Chapter 6. Digital Image Processing

6.1 Introduction

In order to take advantage of and make good use of remote sensing data, we must be able to extract meaningful information from the imagery. This brings us to the topic of discussion in this chapter – interpretation and analysis - the sixth element (F) of the remote sensing process, which we defined in chapter 1 (Figure 1.1). Interpretation and analysis of remote sensing imagery involves the identification and/or measurement of various targets in an image in order to extract useful information about them. Targets in remote sensing images may be any feature or object which can be observed in an image, and have the following characteristics:

- Targets may be a point, line, or area feature. This means that they can have any form, from a bus in a parking lot or plane on a runway, to a bridge or road, to a large area of water or an agricultural field.
- The target must be distinguishable; it must contrast with other features around it in the image.

Much interpretation and identification of targets in remote sensing imagery is performed manually or visually, i.e. by a human interpreter. In many cases this is done using imagery displayed in a pictorial or photograph-type format, independent of what type of sensor was used to collect the data and how the data were collected. In this case we refer to the data as being in analogue format. Remote sensing images can also be represented in a computer as arrays of pixels, with each pixel corresponding to a digital number, representing the brightness level of that pixel in the image. In this case, the data are in a digital format. Visual interpretation may also be performed by examining digital imagery displayed on a computer screen. Both analogue and digital imagery can be displayed as black and white (also called monochrome) images, or as colour images by combining different spectral bands representing different wavelengths.

When remote sensing data are available in digital format, digital processing and analysis may be performed using a computer. Digital processing may be used to enhance data as a preparation to visual interpretation. Digital processing and analysis may also be carried out to automatically identify targets and extract information completely without manual intervention by a human interpreter. However, rarely is digital processing and analysis carried out as a complete replacement for manual interpretation. Often, it is done to supplement and assist the human analyst.

Manual interpretation and analysis dates back to the early beginnings of remote sensing for air photo interpretation. Digital processing and analysis is more recent with the advent of digital recording of remote sensing data and the development of computers. Both manual and digital techniques for interpretation of remote sensing data have their respective advantages and disadvantages. Generally, manual interpretation requires little, if any, specialised equipment, while digital analysis requires specialised, and often expensive, equipment and software. Manual interpretation is often limited to analysing only a single channel of data or a single image at a time due to the difficulty in performing visual interpretation with multiple images. The computer environment is more suitable for handling complex images of several channels or from several dates. In this sense, digital analysis is useful for simultaneous analysis of many spectral bands and can process large data sets much faster than a human interpreter. Manual interpretation is a subjective process, meaning that the results will vary with different interpreters. Digital analysis is based on the manipulation of digital numbers in a computer and is thus more objective, generally resulting in more consistent results. However, determining the validity and accuracy of the results from digital processing can be difficult.

It is important to re-iterate that visual and digital analyses of remote sensing imagery are not mutually exclusive. Both methods have their merits. In most cases, a mix of both methods is usually employed when analysing imagery. In fact, the ultimate decision of the utility and relevance of the information extracted at the end of the analysis process still must be made by humans.

6.2 Digital image processing

In today's world of advanced technology where most remote sensing data are recorded in digital format, virtually all image interpretation and analysis involves some element of digital processing. Digital image processing may involve numerous procedures including formatting and correcting of the data, digital enhancement to facilitate better visual interpretation, or even automated classification of targets and features entirely by computer. In order to process remote sensing imagery digitally, the data must be recorded and available in a digital form suitable for storage on a computer tape or disk. Obviously, the other requirement for digital image processing is a computer system, sometimes referred to as an image analysis system, with the appropriate hardware and software to process the data. Several commercially available software systems have been developed specifically for remote sensing image processing and analysis.

