DIGITAL IMAGE PROCESSING

Minakshi Kumar
Scientist
Photogrammetry and Remote Sensing Division
Indian Institute of Remote Sensing
minakshi@iirs.gov.in

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 are 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. 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. 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 (refer 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.

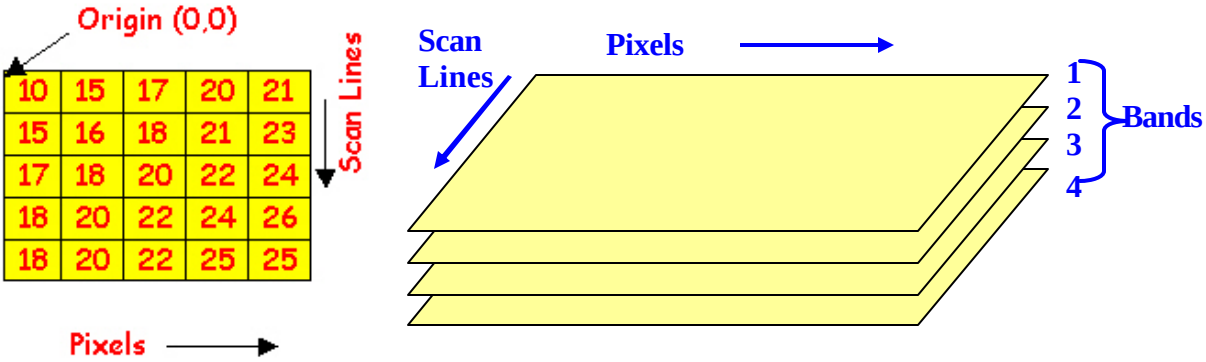


Figure 1: Structure of a multispectral image

DIGITAL IMAGE DATA FORMATS

The image data acquired from Remote Sensing Systems are stored in different types of formats viz. (1) band sequential (BSQ), (2) band interleaved by line (BIL), (3) band interleaved by pixel (BIP). It should be noted, however, that each of these formats is usually preceded on the digital tape by "header" and/or "trailer" information, which consists of ancillary data about the date, altitude of the sensor, attitude, sun angle, and so on. Such information is useful when geometrically or radiometrically correcting the data. The data are normally recorded on nine-track CCTs with data density on the tape of 800, 1600, or 6250 bits per inch (bpi).

Band Sequential Format

The band sequential format requires that all data for a single band covering the entire scene be written as one file. Thus if one wanted the area in the center of a scene in four bands, it would be necessary to read into this location in four separate files to extract the desired information. Many researchers like this format because it is not necessary to read "serially" past unwanted information if certain bands are of no value. The number of tapes may be dependent on the number of bands provided for the scene.

Band Interleaved by Line Format

In this format, the data for the bands are written line by line onto the same tape (i.e. line 1 band 1, line 1 band 2, line 1 band 3, line 1 band 4, etc.). It is a useful format if all the bands are to be used in the analysis. If some bands are not of interest, the format is inefficient since it is necessary to read serially past all the unwanted data.

Band Interleaved by Pixel Format

In this format, the data for the pixel in all bands are written together. Taking the

example of LANDSAT - MSS (Four Bands of Image Data every element in the matrix has four pixel values (one from each spectral band) placed one after the other [i.e., pixel (1,1) of band 1, pixel (1,1) of band 2, pixel (1,1) of band 3, pixel (1,1) of band 4, and then pixel (1,2) of band 1, pixel (1,2) of band 2 and so on]. Again, this is a practical data format if all bands are to be used, otherwise it would be inefficient. This format is not much popular now, but was used extensively by EROS Data Centre for Landsat scene at initial stage.

SOFTWARE CONSIDERATIONS

Digital Image Processing is an extremely broad subject and involves procedures which are mathematically complex. The procedure for digital image processing may be categorized into the following types of computer assisted operations.

1. **Image Rectification** : These operations aim to correct distorted or degraded image data to create a faithful representation of the original scene. This typically involves the initial processing of raw image data to correct for geometric distortion, to calibrate the data radiometrically and to eliminate noise present in the data. Image rectification and restoration procedures are often termed preprocessing operations because they normally precede manipulation and analysis of image data.

2. **Image Enhancement** : These 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 is to create new images from original data in order to increase the amount of information that can be visually interpreted from the data. It includes level slicing, contrast stretching, spatial filtering edge enhancement, spectral ratioing, principal components and intensity-hue-saturation color space transformations.

3. **Image Classification** : The objective of these 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.

COLOR COMPOSITIES

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. In this healthy vegetation shows up in shades of red because vegetation absorbs most of green and red energy but reflects approximately half of incidence Infrared energy. Urban areas reflect equal problem of NIR, R & G, and therefore they appear as steel grey.

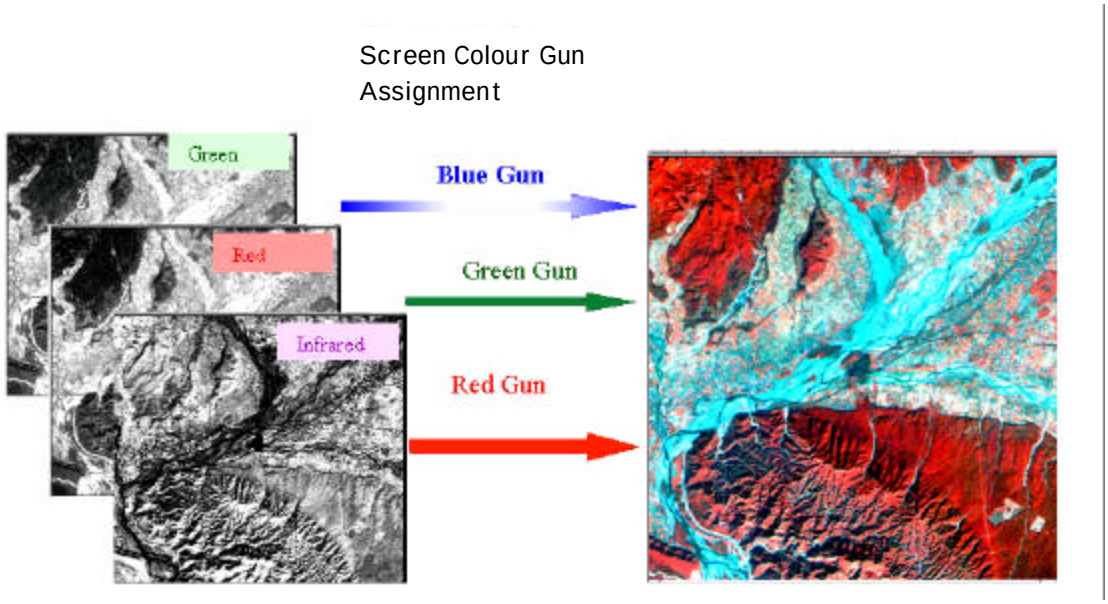


Figure 2: False Colour Composite (FCC 4,2,1) of LISS II Poanta Area

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 the any data extracted from the image useless. This is particularly so if the information is to be compared to other datasets, be it from another image or a GIS dataset. Distortions occur for many reasons. For instance distortions 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.

