed linearly between these extremes. The features or details that were obscure on the original image will be clear in the contrast stretched image. Linear contrast stretch operation can be represented graphically as shown in Fig. 4. To provide optimal contrast and colour variation in colour composites the small range of grey values in each band is stretched to the full brightness range of the output or display unit.

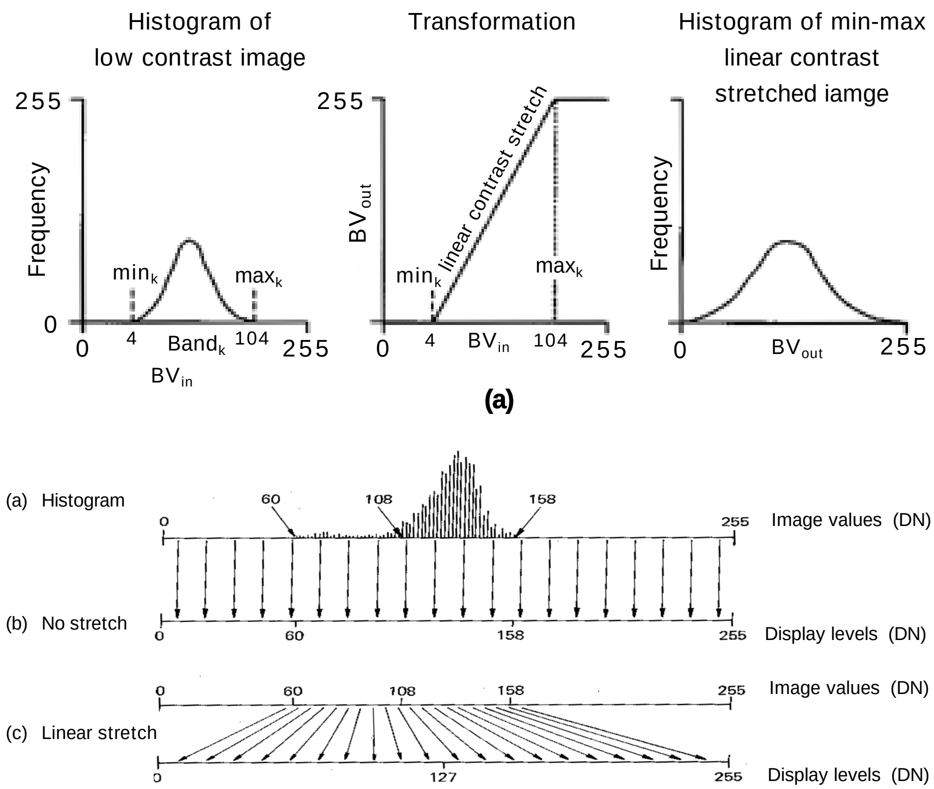

Figure 4: Linear Contrast Stretch (source Lillesand and Kiefer, 1993)

Non-Linear Contrast Enhancement

In these methods, the input and output data values follow a non-linear transformation. The general form of the non-linear contrast enhancement is defined by y = f (x), where x is the input data value and y is the output data value. The non-linear contrast enhancement techniques have been found to be useful for enhancing the colour contrast between the nearly classes and subclasses of a main class.

A type of non linear contrast stretch involves scaling the input data logarithmically. This enhancement has greatest impact on the brightness values found in the darker part of histogram. It could be reversed to enhance values in brighter part of histogram by scaling the input data using an inverse log function.

Histogram equalization is another non-linear contrast enhancement technique. In this technique, histogram of the original image is redistributed to produce a uniform population density. This is obtained by grouping certain adjacent grey values. Thus the number of grey levels in the enhanced image is less than the number of grey levels in the original image.

SPATIAL FILTERING

A characteristic of remotely sensed images is a parameter called spatial frequency defined as number of changes in Brightness Value per unit distance for any particular part of an image. If there are very few changes in Brightness Value once a given area in an image, this is referred to as low frequency area. Conversely, if the Brightness Value changes dramatically over short distances, this is an area of high frequency.

Spatial filtering is the process of dividing the image into its constituent spatial frequencies, and selectively altering certain spatial frequencies to emphasize some image features. This technique increases the analyst's ability to discriminate detail. The three types of spatial filters used in remote sensor data processing are : Low pass filters, Band pass filters and High pass filters.

