Chapter 1
Sources and Characteristics of Remote Sensing Image Data

1.1 Energy Sources and Wavelength Ranges

In remote sensing energy emanating from the earth’s surface is measured using a sensor mounted on an aircraft or spacecraft platform. That measurement is used to construct an image of the landscape beneath the platform, as depicted in Fig. 1.1.

In principle, any energy coming from the earth’s surface can be used to form an image. Most often it is reflected sunlight so that the image recorded is, in many ways, similar to the view we would have of the earth’s surface from an aircraft, even though the wavelengths used in remote sensing are often outside the range of human vision. The upwelling energy could also be from the earth itself acting as a radiator because of its own finite temperature. Alternatively, it could be energy that is scattered up to a sensor having been radiated onto the surface by an artificial source, such as a laser or radar.

Provided an energy source is available, almost any wavelength could be used to image the characteristics of the earth’s surface. There is, however, a fundamental limitation, particularly when imaging from spacecraft altitudes. The earth’s atmosphere does not allow the passage of radiation at all wavelengths. Energy at some wavelengths is absorbed by the molecular constituents of the atmosphere. Wavelengths for which there is little or no atmospheric absorption form what are called atmospheric windows. Figure 1.2 shows the transmittance of the earth’s atmosphere on a path between space and the earth over a very broad range of the electromagnetic spectrum. The presence of a significant number of atmospheric windows in the visible and infrared regions of the spectrum is evident, as is the almost complete transparency of the atmosphere at radio wavelengths. The wavelengths used for imaging in remote sensing are clearly constrained to these atmospheric windows. They include the so-called optical wavelengths covering the visible and infrared, the thermal wavelengths and the radio wavelengths that are used in radar and passive microwave imaging of the earth’s surface.

Whatever wavelength range is used to image the earth’s surface, the overall system is a complex one involving the scattering or emission of energy from the surface, followed by transmission through the atmosphere to instruments mounted
Fig. 1.1 Signal flow in a remote sensing system

Fig. 1.2 The electromagnetic spectrum and the transmittance of the earth’s atmosphere
on the remote sensing platform. The data is then transmitted to the earth’s surface, after which it is processed into image products ready for application by the user. That data chain is shown in Fig. 1.1. It is from the point of image acquisition onwards that this book is concerned. We want to understand how the data, once available in image format, can be interpreted.

We talk about the recorded imagery as *image data*, since it is the primary data source from which we extract usable information. One of the important characteristics of the image data acquired by sensors on aircraft or spacecraft platforms is that it is readily available in digital format. Spatially it is composed of discrete picture elements, or *pixels*. Radiometrically—that is in brightness—it is quantised into discrete levels.

Possibly the most significant characteristic of the image data provided by a remote sensing system is the wavelength, or range of wavelengths, used in the image acquisition process. If reflected solar radiation is measured, images can, in principle, be acquired in the ultraviolet, visible and near-to-middle infrared ranges of wavelengths. Because of significant atmospheric absorption, as seen in Fig. 1.2, ultraviolet measurements are not made from spacecraft altitudes. Most common optical remote sensing systems record data from the visible through to the near and mid-infrared range: typically that covers approximately 0.4–2.5 μm.

The energy emitted by the earth itself, in the thermal infrared range of wavelengths, can also be resolved into different wavelengths that help understand properties of the surface being imaged. Figure 1.3 shows why these ranges are important. The sun as a primary source of energy is at a temperature of about 5950 K. The energy it emits as a function of wavelength is described theoretically by Planck’s black body radiation law. As seen in Fig. 1.3 it has its maximal output at wavelengths just shorter than 1 μm, and is a moderately strong emitter over the range 0.4–2.5 μm identified earlier.

The earth can also be considered as a black body radiator, with a temperature of 300 K. Its emission curve has a maximum in the vicinity of 10 μm as seen in Fig. 1.3. As a result, remote sensing instruments designed to measure surface temperature typically operate somewhere in the range of 8–12 μm. Also shown in Fig. 1.3 is the blackbody radiation curve corresponding to a fire with a temperature of 1000 K. As observed, its maximum output is in the wavelength range 3–5 μm. Accordingly, sensors designed to map burning fires on the earth’s surface typically operate in that range.