Raw , remotely sensed image data gathered by a satellite or an aircraft are representation of irregular surface of earth., or Geometry of an image is distorted with respect to north south orientation of map.

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. 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 georeferenced, 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)

Rectification Procedure involves two steps

- Spatial Interpolation using Coordinate Transformations
- Intensity Interpolation (Resampling)

Spatial Interpolation In this method Geometric relationship between the input pixel location (row, column) and the associated map coordinates of the same area(x,y) are identified. This establishes the transformation parameters to rectify or relocate input pixels at location (x´, y´) to its proper position in the rectified output image (x,y). It involves selecting Ground Control Points (GCPS) and fitting polynomial equations using least squares technique.

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

RESAMPLING

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

This process involves the extraction of a brightness value from a location in the input image and its reallocation in the appropriate coordinate location in the rectified output image. There are three techniques of resampling. In the first, the nearest old cell (based on cell center position) is chosen to determine the value of the new cell. This is called a nearest neighbor rule. In the second, a distance weighted average of the four nearest old cells is assigned to the new cell. This technique is called bilinear interpolation. In the third, a distance weighted average of the 16 nearest old cells is assigned to the new cell. This technique is called cubic convolution. Nearest neighbor resampling should be used when the data values cannot be changed, for example, with categorical data or qualitative data such as soils types. The bilinear and cubic convolution routines are appropriate for quantitative data such as remotely sensed imagery.

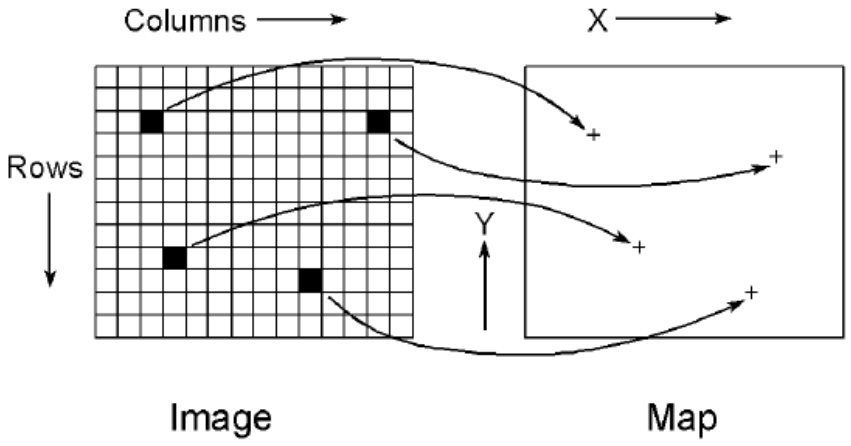


Figure 3 : Image Rectification

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.

Reasons for low contrast of image data

Most of the satellite images lack adequate contrast and require contrast improvement. Low contrast may result from the following causes:

(i) The individual objects and background that make up the terrain may have a nearly uniform electromagnetic response at the wavelength band of energy that is recorded by the remote sensing system. In other words, the scene itself has a low contrast ratio.

(ii) Scattering of electromagnetic energy by the atmosphere can reduce the contrast of a scene. This effect is most pronounced in the shorter wavelength portions.

(iii) The remote sensing system may lack sufficient sensitivity to detect and record the contrast of the terrain. Also, incorrect recording techniques can result in low contrast imagery although the scene has a high-contrast ratio.

Images with low contrast ratio are commonly referred to as `Washed out', with nearly uniform tones of gray. Detectors on the satellite are designed to record a wide range of scene brightness values without getting saturated. They must encompass a range of brightness from black basalt outcrops to white sea ice. However, only a few individual scenes have a brightness range that utilizes the full sensitivity range of remote sensor detectors. The limited range of brightness values in most scenes does not provide adequate contrast for detecting image features. Saturation may also occur when the sensitivity range of a detectors is insufficient to record the full brightness range of a scene. In the case of saturation, the light and dark extremes of brightness on a scene appear as saturated white or black tones on the image.

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.

Contrast enhancement can be effected by a linear or non linear transformation.

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 & 5.

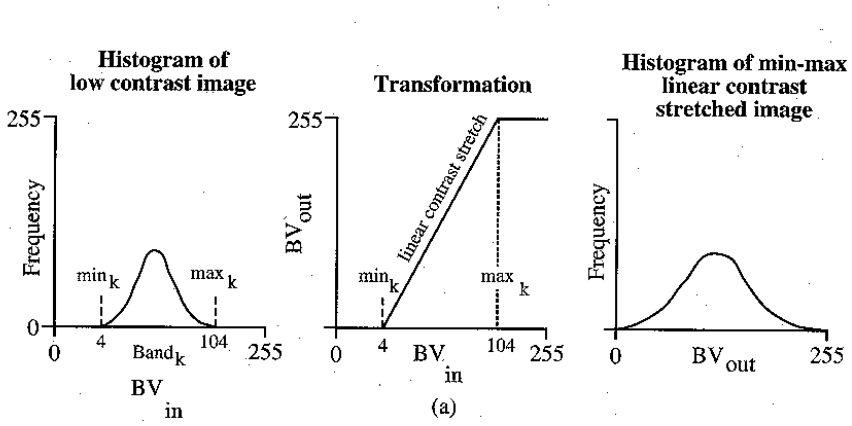


Figure 4 Linear Contrast Stretch - Transformation Function

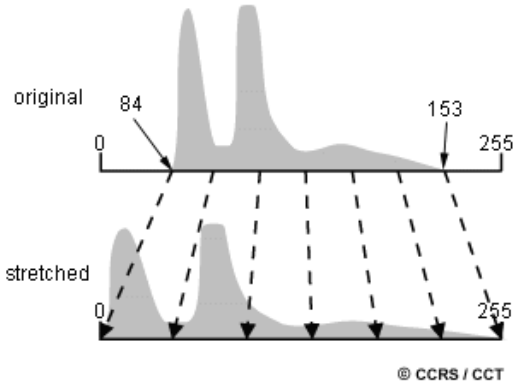


Figure 5 Linear Contrast Stretch (Source: CCRS)

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.

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.

Though there are several non-linear contrast enhancement algorithms available in literature, the use of non-linear contrast enhancement is restricted by the type of application. Good judgment by the analyst and several iterations through the computer are usually required to produce the desired results.

A type of non linear contrast stretch involves scaling the input data logarthemically . 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.(Refer figure 6).