Low-Frequency Filtering in the Spatial Domain

Image enhancements that de-emphasize or block the high spatial frequency detail are low-frequency or low-pass filters. The simplest low-frequency filter evaluates a particular input pixel brightness value, BV_{in}, and the pixels surrounding the input pixel, and outputs a new brightness value, BV_{out} , that is the mean of this convolution. The size of the neighbourhood convolution mask or kernel (n) is usually 3x3, 5x5, 7x7, or 9x9.

The simple smoothing operation will, however, blur the image, especially at the edges of objects. Blurring becomes more severe as the size of the kernel increases.

Using a 3x3 kernel can result in the low-pass image being two lines and two columns smaller than the original image. Techniques that can be applied to deal with this problem include (1) artificially extending the original image beyond its border by repeating the original border pixel brightness values or (2) replicating the averaged brightness values near the borders, based on the image behaviour within a view pixels of the border. The most commonly used low pass filters are mean, median and mode filters.

High-Frequency Filtering in the Spatial Domain

High-pass filtering is applied to imagery to remove the slowly varying components and enhance the high-frequency local variations. Brightness values tend to be highly correlated in a nine-element window. Thus, the high-frequency filtered image will have a relatively narrow intensity histogram. This suggests that the output from most high-frequency filtered images must be contrast stretched prior to visual analysis.

Edge Enhancement in the Spatial Domain

For many remote sensing earth science applications, the most valuable information that may be derived from an image is contained in the edges surrounding various objects of interest. Edge enhancement delineates these edges and makes the shapes and details comprising the image more conspicuous and perhaps easier to analyze. Generally, what the eyes see as pictorial edges are simply sharp changes in brightness value between two adjacent pixels. The edges may be enhanced using either linear or nonlinear edge enhancement techniques.

Linear Edge Enhancement

A straightforward method of extracting edges in remotely sensed imagery is the application of a directional first-difference algorithm and approximates the first derivative between two adjacent pixels. The algorithm produces the first difference of the image input in the horizontal, vertical, and diagonal directions.

The Laplacian operator generally highlights point, lines, and edges in the image and suppresses uniform and smoothly varying regions. Human vision physiological research suggests that we see objects in much the same way. Hence, the use of this operation has a more natural look than many of the other edge-enhanced images.

Band ratioing

Sometimes differences in brightness values from identical surface materials are caused by topographic slope and aspect, shadows, or seasonal changes in sunlight illumination angle and intensity. These conditions may hamper the ability of an interpreter or classification algorithm to identify correctly surface materials or land use in a remotely sensed image. Fortunately, ratio transformations of the remotely sensed data can, in certain instances, be applied to reduce the effects of such environmental conditions. In addition to minimizing the effects of environmental factors, ratios may also provide unique information not available in any single band that is useful for discriminating between soils and vegetation.

The mathematical expression of the ratio function is

$$BV_{i,j,r} = BV_{i,j,k}/BV_{i,j.l}$$

where $BV_{i,j,r}$ is the output ratio value for the pixel at rwo, i, column j; $BV_{i,j,k}$ is the brightness value at the same location in band k, and $BV_{i,j,l}$ is the brightness value in band l. Unfortunately, the computation is not always simple since $BV_{i,j} = 0$ is possible. However, there are alternatives. For example, the mathematical domain of the function is 1/255 to 255 (i.e., the range of the ratio function includes all values beginning at 1/255, passing through 0 and ending at 255). The way to overcome this problem is simply to give any $BV_{i,j}$ with a value of 0 the value of 1.

Ratio images can be meaningfully interpreted because they can be directly related to the spectral properties of materials. Ratioing can be thought of as a method of enhancing minor differences between materials by defining the slope of spectral curve between two bands. We must understand that dissimilar materials having similar spectral slopes but different albedos, which are easily separable on a standard image, may become inseparable on ratio images. Figure 5 shows a situation where Deciduous and Coniferous Vegetation crops out on both the sunlit and shadowed sides of a ridge.

In the individual bands the reflectance values are lower in the shadowed area and it would be difficult to match this outcrop with the sunlit outcrop. The ratio values, however, are nearly identical in the shadowed and sunlit areas and the sandstone outcrops would have similar signatures on ratio images. This removal of illumination differences also eliminates the dependence of topography on ratio images.