The visible, reflective infrared and thermal infrared ranges of wavelength represent only part of the story in remote sensing. We can also image the earth in the microwave or radio range, typical of the wavelengths used in mobile phones, television, FM radio and radar. While the earth does emit its own level of microwave radiation, it is often too small to be measured for most remote sensing purposes. Instead, energy is radiated from a platform onto the earth’s surface. It is by measuring the energy scattered back to the platform that image data is recorded at microwave wavelengths. Such a system is referred to as *active* since the energy

---

source is provided by the platform itself, or by a companion platform. By comparison, remote sensing measurements that depend on an energy source such as the sun or the earth itself are called passive.

### 1.2 Primary Data Characteristics

The properties of digital image data of importance in image processing and analysis are the number and location of the spectral measurements (bands or channels), the spatial resolution described by the pixel size, and the radiometric resolution. These are shown in Fig. 1.4. Radiometric resolution describes the range and discernible number of discrete brightness values. It is sometimes referred to as dynamic range and is related to the signal-to-noise ratio of the detectors used. Frequently, radiometric resolution is expressed in terms of the number of binary digits, or bits, necessary to represent the range of available brightness values. Data with an 8 bit radiometric resolution has 256 levels of brightness, while data with 12 bit radiometric resolution has 4,096 brightness levels.²

---

² See Appendix B.
The size of the recorded image frame is also an important property. It is described by the number of pixels across the frame or *swath*, or in terms of the numbers of kilometres covered by the recorded scene. Together, the frame size of the image, the number of spectral bands, the radiometric resolution and the spatial resolution determine the data volume generated by a particular sensor. That sets the amount of data to be processed, at least in principle.

Image properties like pixel size and frame size are related directly to the technical characteristics of the sensor that was used to record the data. The *instantaneous field of view* (IFOV) of the sensor is its finest angular resolution, as shown in Fig. 1.5. When projected onto the surface of the earth at the operating altitude of the platform, it defines the smallest resolvable element in terms of equivalent ground metres, which is what we refer to as pixel size. Similarly, the *field of view* (FOV) of the sensor is the angular extent of the view it has across the earth’s surface, again as seen in
Fig. 1.5. When that angle is projected onto the surface it defines the swath width in equivalent ground kilometres. Most imagery is recorded in a continuous strip as the remote sensing platform travels forward. Generally, particularly for spacecraft programs, the strip is cut up into segments, equal in length to the swath width, so that a square image frame is produced. For aircraft systems, the data is often left in strip format for the complete flight line flown in a given mission.

1.3 Remote Sensing Platforms

Imaging in remote sensing can be carried out from both satellite and aircraft platforms. In many ways their sensors have similar characteristics but differences in their altitude and stability can lead to differing image properties.

There are two broad classes of satellite program: those satellites that orbit at geostationary altitudes above the earth’s surface, generally associated with weather and climate studies, and those which orbit much closer to the earth and that are generally used for earth surface and oceanographic observations. The low earth orbiting satellites are usually in a sun-synchronous orbit. That means that the orbital plane is designed so that it precesses about the earth at the same rate that the sun appears to move across the earth’s surface. In this manner the satellite acquires data at about the same local time on each orbit.

Low earth orbiting satellites can also be used for meteorological studies. Notwithstanding the differences in altitude, the wavebands used for geostationary and earth orbiting satellites, for weather and earth observation, are very comparable. The major distinction in the image data they provide generally lies in the spatial resolution available. Whereas data acquired for earth resources purposes has pixel sizes of the order of 10 m or so, that used for meteorological purposes (both at geostationary and lower altitudes) has a much larger pixel size, often of the order of 1 km.

The imaging technologies used in satellite remote sensing programs have ranged from traditional cameras, to scanners that record images of the earth’s surface by moving the instantaneous field of view of the instrument across the surface to record the upwelling energy. Typical of the latter technique is that used in the Landsat program in which a mechanical scanner records data at right angles to the direction of satellite motion to produce raster scans of data. The forward motion of the vehicle allows an image strip to be built up from the raster scans. That process is shown in Fig. 1.6.