For discussion purposes, most of the common image processing functions available in image analysis systems can be categorised into the following four categories:

1. Preprocessing;
2. Image Enhancement;
3. Image Transformation;
4. Image Classification and Analysis.

Preprocessing functions involve those operations that are normally required prior to the main data analysis and extraction of information, and are generally grouped as radiometric or geometric corrections. Radiometric corrections include correcting the data for sensor irregularities and unwanted sensor or atmospheric noise, and converting the data so they accurately represent the reflected or emitted radiation measured by the sensor. Geometric corrections include correcting for geometric distortions due to sensor-Earth geometry variations, and conversion of the data to real world co-ordinates (e.g. latitude and longitude) on the Earth's surface.

The objective of the second group of image processing functions, grouped under the term of image enhancement, is solely to improve the appearance of the imagery to assist in visual interpretation and analysis. Examples of enhancement functions include contrast stretching to increase the tonal distinction between various features in a scene, and spatial filtering to enhance (or suppress) specific spatial patterns in an image.

Image transformations are operations similar in concept to those for image enhancement. However, unlike image enhancement operations, which are normally applied only to a single channel of data at a time, image transformations usually involve combined processing of data from multiple spectral bands. Arithmetic operations (i.e. subtraction, addition, multiplication, division) are performed to combine and transform the original bands into "new" images which better display or highlight certain features in the scene. Examples are various methods of spectral or band ratioing (e.g., for deriving vegetation indices, cf. chapter 2), and a procedure called principal components analysis which is used to more efficiently represent the information in multichannel imagery. However, also many spatial filtering techniques (e.g., for edge detection) and segmentation techniques belong to this category.

Calculation of vegetation indices (see chapter 2) is one of the most common transforms applied to image data. It serves to highlight subtle variations in the spectral responses of various surface covers. For instance, by ratioing the data from two different spectral bands, the resultant image enhances variations in the slopes of the spectral reflectance curves between the two different spectral ranges that may otherwise be masked by the pixel brightness variations in each of the bands. The following example illustrates the concept of spectral ratioing. Healthy vegetation reflects strongly in the near-infrared portion of the spectrum while absorbing strongly in the visible red. Other surface types, such as soil and water, show near equal reflectances in both the near-infrared and red portions. Thus, a ratio image of Landsat TM band 4 (near-infrared - 0.76 to 0.90 μm) divided by TM band 3 (red - 0.63 to 0.69 μm) would result in ratios much greater than 1.0 for vegetation, and ratios around 1.0 for soil and water. Thus, the discrimination of vegetation from other surface cover types is significantly enhanced. Also, we may be better able to identify areas of unhealthy or stressed vegetation, which show low near-infrared reflectance, as the ratios would be lower than for healthy green vegetation. Finally, a vegetation index can be used to estimate continuous variables like, e.g., LAI and biomass. So, we can produce a biomass image based on remote sensing observations and application of a simple (statistical) model.

Image classification and analysis operations are used to digitally identify and classify pixels in the data (Figure 6.1). Classification is usually performed on multi-channel data sets (A) and this process assigns each pixel in an image to a particular class or theme (B) based on statistical characteristics of the pixel brightness values. There are a variety of approaches taken to perform digital classification.

In the following sections we will only describe a few digital image processing techniques in more detail.

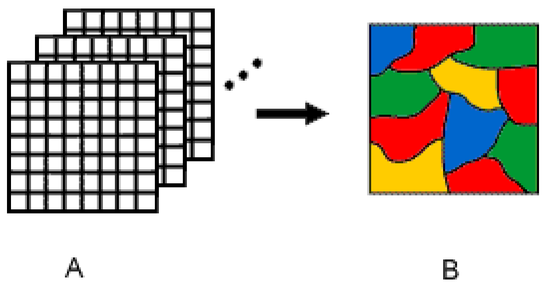

Figure 6.1: Image classification.