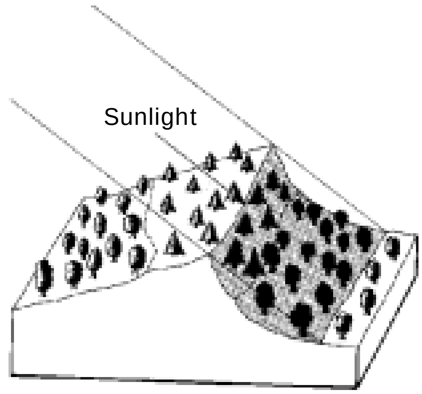

Figure 5: Reduction of Scene Illumination effect through spectral ratioing
(source Lillesand & Kiefer, 1993)

Landcover/ Illumination	Digital Number		Ratio
	Band A	Band B	
Deciduous			
Sunlit	48	50	.96
Shadow	18	19	.95
Coniferous			
Sunlit	31	45	.69
Shadow	11	16	.69

PRINCIPAL COMPONENT ANALYSIS

The multispectral image data is usually strongly correlated from one band to the other. The level of a given picture element on one band can to some extent be predicted from the level of that same pixel in another band.

Principal component analysis is a pre-processing transformation that creates new images from the uncorrelated values of different images. This is accomplished by a linear transformation of variables that corresponds to a rotation and translation of the original coordinate system.

Principal component analysis operates on all bands together. Thus, it alleviates the difficulty of selecting appropriate bands associated with the band ratioing operation. Principal components describe the data more efficiently than the original band reflectance values. The first principal component accounts for a maximum portion of the variance in the data set, often as high as 98%. Subsequent principal components account for successively smaller portions of the remaining variance.

Principal component transformations are used for spectral pattern recognition as well as image enhancement. When used before pattern recognition, the least important principal components are dropped altogether. This permits us to omit the insignificant portion of our data set and thus avoids the additional computer time. The transformation functions are determined during the training stage. Principal component images may be analysed as separate black and white images, or any three component images may be colour coded to form a colour composite. Principal component enhancement techniques are particularly appropriate in areas where little a priori information concerning the region is available.

IMAGE FUSION TECHNIQUES

The satellites cover different portions of the electromagnetic spectrum and record the incoming radiations at different spatial, temporal, and spectral resolutions. Most of these sensors operate in two modes: *multispectral* mode and the *panchromatic* mode.

The *panchromatic* mode corresponds to the observation over a broad spectral band (similar to a typical black and white photograph) and the *multispectral* (color) mode corresponds to the observation in a number of relatively narrower bands. For example in the IRS – 1D, LISS III operates in the multispectral mode. It records energy in the green (0.52 – 0.59 μm), red (0.62-0.68 μm), near infrared (0.77- 0.86 μm) and mid-infrared (1.55 – 1.70 μm). In the same satellite PAN operates in the panchromatic mode. SPOT is another satellite, which has a combination of sensor operating in the multispectral and panchromatic mode. Above information is also expressed by saying that the multispectral mode has a better ***spectral resolution*** than the panchromatic mode.

Now coming to the ***spatial resolution***, most of the satellites are such that the *panchromatic* mode has a better *spatial resolution* than the *multispectral* mode, for e.g. in IRS -1C, PAN has a spatial resolution of 5.8 m whereas in the case of LISS it is 23.5 m. Better is the spatial resolution, more detailed information about a landuse is present in the imagery, hence usually PAN data is used for observing and separating various feature. Both theses type of sensors have their particular utility as per the need of user. If the need of the user is to separate two different kind of landuses, LISS III is used, whereas for a detailed map preparation of any area, PAN imagery is extremely useful.

Image Fusion is the combination of two or more different images to form a new image (by using a certain algorithm).

The commonly applied Image Fusion Techniques are

1. IHS Transformation
2. PCA
3. Brovey Transform
4. Band Substitution

IMAGE CLASSIFICATION

The overall objective of image classification is to automatically categorize all pixels in an image into land cover classes or themes. Normally, multispectral data are used to perform the classification, and the spectral pattern present within the data for each pixel is used as numerical basis for categorization. That is, different feature types manifest different combination of DNs based on their inherent spectral reflectance and emittance properties.

The term *classifier* refers loosely to a computer program that implements a specific procedure for image classification. Over the years scientists have devised many classification strategies. From these alternatives the analyst must select the classifier that will best accomplish a specific task. At present it is not possible to state that a given classifier is "best" for all situations because characteristics of each image and the circumstances for each study vary so greatly. Therefore, it is essential that the analyst understands the alternative strategies for image classification.