Some weather satellites scan the earth’s surface using the spin of the satellite itself while the sensor’s pointing direction is varied along the axis of the satellite. The image data is then recorded in a raster scan fashion.

With the availability of reliable detector arrays based on charge coupled device (CCD) technology, an alternative image acquisition mechanism utilises what is

3 See www.jma.go.jp/jma/jma-eng/satellite/history.html.
commonly called a “push-broom” technique. In this approach a linear CCD imaging array is carried on the satellite normal to the platform motion as shown in Fig. 1.7. As the satellite moves forward the array records a strip of image data, equivalent in width to the field of view seen by the array. Each individual detector records a strip in width equivalent to the size of a pixel. Because the time over which energy emanating from the earth’s surface per pixel can be larger with push broom technology than with mechanical scanners, better spatial resolution is usually achieved.
Two dimensional CCD arrays are also available and find application in satellite imaging sensors. However, rather than record a two-dimensional snapshot image of the earth’s surface, the array is employed in a push broom manner; the second dimension is used to record simultaneously a number of different wavebands for each pixel via the use of a mechanism that disperses the incoming radiation by wavelength. Such an arrangement is shown in Fig. 1.8. Often about 200 channels are recorded in this manner so that the reflection characteristics of the earth’s surface are well represented in the data. Such devices are often referred to as imaging spectrometers and the data described as hyperspectral, as against multispectral when of the order of ten wavebands is recorded.

Aircraft scanners operate essentially on the same principles as those found with satellite sensors. Both mechanical scanners and CCD arrays are employed.

The logarithmic scale used in Fig. 1.3 hides the fact that each of the curves shown extends to infinity. If we ignore emissions associated with a burning fire, it is clear that the emission from the earth at longer wavelengths far exceeds reflected solar energy. Figure 1.9 re-plots the earth curve from Fig. 1.3 showing that there is continuous emission of energy right out to the wavelengths we normally associate with radio transmissions. In the microwave energy range, where the wavelengths are between 1 cm and 1 m, there is, in principle, measurable energy coming from the earth’s surface. As a result it is possible to build remote sensing instruments that form microwave images the earth. If those instruments depend on measuring the naturally occurring levels shown in Fig. 1.9, then the pixels tend to be very large because of the extremely low levels of energy available. Such large pixels are necessary to collect enough signal so that noise from the receiver electronics and the environment does not dominate the information of interest.

Fig. 1.8 Image formation by push broom scanning with an array that allows the recording of several wavelengths simultaneously
More often, we take advantage of the fact that the very low naturally occurring levels of microwave emission from the surface permits us to assume that the earth is, for all intents and purposes, a zero emitter. That allows us to irradiate the earth's surface artificially with a source of microwave radiation at a wavelength of particular interest. In principle, we could use a technique not unlike that shown in Fig. 1.6 to build up an image of the earth at that wavelength. Technologically, however, it is better to use the principle of synthetic aperture radar to create the image. We now describe that technique by reference to Fig. 1.10.

**Fig. 1.9** Illustration of the level of naturally emitted energy from the earth in the microwave range of wavelengths

**Fig. 1.10** Synthetic aperture radar imaging; as the antenna beam travels over features on the ground many echoes are received from the pulses of energy transmitted from the platform, which are then processed to provide a very high resolution image of those features.
A pulse of electromagnetic energy at the wavelength of interest is radiated to the side of the platform. It uses an antenna that produces a beam that is broad in the across-track direction and relatively narrow in the along-track direction, as illustrated. The cross track beamwidth defines the swath width of the recorded image. Features are resolved across the track by the time taken for the pulse to travel from the transmitter, via scattering from the surface, and back to the radar instrument. Along the track, features are resolved spatially using the principle of aperture synthesis, which entails recording many reflections from each spot on the ground and using signal processing techniques to synthesise high spatial resolution from a system that would otherwise record features at a detail too coarse to be of value. The technical details of how the image is formed are beyond the scope of this treatment but can be found in standard texts on radar remote sensing. What is important here is the strength of the signal received back at the radar platform since that determines the brightnesses of the pixels that constitute the radar image. As with optical imaging, the image properties of importance in radar imaging include the spatial resolution, but now different in the along and cross track directions, the swath width, and the wavebands at which the images are recorded.