6.3 Image enhancement

Enhancements are used to make it easier for visual interpretation and understanding of imagery. The advantage of digital imagery is that it allows us to manipulate the digital pixel values in an image. Although radiometric corrections for illumination, atmospheric influences, and sensor characteristics may be done prior to distribution of data to the user, the image may still not be optimised for visual interpretation. Remote sensing devices, particularly those operated from satellite platforms, must be designed to cope with levels of target/background energy, which are typical of all conditions likely to be encountered in routine use. With large variations in spectral response from a diverse range of targets (e.g. forest, deserts, snowfields, water, etc.) no generic radiometric correction could optimally account for and display the optimum brightness range and contrast for all targets. Thus, for each application and each image, a customised adjustment of the range and distribution of brightness values is usually necessary.

In raw imagery, the useful data often covers only a small portion of the available range of digital values (commonly 8 bits or 256 levels). Contrast enhancement involves changing the original values so that more of the available range is used, thereby increasing the contrast between targets and their backgrounds. The key to understanding contrast enhancements is to understand the concept of an image histogram. A histogram is a graphical representation of the brightness values that comprise an image. The brightness values (i.e. 0-255) are displayed along the x-axis of the graph. The frequency of occurrence of each of these values in the image is shown on the y-axis (Figure 6.2).

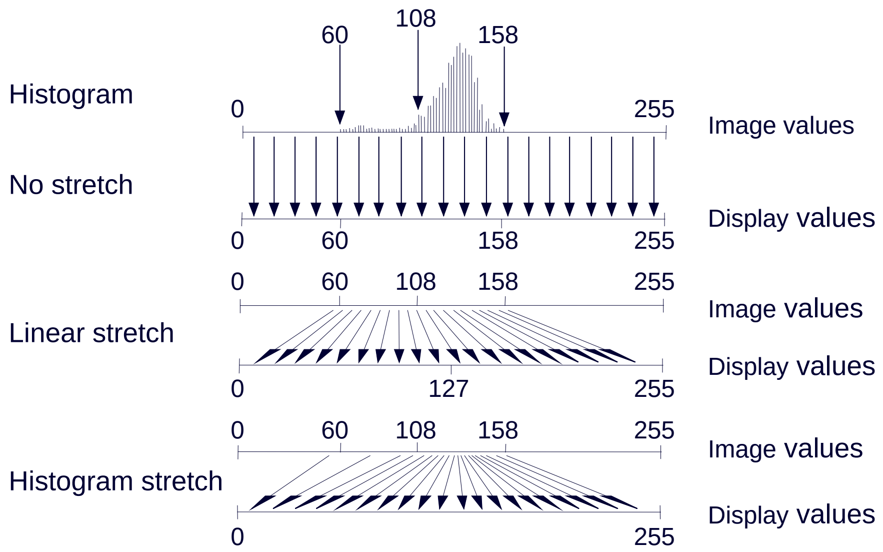

Figure 6.2: Principle of image enhancement.

By manipulating the range of digital values in an image, graphically represented by its histogram, we can apply various enhancements to the data. There are many different techniques and methods of enhancing contrast and detail in an image; we will cover only a few common ones here. The simplest type of enhancement is a linear contrast stretch (Figure 6.2). This involves identifying lower and upper bounds from the histogram (usually the minimum and maximum brightness values in the image) and applying a transformation to stretch this range to fill the full range. In our example, the minimum value (occupied by actual data) in the histogram is 60 and the maximum value is 158. These 99 levels occupy less than half of the full

256 levels available. A linear stretch uniformly expands this small range to cover the full range of values from 0 to 255. This enhances the contrast in the image with light toned areas appearing lighter and dark areas appearing darker, making visual interpretation much easier.

A uniform distribution of the input range of values across the full range may not always be an appropriate enhancement, particularly if the input range is not uniformly distributed. In this case, a histogram-equalised stretch may be better (Figure 6.2). This stretch assigns more display values (range) to the frequently occurring portions of the histogram. In this way, the detail in these areas will be better enhanced relative to those areas of the original histogram where values occur less frequently.