The traditional methods of classification mainly follow two approaches: unsupervised and supervised. The unsupervised approach attempts spectral grouping that may have an unclear meaning from the user's point of view. Having established these, the analyst then tries to associate an information class with each group. The unsupervised approach is often referred to as clustering and results in statistics that are for spectral, statistical clusters. In the supervised approach to classification, the image analyst supervises the pixel categorization process by specifying to the computer algorithm; numerical descriptors of the various land cover types present in the scene. To do this, representative sample sites of known cover types, called training areas or training sites, are used to compile a numerical interpretation key that describes the spectral attributes for each feature type of interest. Each pixel in the data set is then compared numerically to each category in the interpretation key and labeled with the name of the category it looks most like. In the supervised approach the user defines useful information categories and then examines their spectral separability whereas in the unsupervised approach he first determines spectrally separable classes and then defines their informational utility.

It has been found that in areas of complex terrain, the unsupervised approach is preferable to the supervised one. In such conditions if the supervised approach is used, the user will have difficulty in selecting training sites because of the variability of spectral response within each class. Consequently, a prior ground data collection can be very time consuming. Also, the supervised approach is subjective in the sense that the analyst tries to classify information categories, which are often composed of several spectral classes whereas spectrally distinguishable classes will be revealed by the unsupervised approach, and hence ground data collection requirements may be reduced. Additionally, the unsupervised approach has the potential advantage of revealing discriminable classes unknown from previous work. However, when definition of representative training areas is possible and statistical information classes show a close correspondence, the results of supervised classification will be superior to unsupervised classification.

Unsupervised classification

Unsupervised classifiers do *not utilize* training data as the basis for classification. Rather, this family of classifiers involves algorithms that examine the unknown pixels in an image and aggregate them into a number of classes based on the natural groupings or clusters present in the image values. It performs very well in cases where the values within a given cover type are close together in the measurement space, data in different classes are comparatively well separated.

The classes that result from unsupervised classification are spectral classes because they are based solely on the natural groupings in the image values, the identity of the spectral classes will not be initially known. The analyst must compare the classified data with some form of reference data (such as larger scale imagery or maps) to determine the identity and informational value of the spectral classes. In the supervised approach we define useful information categories and then examine their spectral separability; in the unsupervised approach we determine spectrally separable classes and then define their informational utility.

There are numerous clustering algorithms that can be used to determine the natural spectral groupings present in data set. One common form of clustering, called the "K-means" approach also called as ISODATA (Interaction Self-Organizing Data Analysis Technique) accepts from the analyst the number of clusters to be located in the data. The algorithm then arbitrarily "seeds", or locates, that number of cluster centers in the multidimensional measurement space. Each pixel in the image is then assigned to the cluster whose arbitrary mean vector is closest. After all pixels have been classified in this manner, revised mean vectors for each of the clusters are computed. The revised means are then used as the basis of reclassification of the image data. The procedure continues until there is no significant change in the location of class mean vectors between successive iterations of the algorithm. Once this point is reached, the analyst determines the land cover identity of each spectral class. Because the K-means approach is iterative, it is computationally intensive. Therefore, it is often applied only to image sub-areas rather than to full scenes.

Supervised classification

Supervised classification can be defined normally as the process of samples of known identity to classify pixels of unknown identity. Samples of known identity are those pixels located within training areas. Pixels located within these areas term the training samples used to guide the classification algorithm to assigning specific spectral values to appropriate informational class.

The basic steps involved in a typical supervised classification procedure are illustrated on Fig. 6.

The training stage
Feature selection
Selection of appropriate classification algorithm
Post classification smoothening
Accuracy assessment