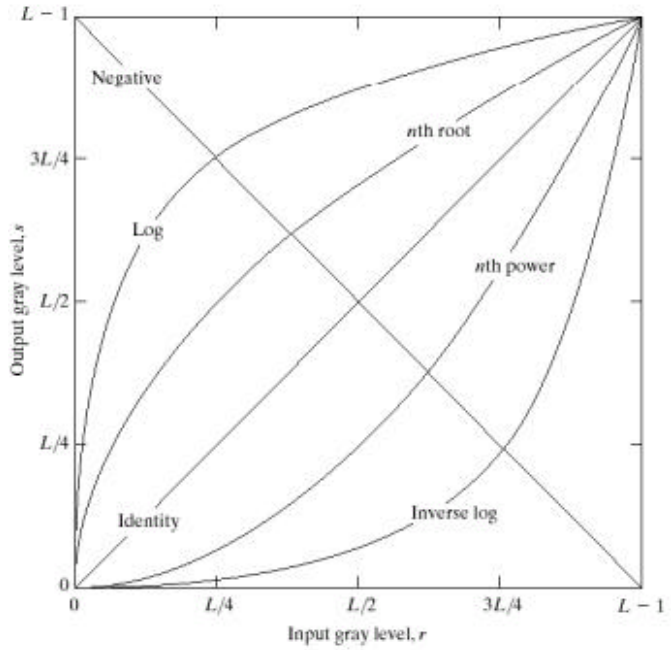


Figure 6: Logic of a Non Linear Logarithmic and Inverse Log Contrast Stretch Algorithms

HISTOGRAM EQUALIZATION

This 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.

The redistribution of the histogram results in greatest contrast being applied to the most populated range of brightness values in the original image. In this process the light and dark tails of the original histogram are compressed, thereby resulting in some loss of detail in those regions. This method gives large improvement in image quality when the histogram is highly peaked.

With any type of contrast enhancement, the relative tone of different materials is modified. Simple linear stretching has the least effect on relative tones, and brightness differences can still be related to the differences in reflectivity. In other cases, the relative tone can no longer be meaningfully related to the reflectance of materials. An analyst must therefore be fully cognizant of the processing techniques that have been applied to the data.

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 change 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 enhancement that de-emphasize or block the high spatial frequency detail are low-frequency or low-pass filters. The simplest low-frequency filter (LEF) 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 behavior within a view pixels of the border.

The neighborhood ranking median filter is useful for removing noise in an image, especially shot noise by which individual pixels are corrupted or missing. Instead of computing the average (mean) of the nine pixels in 3x3 convolution, the median filter ranks the pixels in the neighborhood from lowest to highest and selects the median value, which is then placed in the central value of the mask.

A median filter has certain advantages when compared with weighted convolution filters, including (1) it does not shift boundaries, and (2) the minimal degradation to edges allows the median filter to be applied repeatedly, which allows fine detail to be erased and large regions to take on the same brightness value.

A mode filter is used for removing random noise present in the imagery. In the mode filter, the central pixel value is the window make is replaced by the most frequently occurring value. This is a post classification filter.

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 shaped 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.

Compass gradient masks may be used to perform two-dimensional, discrete differentiation directional edges enhancement .Laplacian convolution masks may be

applied to imagery to perform edge enhancement. The Laplacian is a second derivative (as opposed to the gradient, which is a first derivative) and is invariant to rotation, meaning that it is insensitive to the direction in which the discontinuities (points, line, and edges) run.

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.

Numerous coefficients can be placed in the convolution masks. Usually, the analyst works interactively with the remotely sensed data, trying different coefficients and selecting those that produce the most effective results. It is also possible to use combinations of operation for edge detection. For example, a combination of gradient and Lapacian edge operation may be superior to using either edge enhancement alone.

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

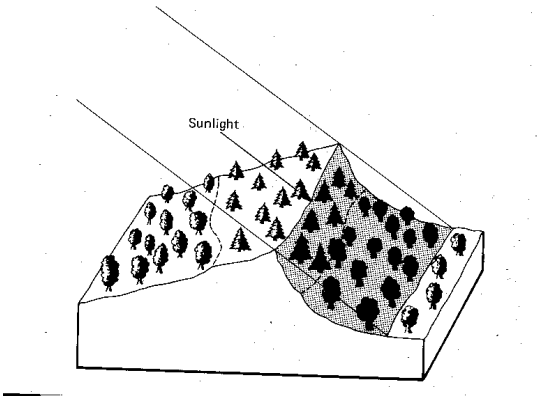
$$BV_{i,j,r} = BV_{i,j,k}/BV_{i,j,l}$$

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. Alternatively, some like to add a small value (e.g.0.1) to the denominator if it equals zero.

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 7 shows a situation where Decidous and Coniferous Vegetation crops out on both the sunlit and shadowed sides of a ridge.



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

Figure 7 Reduction of Scene Illumination effect through spectral ratioing

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.

The potential advantage of band ratioing is that greater contrast between or within classes might be obtained for certain patterns of spectral signatures. Ratioing is a non-linear operation and has the property of cancelling or minimizing positively correlated variations in the data while emphasizing negatively correlated variations. In other words, a ratio image will enhance contrast for a pair of variables which exhibit negative correlation between them.

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 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.

This transformation is conceptualized graphically by considering the two-dimensional distribution of pixel values obtained in two bands, which are labeled simply X_1 and X_2. A scatterplot of all the brightness values associated with each pixel in each band is shown in Figure 8, along with the location of the respective means, μ_1 μ_2. The spread or variance of the distribution of points is an indication of the correlation and quality of information associated with both bands. If all the data points clustered in an extremely tight zone in the two-dimensional space, these data would probably provide very little information as they are highly correlated.

The initial measurement coordinate axes (X_1 and X_2) may not be the best arrangement in multispectral feature space to analyze the remote sensor data associated with these two bands. The goal is to use principal components analysis to translate and/or route the original axes so that the original brightness values on axes X_1 and X_2 are redistributed (reprojected) onto a new set of axes or dimensions, X'_1 and X'_2. For example, the best translation for the original data points from X_1 to X'_1 and from X_2 to X'_2 coordinate systems might be the simple relationship $X'_1 = X_1 - \mu_1$ and $X'_2 = X - \mu_2$. Thus, the origin of the new coordinate system (X'_1 and X'_2) now lies at the location of both means in the original scatter of points

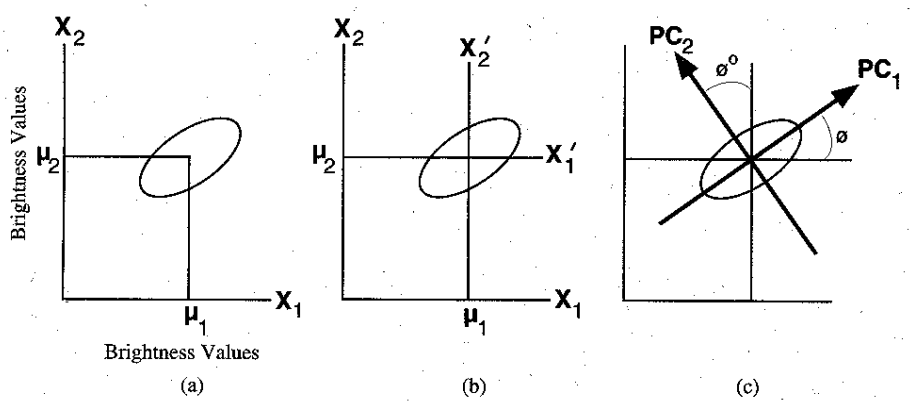


Figure 8: Diagram representing relationship between two principal components.