In other cases, it may be desirable to enhance the contrast in only a specific portion of the histogram. For example, suppose we have an image of the mouth of a river, and the water portions of the image occupy the digital values from 40 to 76 out of the entire image histogram. If we wished to enhance the detail in the water, perhaps to see variations in sediment load, we could stretch only that small portion of the histogram represented by the water (40 to 76) to the full grey level range (0 to 255). All pixels below or above these values would be assigned to 0 and 255, respectively, and the detail in these areas would be lost. However, the detail in the water would be greatly enhanced.

Of course, such image enhancement is not restricted to one spectral band, but can be applied to all spectral bands individually, yielding an enhanced colour composite image.

6.4 Image classification

6.4.1 Introduction

A human analyst attempting to classify features in an image uses the elements of visual interpretation to identify homogeneous groups of pixels, which represent various features or land cover classes of interest. Digital image classification (Figure 6.1) uses the spectral information represented by the digital numbers in one or more spectral bands, and attempts to classify each individual pixel based on this spectral information. This type of classification is termed spectral pattern recognition. In either case, the objective is to assign all pixels in the image to particular classes or themes (e.g. water, coniferous forest, deciduous forest, corn, wheat, etc.). The resulting classified image is comprised of a mosaic of pixels, each of which belong to a particular theme, and is essentially a thematic "map" of the original image.

When talking about classes, we need to distinguish between information classes and spectral classes. Information classes are those categories of interest that the analyst is actually trying to identify in the imagery, such as different kinds of crops, different forest types or tree species, different geologic units or rock types, etc. Spectral classes are groups of pixels that are uniform (or near-similar) with respect to their brightness values in the different spectral bands of the data. The objective is to match the spectral classes in the data to the information classes of interest. Rarely is there a simple one-to-one match between these two types of classes. Rather, unique spectral classes may appear which do not necessarily correspond to any information class of particular use or interest to the analyst. Alternatively, a broad information class (e.g. forest) may contain a number of spectral sub-classes with unique spectral variations. Using the forest example, spectral sub-classes may be due to variations in age, species, and density, or perhaps as a result of shadowing or variations in scene illumination. It is the analyst's job to decide on the utility of the different spectral classes and their correspondence to useful information classes.

Common classification procedures can be broken down into two broad subdivisions based on the method used: supervised classification and unsupervised classification. In a supervised classification, the analyst identifies in the imagery homogeneous representative samples of the different surface cover types (information classes) of interest. These samples are referred to as training areas. The selection of appropriate training areas is based on the analyst's familiarity with the geographical area and the knowledge of the actual surface cover types present in the image. Thus, the analyst is "supervising" the categorisation of a set of specific classes. The numerical information in all spectral bands for the pixels comprising these areas are used to "train" the computer to recognise spectrally similar areas for each class. The computer uses a special program or algorithm (of which there are several variations), to determine the numerical "signatures" for each training class. Once the computer has determined the signatures for each class, each pixel in the image is compared to these signatures and labelled as the class it most closely "resembles" digitally (see section 6.4.3). Thus, in a supervised classification we are first identifying the information classes, which are then used to determine the spectral classes to represent them.

Unsupervised classification in essence reverses the supervised classification process. Spectral classes are grouped first, based solely on the numerical information in the data, and are then matched by the analyst to information classes (if possible). Programs, called clustering algorithms, are used to determine the natural (statistical) groupings or structures in the data. Usually, the analyst specifies how many groups or clusters are to be looked for in the data. In addition to specifying the desired number of classes, the analyst may also specify parameters related to the separation distance among the clusters and the variation within each cluster. The final result of this iterative clustering process may result in some clusters that the analyst will want to subsequently combine, or clusters that should be broken down further - each of these requiring a further application of the clustering algorithm. Thus, unsupervised classification is not completely without human intervention. However, it does not start with a pre-determined set of classes as in a supervised classification.