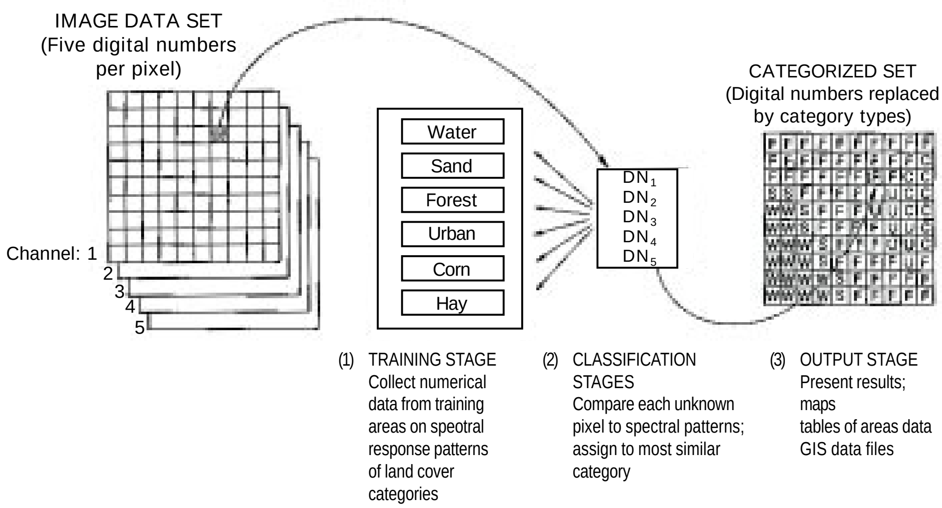

Figure 6: Basic Steps in Supervised Classification

Training data

Training fields are areas of known identity delineated on the digital image, usually by specifying the corner points of a rectangular or polygonal area using line and column numbers within the coordinate system of the digital image. The analyst must, of course, know the correct class for each area. Usually the analyst begins by assembling maps and aerial photographs of the area to be classified. Specific training areas are identified for each informational category following the guidelines outlined below. The objective is to identify a set of pixels that accurately represents spectral variation present within each information region (Fig. 7a).

Select the Appropriate Classification Algorithm

Various supervised classification algorithms may be used to assign an unknown pixel to one of a number of classes. The choice of a particular classifier or decision rule depends on the nature of the input data and the desired output. Parametric classification algorithms assume that the observed measurement vectors X_c for each class in each spectral band during the training phase of the supervised classification are Gaussian in nature; that is, they are normally distributed. Nonparametric classification algorithms make no such assumption. Among the most frequently used classification algorithms are the parallelepiped, minimum distance, and maximum likelihood decision rules.

Parallelepiped Classification Algorithm

This is a widely used decision rule based on simple Boolean "and/or" logic. Training data in n spectral bands are used in performing the classification. Brightness values from each pixel of the multispectral imagery are used to produce an n-dimensional mean vector, $M_c = (\mu_{ck1}, \mu_{c2}, \mu_{c3}, ... \mu_{cn})$ with μ_{ck} being the mean value of the training data obtained for class c in band k out of m possible classes, as previously defined. S_{ck} is the standard deviation of the training data class c of band k out of m possible classes.

The decision boundaries form an n-dimensional parallelepiped in feature space. If the pixel value lies above the lower threshold and below the high threshold for all n bands evaluated, it is assigned to an unclassified category (Figs. 7c and 7d). Although it is only possible to analyze visually up to three dimensions, as described in the section on computer graphic feature analysis, it is possible to create an n-dimensional parallelepiped for classification purposes.

The parallelepiped algorithm is a computationally efficient method of classifying remote sensor data. Unfortunately, because some parallelepipeds overlap, it is possible that an unknown candidate pixel might satisfy the criteria of more than one class. In such cases it is usually assigned to the first class for which it meets all criteria. A more elegant solution is to take this pixel that can be assigned to more than one class and use a minimum distance to means decision rule to assign it to just one class.

Minimum Distance to Means Classification Algorithm

This decision rule is computationally simple and commonly used. When used properly it can result in classification accuracy comparable to other more computationally intensive algorithms, such as the maximum likelihood algorithm. Like the parallelepiped algorithm, it requires that the user provide the mean vectors for each class in each hand μ_{ck} from the training data. To perform a minimum distance classification, a program must calculate the distance to each mean vector, μ_{ck} from each unknown pixel (BV_{ijk}). It is possible to calculate this distance using Euclidean distance based on the Pythagorean theorem (Fig. 7b).

The computation of the Euclidean distance from point to the mean of Class-1 measured in band relies on the equation

$$\text{Dist} = \text{SQRT}\{ (BV_{ijk} - \mu_{ck})^2 + (BV_{ijl} - \mu_{cl})^2 \}$$

Where μ_{ck} and μ_{cl} represent the mean vectors for class c measured in bands k and l.