The X' coordinate system might then be rotated about its new origin (μ_1, μ_2) in the new coordinate system some ϕ degree so that the axis X'_1 is associated with the maximum amount of variance in the scatter of points . This new axis is called the first principal component ($PC_1 = \lambda_1$). The second principal component ($PC_2 = \lambda_2$) is perpendicular (Orthogonal) to PC_1. Thus, the major and minor axes of the ellipsoid of points in bands X_1 and X_2 are called the principal components. The third, fourth, fifth, and so on, components contain decreasing amounts of the variance found in the data set.

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 priori information concerning the region is available.

IMAGE FUSION TECHNIQUES

Introduction

Analysis of the remotely sensed data has become a common source of gathering the information about various landcovers and landuses. This data is obtained from the earth observation satellites, which have been launched from time to time. These satellites cover different portions of the electromagnetic spectrum and record the incoming radiations at different spatial, temporal, and spectral resolutions (Pohl and Genderen 1998).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 band. For example in the IRS – 1C, LISS III operates in the multispectral mode. It records energy in the green (0.52 – 0.59μm), red (0.62-0.68μm), near infrared (0.77- 0.86μm) and mid-infrared (1.55 –1.70). 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 these 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 5.8m whereas in the case of LISS it is 23.5m. 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 has 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.

Whenever one talks of fusion there are two terms, which are often employed. More often than once these terms are often confused for one another. They are *Data Fusion* and *Image fusion.*

Data Fusion *is a process dealing with data and information from multiple sources to achieve refined/improved information for decision-making* (Hall 1992).Often, in case of data fusion, not only remote sensing images are fused but also further ancillary data (e.g. topographic maps, GPS coordinates, geophysical information etc.) contribute to the resulting image. (Harris and Murray 1989)

Image Fusion *is the combination of two or more different images to form a new image by using a certain algorithm* (Genderen and Pohl 1994).

The actual procedure of fusion does not consume much time and energy. It is the preprocessing work and the efforts to bring the diverse data to a common platform that requires a greater attention and effort.

Proper Pre-Processing of the image data

Fusion involves the interaction of the images having different spatial resolution. These images have different pixel size, which creates problems while merging. Hence the image data is resampled to a common pixel spacing and map projection. The common pixel spacing should be the one, which is required in the desired fused image. Besides all sensor-specific corrections and enhancements of image data have to be applied prior to image fusion. After the fusion process the contribution of each sensor cannot be distinguished or quantified. Therefore as a general rule one must attempt to produce the best single sensor geometry and radiometry and then fuse the images. Any spatial enhancement performed prior to image fusion is of benefit to the resulting fused image.

Image Fusion Techniques

In general the fusion techniques can be grouped into two classes:

1) Color related techniques.
2) Statistical/numerical techniques

Color related techniques

Color Transformations: RGB to IHS

IHS. (Intensity, Hue, Saturation.) Transformation

The *intensity* (I) component is not associated with any color and its value varies from black (a value of 0) to white (that is 1). In other words for any color intensity is a measure of the degree of whitishness. That is there might be a lighter shade of red and there might be a darker shade of red.

The *Hue* (H) component represents the dominant wavelength of color, i.e. it is this component, which provides the information whether, and a color is red or green. That is, the separation between the colors is provided by this component.

The *saturation* (S) component represents the purity of the color and thus describes the separation within a color.(John R. Jensen 2nd edition)

For our purpose of Fusion the most important thing to remember is that this color coordinate system effectively separates the spatial and spectral components of any color**.**

Thus *Intensity* is a spatial component whereas *hue* and *saturation* are the spectral components

IHS Transformation and its utility in Image fusion

IHS transformation has its own inherent advantages. It relates to human color perception parameters. Its application in data merging involves the following steps:

- *The transformation from the RGB domain to the IHS domain*
- *Contrast Manipulation*
- *Substitution of the Intensity component by the High Resolution imagery*
- *Reverse transformation from the IHS domain to the RGB domain*

This method provides an effective method to obtain an image, which has increased spatial resolution without any appreciable change in the spectral content. In any display device there are three color guns, Red, green, and blue. While displaying any multispectral imagery we can display at a time a color composite comprising of three bands, with one band assigned to each color gun. Hence an obvious drawback of this method is that at a time only three datasets can be transformed into the IHS domain. Hence a simultaneous operation which improves the spatial information of all the bands is not possible in this technique. Of all the methods used to merge multiresolution data, the IHS method distorts the spectral characteristics most.

Statistical/Numerical Techniques

Application of PCA in image fusion

The process of fusion involves the following steps:

1. Calculation of covariance (unstandardized PCA) or correlation matrix (standardized PCA).
2. Calculation of eigenvalues, eigenvectors of the above matrix.
3. Computation of the principle components.
4. Contrast stretching the high-resolution data to have approximately the same variance and average as the first principal component image. The stretched high-resolution data is substituted for the first principal component image. This mage can be substituted for the first principal component image because the first component image usually contains all the information that is common to all bands while the spectral information unique to any of the input bands is mapped to other components. An additional advantage is that PC1 accounts for maximum variance, which can maximize the effect of the high-resolution data in the fused image.
5. The inverse transformation of the PC's to the RGB imagery is the next step to be followed.

BROVEY Transform

The Brovey transform equation is given by:

$$DN_{fused} = \left(\frac{DN_{b1}}{DN_{b1} + DN_{b2} + DN_{bn}} \right) \left(DN_{highres} \right)$$

DN_{b1} is the pixel value for the band 1 of th multispctr al imagery,

DN_{b2} is the pixel value for the band 2 of the multispect ral imagery and so on.

$DN_{highres}$ is the pixel value for the high spatial resolution image.

DN_{fused} is the pixel value for the resultant fused image.

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 situation because characteristics of each image and the circumstances for each study vary so greatly. Therefore, it is essential that analyst understand 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(the user) 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 potential advantage of revealing discriminable classes

unknown from previous work. However, when definition of representative training areas is possible and statistical an information classes show a close correspondence, the results of supervised classification will be superior to unsupervised classification.