6.4.2 Image space – feature space

Multispectral classification methods can be explained with an image for which radiation values have been recorded for each pixel in two different spectral bands. All pixels have their position in the image, which we will call the image space (Figure 6.3). Each pixel will have a digital pixel value (or digital number) in spectral band 1 and another value in band 2. These are called the features of a pixel. In Figure 6.3 (top part) only a small part of one band is depicted. In principle, this offers us the spatial patterns present in an image.

Another space may be defined by plotting the digital numbers in band 1 against those in band 2 (bottom part of Figure 6.3). This is called the feature space and the graphical representation is called a feature space plot. This offers us the spectral patterns present in an image. In the feature space of Figure 6.3 we see three clusters occurring. By assigning labels to these clusters according to some rule (algorithm, see next section) and then depicting their position again in the image space, we end up with a classified image. This is illustrated in Figure 6.4.

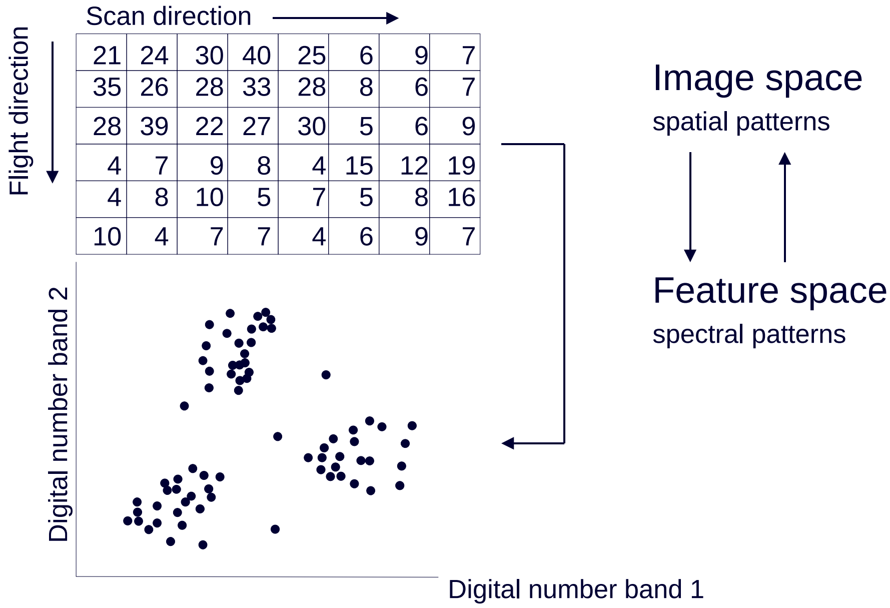

Figure 6.3: Illustration of the image space (top) and the feature space (bottom).