Many minimum-distance algorithms let the analyst specify a distance or threshold from the class means beyond which a pixel will not be assigned to a category even though it is nearest to the mean of that category.

Maximum Likelihood Classification Algorithm

The maximum likelihood decision rule assigns each pixel having pattern measurements or 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, that is, Gaussian. In other words, training data with bi-or trimodal histograms in a single band are not ideal. In such cases, the individual modes probably represent individual classes that should be trained upon individually and labeled as separate classes. This would then produce unimodal, Gaussian training class statistics that would fulfil the normal distribution requirement.

The Bayes's decision rule is identical to the maximum likelihood decision rule that it does not assume that each class has equal probabilities. A priori probabilities have been used successfully as a way of incorporating the effects of relief and other terrain characteristics in improving classification accuracy. The maximum likelihood and Bayes's classification require many more computations per pixel than either the parallelepiped or minimum-distance classification algorithms. They do not always produce superior results.

Classification Accuracy Assessment

Quantitatively assessing classification accuracy requires the collection of some in situ data or a priori knowledge about some parts of the terrain which can then be compared with the remote sensing derived classification map. Thus to assess classification accuracy it is necessary to compare two classification maps 1) the remote sensing derived map, and 2) assumed true map (in fact it may contain some error). The assumed true map may be derived from in situ investigation or quite often from the interpretation of remotely sensed data obtained at a larger scale or higher resolution.

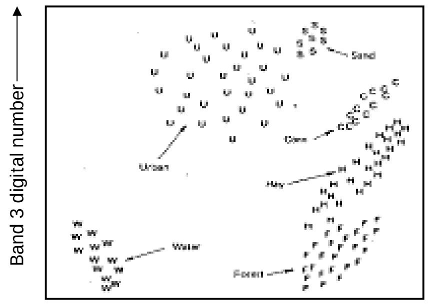

Figure 7a: Pixel observations from selected training sites plotted on scatter diagram

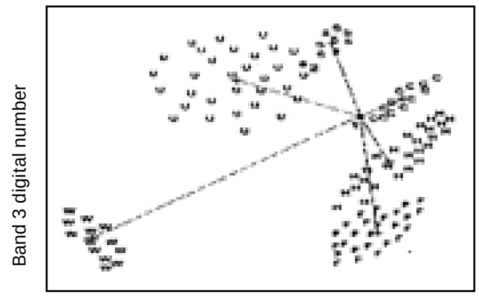

Figure 7b: Minimum Distance to Means Classification strategy

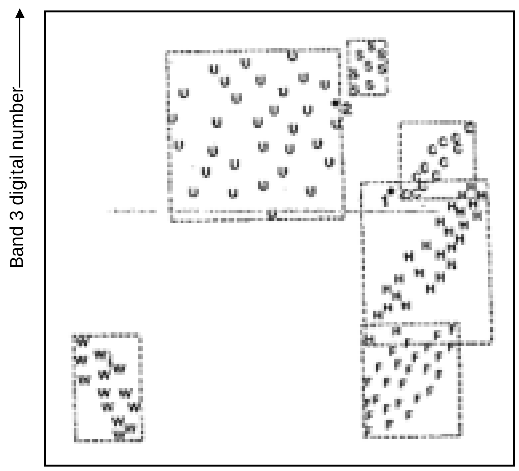

Figure 7c: Parallelepiped classification strategy

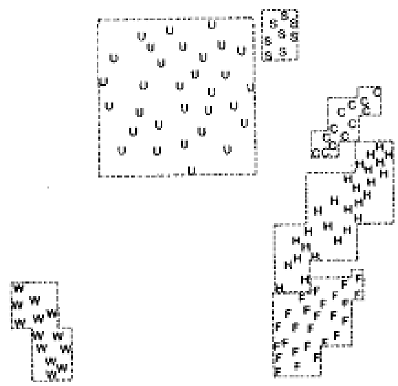

Figure 7d: Stepped parallelepipeds to avoid overlap (source Lillesand and Kiefer, 1993)

Classification Error Matrix

One of the most common means of expressing classification accuracy is the preparation of classification error matrix sometimes called confusion or a contingency table. Error matrices compare on a category by category basis, the relationship between known reference data (ground truth) and the corresponding results of an automated classification. Such matrices are square, with the number of rows and columns equal to the number of categories whose classification accuracy is being assessed. Table 1 is an error matrix that an image analyst has prepared to determine how well a Classification has categorized a representative subset of pixels used in the training process of a supervised classification. This matrix stems from classifying the sampled training set pixels and listing the known cover types used for training (columns) versus the Pixels actually classified into each land cover category by the classifier (rows).