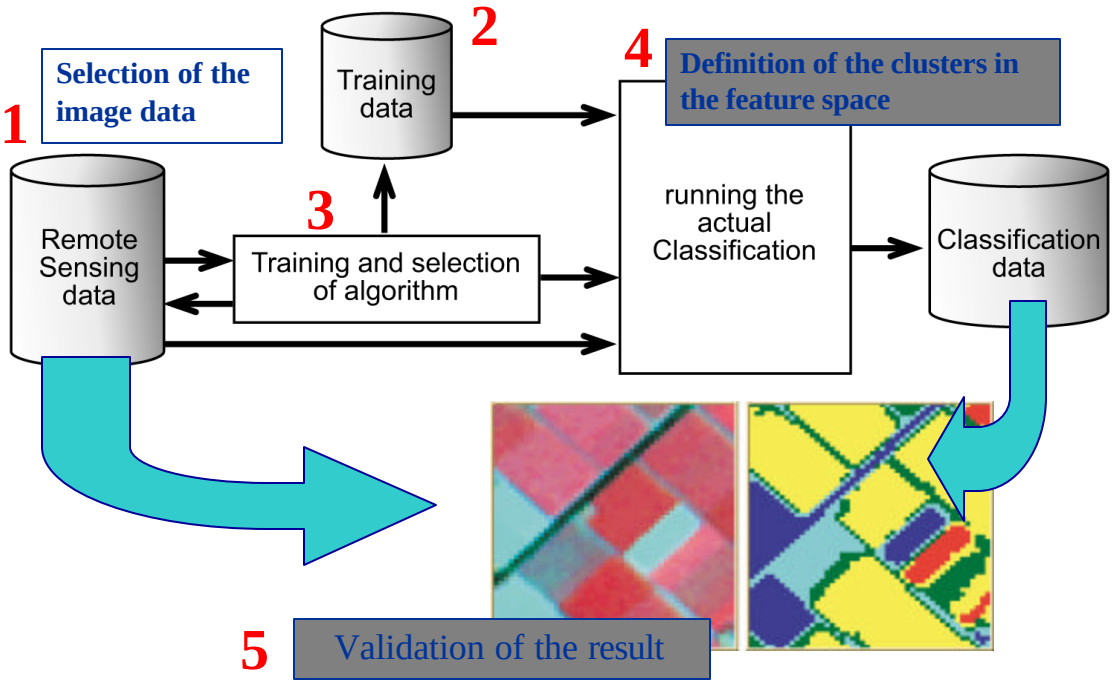


Figure 9: Basic Steps in Supervised Classification (Source ITC slides)

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 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 subareas rather than to full scenes.

There are thousands of pixels in a remotely sensed image, if they are considered individually as prospective members of groups, the distances to other pixels can always be used to define group membership. How can such distances be calculated? A number of methods for finding distances in multidimensional data space are available. One of the simplest is Euclidean distance:

$$Dab = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$$

Where i represents one of n spectral bands, a and b are pixels, and D_{ab} is the distance between the two pixels. The distance calculation is based on the Pythagorean theorem.

$$C = \sqrt{\sum (a^2 + b^2)}$$

In this instance we are interested c; a, b and c measured in units of the two spectral channels.

$C = D_{ab}$

Advantages

Advantages of unsupervised classification (relative to supervised classification) can be enumerated as follows:

1. No extensive prior knowledge of the region is required.
2. The opportunity for human error is minimized.
3. Unique classes are recognized as distinct units.

Disadvantages and limitations

Disadvantages and limitations arise primarily from a reliance on "natural" grouping and difficulties in matching these groups to the informational categories that are of interest to the analyst.

- Unsupervised classification identifies spectrally homogeneous classes within the data; these classes do not necessarily correspond to the informational categories that are of interest to analyst. As a result, the analyst is faced with the problem of matching spectral classes generated by the classification to the informational classes that are required by the ultimate user of the information.

    - Spectral properties of specific information classes will change over time (on a seasonal basis, as well as over the year). As a result, relationships between informational classes and spectral classes are not constant and relationships defined for one image can seldom be extended to others.

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.

1. The training stage
2. Feature selection
3. Selection of appropriate classification algorithm
4. Post classification smoothening
5. Accuracy assessment

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.

KEY CHARACTERISTICS OF TRAINING AREAS

Shape

Shapes of training areas are not important provided that shape does not prohibit accurate delineating and positioning of outlines of regions on digital images. Usually it is easiest to define rectangular or polygonal areas, as such shapes minimize the number of vertices that must be specified, which is usually the most bothersome task for the analyst.

Location

Location is important as each informational category should be represented by several training areas positioned throughout the image. Training areas must be positioned in locations that favour accurate and convenient transfer of their outlines from maps and aerial photographs to the digital image. As the training data are to represent variation within the image, they must not be clustered in favoured regions of the image, which may not typify conditions encountered throughout the image as a whole. As it is desirable for the analyst to use direct field observations in the selection of training data, the requirement for an even distribution of training fields often conflicts with practical constraints, as it may not be practical to visit remote or inaccessible sites that may seem to form good areas for training data. Often aerial observation, or use of good maps and aerial photographs, can provide the basis for accurate delineation of training fields that cannot be inspected in the field.

Number

The optimum number of training areas depends on the number of categories to be mapped, their diversity, and the resources that can be devoted to delineating training areas. Ideally, each information category, or each spectral subclass, should be represented by a number (perhaps 5 to 10 at minimum) of training areas to ensure that spectral properties of each category are represented. Experience indicates that it is usually better to define many small training fields than to use only a few large fields.

Placement

Placement of training areas may be important. Training areas should be placed within the image in a manner that permits convenient and accurate location with respect to distinctive features, such as water bodies, or boundaries between distinctive features on the image. They should be distributed throughout the image so that they provide a basis for representation of diversity present within the scene. Boundaries of training fields should be placed well away from the edges of contrasting parcels so that they do not encompass edge pixels.

Uniformity

Perhaps the most important property of a good training area is its uniformity, homogeneity. Data within each training area should exhibit a unimodel frequency distribution for each spectral band to be used.

Evaluating Signatures:

There are tests to perform that can help determine whether the signature data are a true representation of the pixels to be classified for each class. You can evaluate signatures

that were created either from supervised or unsupervised training. There are number of methods for evaluation of the signatures:

A. Graphical Method:

Ellipse—view ellipse diagrams and scatter plots of data file values for every pair of bands.

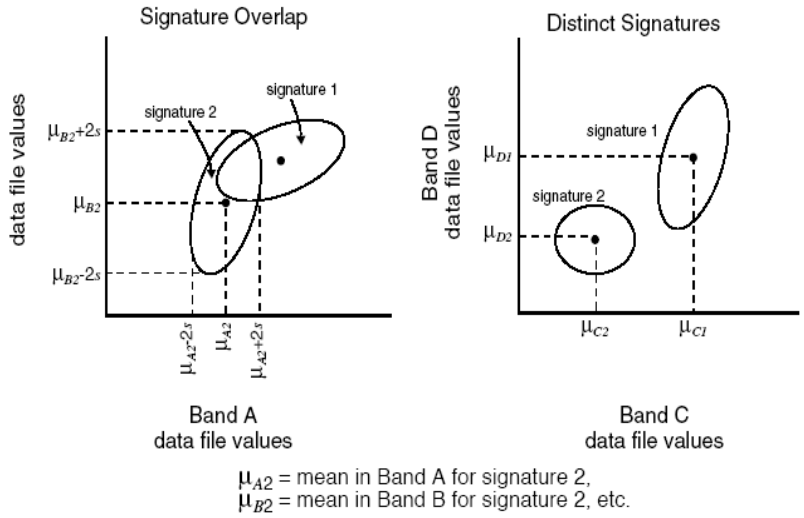


Figure 10 : Signature Separability

B. Signature Separability: **Signature separability** is a statistical measure of distance between two signatures. Separability can be calculated for any combination of bands that is used in the classification, enabling you to rule out any bands that are not useful in the results of the classification.