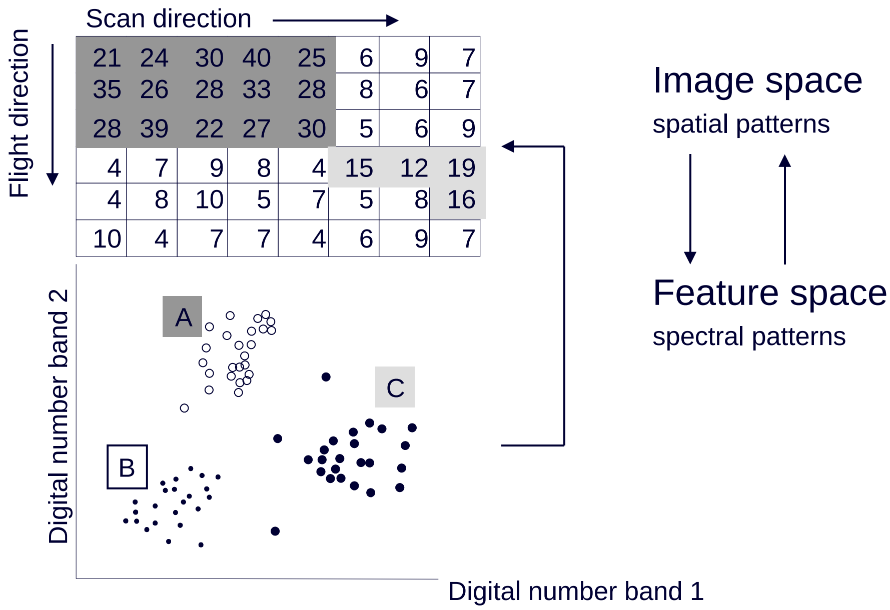

Figure 6.4: Label assignment in the feature space and presentation of the classified image in the image space.

6.4.3 Classification algorithms

In order to assign an unknown pixel to an existing cluster or class, a measure of distance must be defined as the distance of a point to a certain cluster. By calculating these distances for a certain pixel to various surrounding clusters, the pixel may be assigned to the nearest pixel. If all distances exceed a certain threshold, so that a pixel does not belong to any of the clusters, the pixel is not labelled or it gets a label like “unknown”. Four examples of decision rules will be given in this section.

(1) The "minimum distance to means" classification. The centre point of each cluster in the feature space is determined. The distances of a new point to these centres are calculated. The point is assigned to the cluster with the centre nearest to it (Figure 6.5).

(2) The "*k*-nearest neighbours" classification. From all clusters, *k* points per cluster are selected nearest to the point to be assigned. This new point is assigned to the cluster for which the average distance of *k* points to this new point is minimum (Figure 6.6). The number *k* can be selected, 1, 2 or 3, etc. Even if $k = 1$, the amount of computation in the *k*-nearest neighbours method is considerably larger than with the first-mentioned method, but the *k*-nearest neighbours method takes the forms and sizes of the clusters into account.

(3) Maximum likelihood classification. For each cluster, ellipses are drawn about the mean, dependent on the distribution of points in the cluster about that mean, assuming a Gaussian probability distribution. The probability of a point falling inside such an ellipse, if it actually belongs to that cluster, can be computed. The larger the ellipse the greater the probability a point falls inside the ellipse if it belongs to the cluster concerned, and the smaller the probability of falling outside. The more remote from the mean, the smaller the probability of finding a point that belongs to this cluster. For a new point to be assigned, this latter probability is computed for all surrounding clusters on the basis of its position in the feature space. The point concerned is assigned to the cluster with the greatest probability (likelihood) to have a point in that position (Figure 6.7).

(4) Parallelepiped (box) classification. Rectangular areas are drawn about the clusters according to the range of pixel values in each cluster. If a new point falls inside such a rectangle it is assigned to the cluster concerned (Figure 6.8). It is a fast classifier, but difficulties arise in overlapping zones.

These examples of decision rules presume the recording of two radiation values for each pixel. The line of thought may be extended to more radiation values (more spectral bands). The feature space becomes multi-dimensional, one additional axis for each spectral band. The computation formulae pertinent to the various decision rules may be extended to an arbitrary number of dimensions.

6.5 References and further reading

Chuvieco, E., 2016. Fundamentals of Satellite Remote Sensing: An Environmental Approach. Second Edition, CRC Press, Boca Raton, 468 pp.

Jensen, J.R., 2005. Introductory digital image processing. Third Edition, Prentice Hall, Upper Saddle River, NJ, 526 pp.

Liang, S., 2004. Quantitative remote sensing of land surfaces. John Wiley and Sons, Hoboken NJ, 534 pp.

Lillesand T.M., R.W. Kiefer & J.W. Chipman, 2008. Remote Sensing and Image Interpretation. Sixth Edition, John Wiley & Sons, Hoboken NJ. 756 pp.