Whereas there may be as many as 200 wavebands with optical instruments, there are rarely more than three or four with radar at this stage of our technology. However there are other radar parameters. They include the angle with which the earth’s surface is viewed out from the platform (the so-called look angle) and the polarisation of both the transmitted and received radiation. As a consequence, the parameters that describe a radar image can be more complex than those that describe an optical image. Nevertheless, once a radar image is available, the techniques of this book become relevant to the processing and analysis of radar image data. There are, however, some peculiarities of radar data that mean special techniques more suited to radar imagery are often employed.

1.4 What Earth Surface Properties are Measured?

In the visible and infrared wavelength ranges all earth surface materials absorb incident sunlight differentially with wavelength. Some materials detected by satellite sensors show little absorption, such as snow and clouds in the visible and near infrared. In general, though, most materials have quite complex absorption characteristics. Early remote sensing instrumentation, and many current instruments, do not have sufficient spectral resolution to be able to recognise the

---


5 See Richards, loc. cit., for information on image analysis tools specifically designed for radar imagery.
absorption spectra in detail, comparable to how those spectra might be recorded in a laboratory. Instead, the wavebands available with some detectors allow only a crude representation of the spectrum, but nevertheless one that is more than sufficient for differentiating among most cover types. Even our eyes do a crude form of spectroscopy by allowing us to differentiate earth surface materials by the colours we see, even though the colours are composites of the red, green and blue signals that reach our eyes after incident sunlight has scattered from the natural and built environment.

More modern instruments record many, sufficiently fine spectral samples over the visible and infrared range that we can get very good representations of reflectance spectra, as we will see in the following.
1.4.1 Sensing in the Visible and Reflected Infrared Ranges

In the absence of burning fires, Fig. 1.3 shows that the upwelling energy from the earth’s surface up to wavelengths of about 3 \( \mu m \) is predominantly reflected sunlight. It covers the range from the ultraviolet, through the visible, and into the infrared range. Since it is reflected sunlight the infrared is usually called reflected infrared, although it is then broken down into the near-infrared, short wavelength infrared and middle-infrared ranges. Together, the visible and reflected infrared ranges are called optical wavelengths as noted earlier. The definitions and the ranges shown in Fig. 1.3 are not fixed; some variations will be seen over different user communities.

Most modern optical remote sensing instrumentation operates somewhere in the range of 0.4–2.5 \( \mu m \). Figure 1.11 shows how the three broad surface cover types of vegetation, soil and water reflect incident sunlight over those wavelengths. In contrast, if we were to image a perfect reflector the reflection characteristics would be a constant at 100% reflectance over the range. The fact that the reflectance curves of the three fundamental cover types differ from 100% is indicative of the selective absorption characteristics associated with their biophysical and biochemical compositions.\(^6\) It is seen in Fig. 1.11 that water reflects about 10% or less in the blue-green range of wavelengths, a smaller percentage in the red and almost no energy at all in the infrared range. If water contains suspended sediments, or if a clear body of water is shallow enough to allow reflection from the bottom, then an increase in apparent water reflection will occur, including a small but significant amount of energy in the near infrared regime. That is the result of reflection from the suspension or bottom material.

Soils have a reflectance that increases approximately monotonically with wavelength, however with dips centred at about 1.4, 1.9 and 2.7 \( \mu m \) owing to moisture content. Those water absorption bands are almost unnoticeable in very dry soils and sands. In addition, clay soils have hydroxyl absorption bands at 1.4 and 2.2 \( \mu m \).