1. Euclidian Distance:

$$D = \sqrt{\sum_{i=1}^{n} (d_i - e_i)^2}$$

Where:
D = spectral distance
n = number of bands (dimensions)
i = a particular band
*d*i = data file value of pixel d in band *i*
*e*i = data file value of pixel *e* in band *i*

2. Divergence:

$$D_{ij} = \frac{1}{2}tr((C_i - C_j)(C_i^{-1} - C_j^{-1})) + \frac{1}{2}tr((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T)$$

Where:

i and j	=	the two signatures (classes) being compared
C_i	=	the covariance matrix of signature i
μ_i	=	the mean vector of signature i
tr	=	the trace function (matrix algebra)
T	=	the transposition function

Source: Swain and Davis, 1978

3. Transformed Divergence

$$D_{ij} = \frac{1}{2}tr((C_i - C_j)(C_i^{-1} - C_j^{-1})) + \frac{1}{2}tr((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T)$$

$$TD_{ij} = 2000\left(1 - \exp\left(\frac{-D_{ij}}{8}\right)\right)$$

Where:

i and j	=	the two signatures (classes) being compared
C_i	=	the covariance matrix of signature i
μ_i	=	the mean vector of signature i
tr	=	the trace function (matrix algebra)
T	=	the transposition function

Source: Swain and Davis, 1978

The scale of the divergence values can range from 0 to 2,000. As a general rule, if the result is greater than 1,900, then the classes can be separated. Between 1,700 and 1,900, the separation is fairly good. Below 1,700, the separation is poor (Jensen, 1996).

4. Jeffries-Matusita Distance

$$\alpha = \frac{1}{8}(\mu_i - \mu_j)^T\left(\frac{C_i + C_j}{2}\right)^{-1}(\mu_i - \mu_j) + \frac{1}{2}\ln\left(\frac{|(C_i + C_j)/2|}{\sqrt{|C_i| \times |C_j|}}\right)$$

$$JM_{ij} = \sqrt{2(1 - e^{-\alpha})}$$

Where:

i and j	=	the two signatures (classes) being compared
C_i	=	the covariance matrix of signature i
μ_i	=	the mean vector of signature i
ln	=	the natural logarithm function
$\lvert C_i \rvert$	=	the determinant of C_i (matrix algebra)

Source: Swain and Davis, 1978

Range of JM is between 0 and 1414. The JM distance has a saturating behavior with increasing class separation like transformed divergence. However, it is not as computationally efficient as transformed divergence" (Jensen, 1996).

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. It is instructive to review the logic of several of the classifiers. Among the most frequently used classification algorithms are the parallelepiped, minimum distance, and maximum likelihood decision rules.

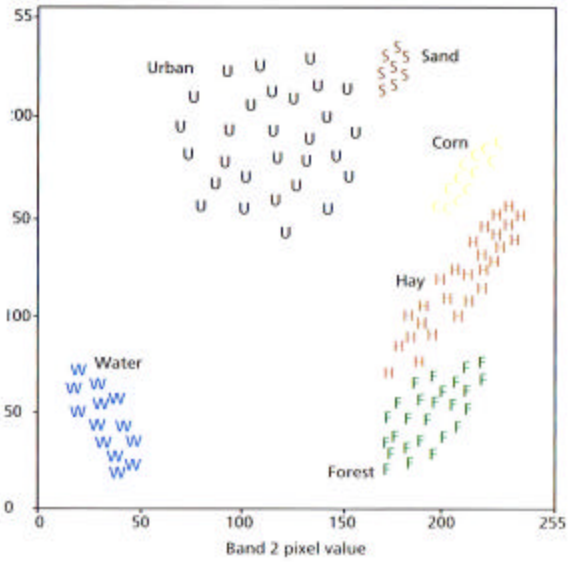


Figure 11 : The training Samples

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, $M_c = (\mu_{ck1}, \mu_{c2}, \mu_{c3}, \ldots \mu_{cn})$ 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.

Using a one-standard deviation threshold ,a parallelepiped algorithm decides BV_{ijk} is in class c if, and only if,

$$\mu_{ck} - S_{ck} \leq BV_{ijk} \leq \mu_{ck} + S_{ck}$$

where

c = 1, 2, 3, ..., m, number of classes
k = 1, 2, 3, ..., n, number of bands
Therefore, if the low and high decision boundaries are defined as

$L_{ck} = \mu_{ck} - S_{ck}$

And

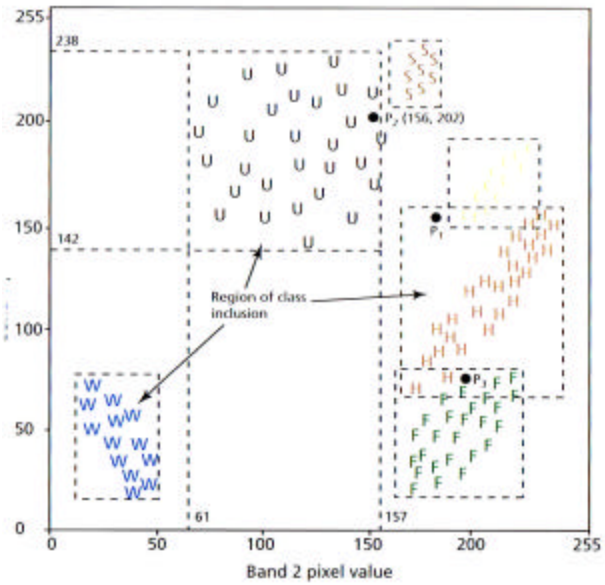
$H_{ck} = \mu_{ck} + + S_{ck}$

The parallelepiped algorithm becomes

$L_{ck} \leq BC_{ijk} \leq H_{ck}$

These 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 al n bands evaluated, it is assigned to an unclassified category. 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.



a b

Fig 12: Parallelepiped Classification Algorithm -B: Parallepipeds in the feature space

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

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

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.