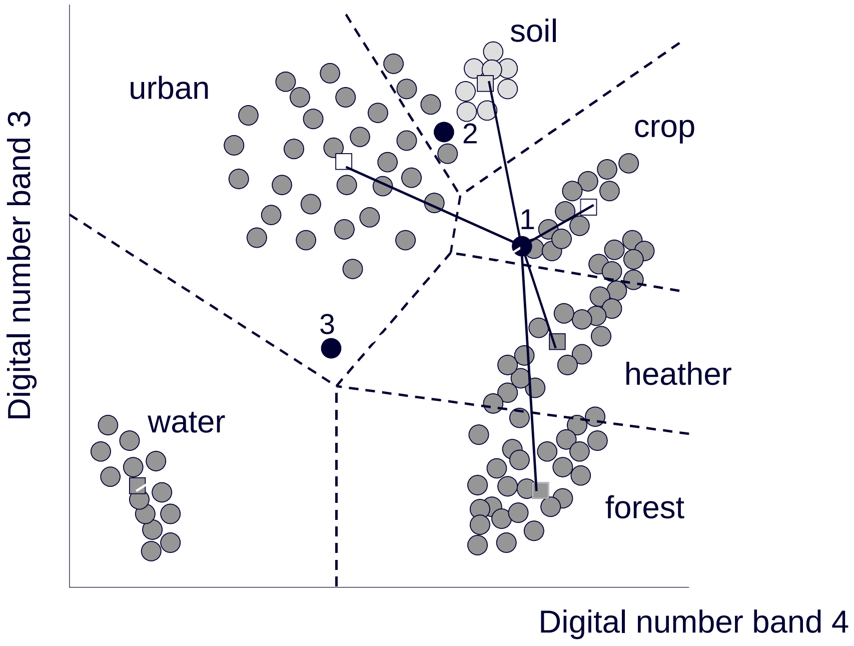

Figure 6.5: The minimum distance to means classification strategy. The decision boundaries are denoted by ----.

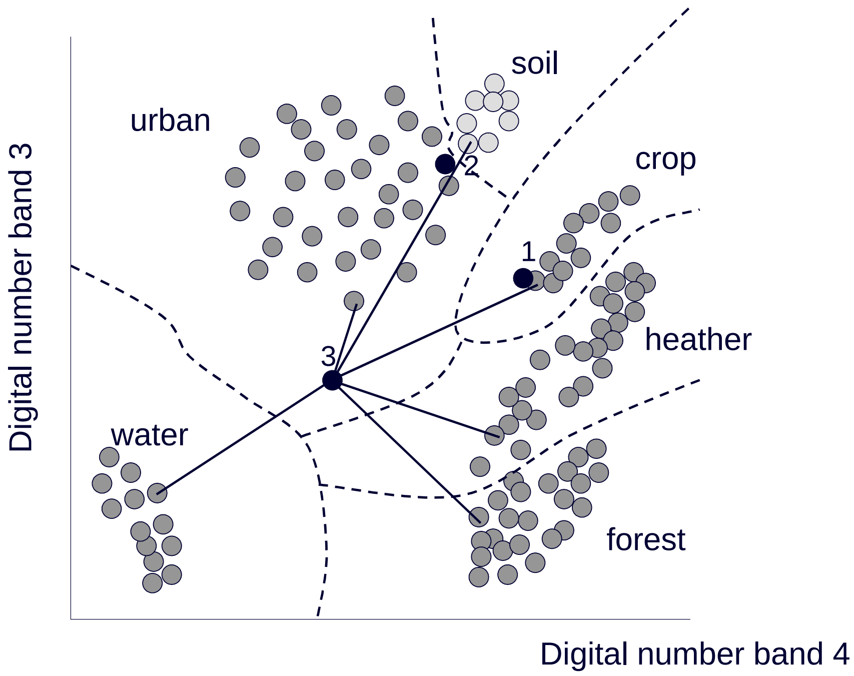

Figure 6.6: Nearest neighbour classification (k = 1) strategy. The decision boundaries are indicated by ----.