Table 1. Error Matrix resulting from classifying training Set pixels

	W	S	F	U	C	H	Row Total
W	480	0	5	0	0	0	485
S	0	52	0	20	0	0	72
F	0	0	313	40	0	0	353
U	0	16	0	126	0	0	142
C	0	0	0	38	342	79	459
H	0	0	38	24	60	359	481
Column Total	480	68	356	248	402	438	1992

Classification data Training set data (Known cover types) ⟶

Producer's Accuracy

W=480/480 = 100%

S = 052/068 = 16%

F = 313/356 = 88%

U =126/241l = 51%

C =342/402 = 85%

H = 359/438 = 82%

Users Accuracy

W= 480/485 = 99%

S = 052/072 = 72%

F = 313/352 = 87%

U = 126/147 = 99%

C = 342/459 = 74%

H = 359/481 = 75%

Overall accuracy = (480 + 52 + 313+ 126+ 342 +359)/1992= 84%

W, water; S, sand; F, forest; U, urban; C, corn; H, hay
(source Lillesand and Kiefer, 1993).

An error matrix expresses several characteristics about classification performance. For example, one can study the various classification errors of omission (exclusion) and commission (inclusion). Note in Table 1 the training set pixels that are classified into the proper land cover categories are located along the major diagonal of the error matrix (running from upper left to lower right). All non-diagonal elements of the matrix represent errors of omission or commission. Omission errors correspond to non-diagonal column elements (e.g. 16 pixels that should have classified as "sand" were omitted from that category). Commission errors are represented by non-diagonal row elements (e.g. 38 urban pixels plus 79 hay pixels were improperly included in the corn category).

Several other measures for e.g. the overall accuracy of classification can be computed from the error matrix. It is determined by dividing the total number correctly classified pixels (sum of elements along the major diagonal) by the total number of reference pixels. Likewise, the accuracies of individual categories can be calculated by dividing the number of correctly classified pixels in each category by either the total number of pixels in the corresponding rows or column. Producers accuracy which indicates how well the training sets pixels of a given cover type are classified can be determined by dividing the number of correctly classified pixels in each category by number of training sets used for that category (column total). Users 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). This figure is a measure of commission error and indicates the probability that a pixel classified into a given category actually represents that category on ground.

Note that the error matrix in the table indicates an overall accuracy of 84%. However producers accuracy ranges from just 51% (urban) to 100% (water) and users accuracy ranges from 72% (sand) to 99% (water). This error matrix is based on training data. If the results are good it indicates that the training samples are spectrally separable and the classification works well in the training areas. This aids in the training set refinement process, but indicates little about classifier performance else where in the scene.

Kappa coefficient

Kappa analysis is a discrete multivariate technique for accuracy assessment. Kappa analysis yields a Khat statistic that is the measure of agreement of accuracy. The Khat statistic is computed as

$$\text{Khat} = \frac{N\sum^{r} x_{ii}(\sum x_{i} + {*}x_{+i})}{N^{2} - \sum^{r}(x_{i+} + {*}x_{+i})^{r}}$$

Where r is the number of rows in the matrix xii is the number of observations in row i and column i, and x_{i+} and x_{+i} are the marginal totals for the row i and column i respectively and N is the total number of observations.

CONCLUSIONS

Digital image processings of satellite data can be primarily grouped into three categories : Image Rectification and Restoration, Enhancement and Information extraction. Image rectification is the pre-processing of satellite data for geometric and radiometric connections. Enhancement is applied to image data in order to effectively display data for subsequent visual interpretation. Information extraction is based on digital classification and is used for generating digital thematic map.

REFERENCES

Campbell, J.B. 1996. Introduction to Remote Sensing. Taylor & Francis, London.

ERDAS IMAGINE 8.4 Field Guide: ERDAS Inc.

Jensen, J.R. 1996. Introduction to Digital Image Processing : A Remote Sensing Perspective. Practice Hall, New Jersey.

Lillesand, T.M. and Kiefer, R. 1993. Remote Sensing Image Interpretation. John Wiley, New York.