When more than two bands are evaluated in a classification, it is possible to extend the logic of computing the distance between just two points in n space using the equation

$$Dab = \sqrt{\sum_{l=1}^{n} (a_i - b_i)^2}$$

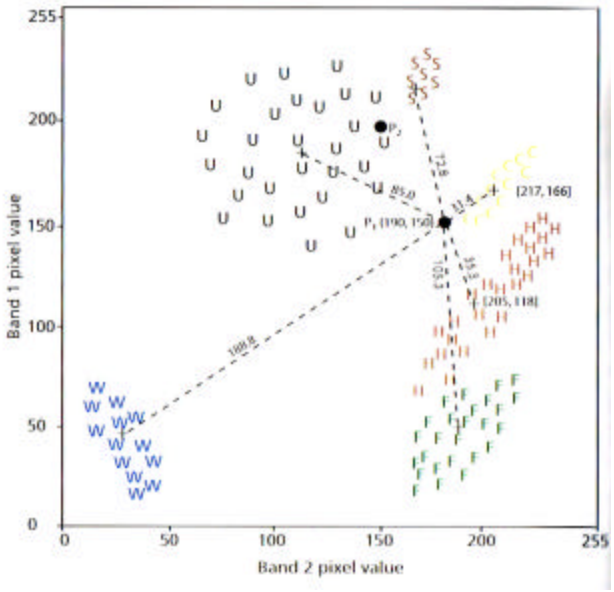


Fig 13: Minimum Distance to means Algorithm

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.

Maximum likelihood classification makes use of the statistics already computed and discussed in previous sections, including the mean measurement vector. M_c for each class and the covariance matrix of class c for bands k through l, V_c. The decision rule applied to the unknown measurement vector X is.

Decide X is in class c if, and only if,
$p_c \geq p_i$, where i = 1, 2, 3, ..., m possible classes
and
$p_c = \{-0.5\log_e[\det(V_c)]\} - [0.5(X-M_c)^T V_c^{-1} (X-M_c)]$
and det (V_c) is the determinant 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.

Now let us consider the computations required. In the first pass, p_1 is computed, with V_1 and M_1 being the covariance matrix and mean vectors for class 1. Next p_2 is computed using V_2 and M_2 This continues for all m classes. The pixel or measurement vector X is assigned to the class that produces the largest or maximum p_c. The measurement vector X used in each step of the calculation consists of n elements (the number of bands being analyzed). For example, if all six bands were being analyzed, each unknown pixel would have a measurement vector X of

$$X = \begin{matrix} BV_{i,j,l} \\ BV_{ij,2} \\ BV_{ij,3} \\ BV_{ij,4} \\ BV_{ij,5} \\ BV_{ij,6} \end{matrix}$$

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.

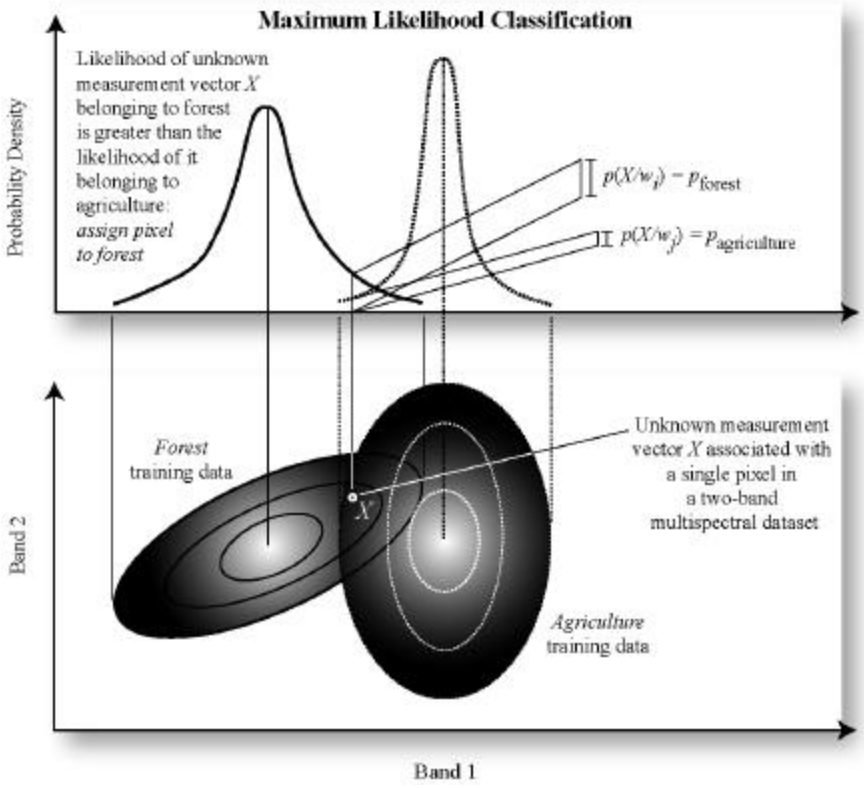


Figure 14: Maximum Likelihood Classification Algorithm

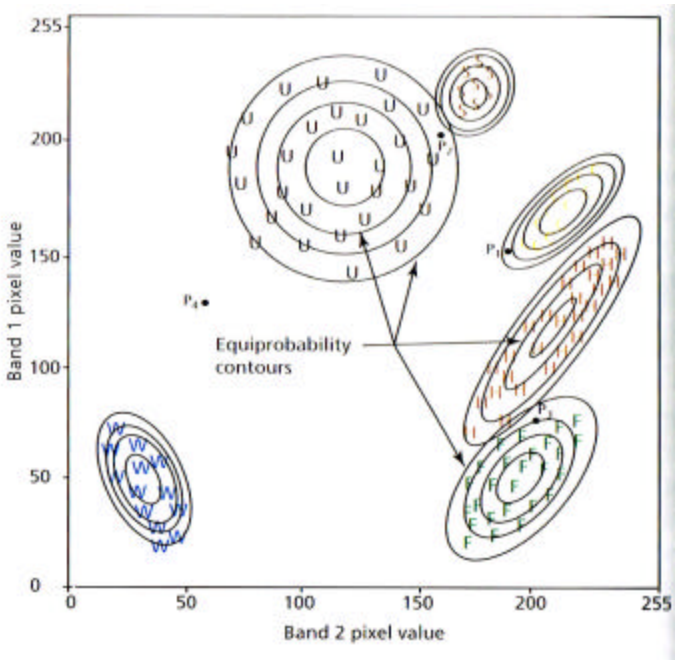


Figure 15 Equal Probability Contours

Feature Selection

In remote sensing the user may often have to work with multi-channel dataset to derive useful information. With advancement in data reception technology these volumes of data can often frustrate the user to select an optimal dataset. However, there are schemes which allow the user to select those dimensions most suitable for processing without unnecessarily increasing the computational time. This step is called feature selection or separability analysis. Selection of variables showing strong discriminatory power and hence to avoid redundancies in feature set is a key problem in pattern recognition and is termed as feature selection or feature extraction.