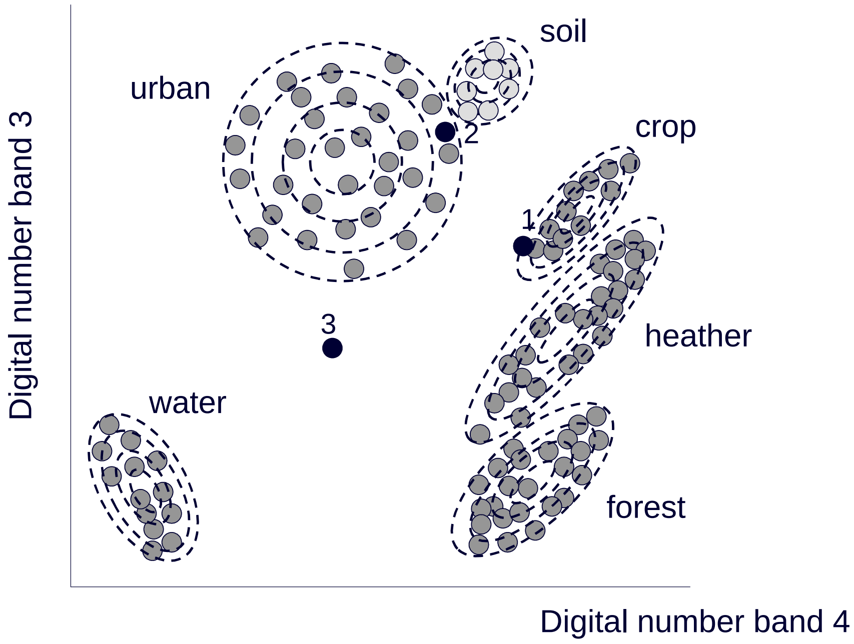

Figure 6.7: Maximum likelihood classification strategy. Ellipses of equal probability of a point falling inside such an ellipse are drawn for various clusters.

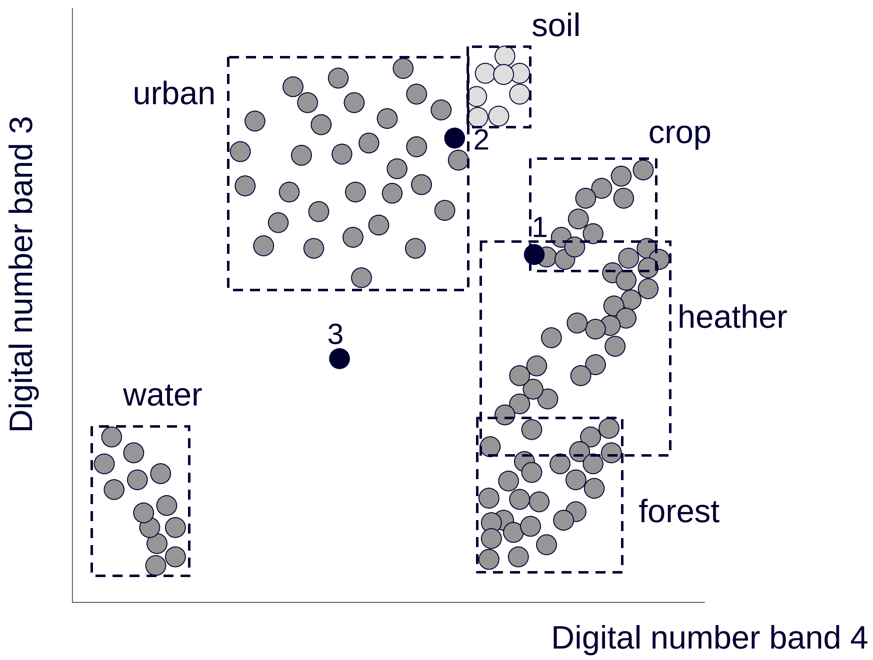

Figure 6.8: Parallelepiped classification strategy.