Various methods of feature selection schemes are used depending upon the classes and their statistics. If statistics are now known, then the best feature selection scheme is provided by principal component analysis. (For details refer to the Notes of Lecture "Bank Rationing and Principal Component Analysis"). If classes and their statistics are known then divergence analysis is carried out by variance covariance method, and then finding correlation matrix to find out, which bands are correlated, and which are not. The correlation of data among various bands can also be estimated graphically from the scatterplot of the data between two bands. In the process of feature selection, the highly correlated bands are rejected, and those with lesser or no correlation are selected for efficient analysis.

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 asses 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.

Overall Classification Map Accuracy Assessment

To determine the overall accuracy of a remotely sensed classified map it is necessary to ascertain whether the map meets or exceeds some predetermined classification accuracy criteria. Overall accuracy assessment evaluates the agreement between the two maps in total area or each category. They usually do not evaluate construction errors that occur in the various categories

Site Specific Classification Map Accuracy Assessment

This type of error analysis compares the accuracy of the remote sensing derived classification map pixel by pixel with the assumed true land use map. First, it is possible to

conduct a site-specific error evaluation based only on the training pixels used to train the classifier in a supervised classification. This simply means that those pixel locations i, j used to train the classifier are carefully evaluated on both the classified map from remote sensing data products and the assumed true map. If training samples are distributed randomly throughout the study area, this evaluation may consider representative of the study area. If they act biased by the analyst a prior knowledge of where certain land cover types exist in the scene. Because of this bias, the classification accuracy for pixels found within the training sites are generally higher than for the remainder of the map because these are the data locations that were used to train the classifier.

Conversely if others test locations in the study area are identified and correctly labeled prior to classification and if these are not used in the training of the classification algorithm they can be used to evaluate the accuracy of the classification map. This procedure generally yields a more credible classification accuracy assessment. However additional ground truth is required for these test site coupled with problem of determining how many pixels are necessary in each test site class. Also the method of identifying the location of the test sites prior to classification is important since many statistical tests require that locations be randomly selected (e .g using a random number generator for the identification off unbiased row and column coordinates) so that the analyst does not bias their selection.

Once the Criterion for objectively identifying the location of specific pixels to be compared is determined, it in necessary to identify the class assigned to each pixel in both the remote sensing derived map and the assumed true map. These data are tabulated and reported in a contingency table(error matrix), where overall classification accuracy and misclassification between categories are identified.

It takes the form of an m x m matrix, where m is the number of classes under investigation. The rows in the matrix represent the assumed true classes, while the columns are associated with the remote sensing derived land use. The entries in the contingency table represent the raw number of pixels encountered in each condition; however, they may be expressed as percentages, if the number becomes too large. One of the most important characteristics of such matrices is their ability to summarize errors of omission and commission. These procedures allow quantitative evaluation of the classification accuracy. Their proper use enhances the credibility, of using remote sensing derived land use information.

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)

		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**	

Table 1 Error Matrix resulting from classifying training Set pixels
Training set data (Reference)

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

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 accuracy's 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). Whereas the 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 represent that category on ground.

Note the error matrix in the table indicates an overall accuracy of 84%. However producers accuracy range from just 51%(urban) to 100% (water) and users accuracy range 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

Discrete multivariate techniques have been used to statistically evaluate the accuracy of remote sensing derived maps and error matrices since 1983 and are widely adopted. These techniques are appropriate as the remotely sensed data are discrete rather than continuous and are also binomially or multinomially distributed rather than normally distributed.

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

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

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.

Sample Considerations

Test areas of representative uniform land cover, which are different and more extensive than the sample from training sets. They are often located during the training stage of supervised classification by designating more candidate training areas than are actually needed to develop the classification statistics. A subset of these may then be withheld for the post classification accuracy assessment; the accuracies obtained in these areas represent at least a first approximation to classification performance through out the scene. However, being homogeneous test areas might not provide a valid indication of classification accuracy as the individual pixel level of land cover variability.

One way that would appear to ensure adequate accuracy assessment at the pixel level of specificity would be to compare the land cover classification at every pixel in an image with reference source. While such "wall to wall" comparisons may have value in research situations, assembling reference land cover information for an entire project area is expensive and defeats the whole purpose of performing a remote sensing based classification in the first place.

Random sampling of pixels circumvents the above problems, but it is plagued with its own set of limitations. First, collection of reference data for a large sample of randomly distributed points is often very difficulty and costly. For e.g. travel distance and access to random sites might be prohibitive. Second, the validity of random sampling depends on the ability to precisely register the reference data to the image data. This is often difficult to do. One way to overcome this problem is to sample only pixels whose identity is not influenced by potential registration errors (for example points at least several pixels away from field boundaries).

Another consideration is making certain that the randomly selected test pixels of areas are geographically representative of the data set under analysis. Simple random sampling tends to under sample small but potentially important areas. Stratified random sampling, where each land cover category may be considered a stratum, is frequently used in such cases. Clearly, the sampling approach appropriate for an agricultural inventory would from that of wetlands mapping activity. Each sample design must account for the area being studied and the Cover type being classified

One common means of accomplishing random sampling is to overlay classified output data with a grid. Test cell within those grids is then selected randomly and groups of pixels within the test cells are evaluated. The cover types present are determined through ground verification (of other reference) and compared to the classification data.

Several papers have been written about the proper sampling scheme to be used for accuracy assessment under various conditions and opinions vary among researchers

One suggestion his been the concept of combining both random and systematically sampling. Such a technique may use systematically sampled areas to collect some accuracy assessment data early in a project (perhaps as part of the training area selection process) and random sampling within strata after classification is complete.

Consideration must also be given to the sample unit employed in accuracy assessment. Depending upon the application, the appropriate sample unit might be individual pixel clusters or polygons. Sample size must also weigh heavily in the development and interpretation of classification accuracy figures. As a broad guideline it has been suggested that a minimum of 50 samples of each vegetation or landuse category can be included in the error matrix. If the area is too large i.e. more than a million hectares or the classification scheme has large number of landuse categories (more than 12) the minimum number of samples has to be increased to 75-100 samples per category. Based on the importance of a particular or depending upon the variability within that class, the number of samples in that class can be adjusted accordingly (e.g. more samples can be taken in more variable class such as wetlands and fewer in water i.e. less variable).

There are three other facets of classification accuracy. The first relates to the fact that the quality of any accuracy estimate is only as good as the information used to establish the true land cover types present in the test sites. It is not uncommon to have the accuracy of the reference data influenced by such factors as spatial misregistration, photo interpretation errors, data entry errors and changes in land cover between the date of the classified image and the date of the reference data. The second point being that the accuracy assessment procedure must be designed to reflect the intended use of the classification. For e.g. a single pixel misclassified as wetland in the midst of cornfield might be of little significant in the development of regional land use plan. However the same error might be intolerable if the classified map forms the basis for land taxation or fore-enforcement of wetland preservation legislation. Finally it should be noted that remotely sensed data is just a small subset of many possible forms of data resident in a GIS. How errors accumulate through the multiple layers of information in a GIS is the subject of ongoing research.