The vegetation curve is more complex than the other two. In the middle infrared range it is dominated by the water absorption bands near 1.4, 1.9 and 2.7 \( \mu m \). The plateau between about 0.7 and 1.3 \( \mu m \) is dominated by plant cell structure, while in the visible range of wavelengths plant pigmentation is the major determinant of shape. The curve shown in Fig. 1.11 is for healthy green vegetation. That has chlorophyll absorption bands in the blue and red regions leaving only green reflection of any significance in the visible. That is why we see chlorophyll pigmented plants as green. If the plant matter has different pigmentation then the shape of the curve in the visible wavelength range will be different. If healthy green vegetation dies the action of chlorophyll ceases and the absorption dips in the blue and red fill up, particularly the red. As a result, the vegetation appears yellowish, bordering on white when completely devoid of pigmentation.

---

Inspection of Fig. 11.1 shows why the wavebands for different remote sensing missions have been located in the positions indicated. They are arranged so that they detect those features of the reflectance spectra of earth surface cover types that are most helpful in discriminating among the cover types and in understanding how they respond to changes related to water content, disease, stage of growth and so on. In the case of the Hyperion instrument the number of wavebands available allows an almost full laboratory-like reconstruction of the reflectance spectrum of the earth surface material. We will see later in this book that such a rendition allows scientific spectroscopic principles to be used in analysing what the spectrum tells us about a particular point on the ground.

It is important to recognise that the information summarised in Fig. 1.11 refers to the reflection characteristics of a single pixel on the earth’s surface. With imaging spectrometers such as Hyperion we have the ability to generate full reflectance spectrum information for each pixel and, in addition, to produce a map showing the spatial distribution of reflectance information because of the lines and columns of pixels recorded by the instrument. With so many spectral bands available, we have the option of generating the equivalent number of images, or of combining the images corresponding to particular wavebands into a colour product that captures, albeit in a summary form, some of the spectral information. We will see in Chap. 3 how we cope with forming such a colour product.

Although our focus in this book will tend to be on optical remote sensing when demonstrating image processing and analysis techniques, it is of value at this point to note the other significant wavelength ranges in which satellite and aircraft remote sensing is carried out.

### 1.4.2 Sensing in the Thermal Infrared Range

Early remote sensing instruments that contained a thermal infrared band, such as the Landsat Thematic Mapper, were designed to use that band principally for measuring the earth’s thermal emission over a broad wavelength range. Their major applications tended to be in surface temperature mapping and in assessing properties that could be derived from such a measurement. If a set of spectral measurements is available over the wavelength range associated with thermal infrared emission, viz. 8–12 μm, thermal spectroscopic analysis is possible, allowing a differentiation among cover types.

If the surface being imaged were an ideal black body described by the thermal curve in Fig. 1.3 the upwelling thermal radiance measured by the satellite is proportional to the energy given by Planck’s radiation law. The difference between the radiation emitted by a real surface and that described by ideal black body behaviour is defined by the emissivity of the surface, which is a quantity equal to or less than one, and is a function of wavelength, often with strong absorption dips that correspond to diagnostic spectroscopic features. The actual measured upwelling radiance is complicated by the absorbing and emitting properties of the atmosphere; in practice they are removed by...
Fig. 1.12 Some emissivity spectra in the thermal infrared range; not to scale vertically. The quartz spectrum is, with permission of the IEEE, based on Fig. 1 of G.C. Hulley and S.J. Hook, Generating consistent land surface temperature and emissivity products between ASTER and MODIS data for earth science research, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 4, April 2011, pp. 1304–1315; the gypsum spectrum is, with permission of the IEEE, based on Fig. 1 of T. Schmugge, A. French, J. Ritchie, M. Chopping and A. Rango, ASTER observations of the spectral emissivity for arid lands, *Proc. Int. Geoscience and Remote Sensing Symposium*, vol. II, Sydney, Australia, 9–13 July 2001, pp. 715–717; the benzene spectrum was taken from D. Williams, Thermal multispectral detection of industrial chemicals, 2010, *personal communication*.

Fig. 1.13 a Ammonia spectrum recorded by the AHI thermal imaging spectrometer (Airborne Hyperspectral Imager, Hawaii Institute of Geophysics and Planetology (HGIP) at the University of Hawaii; this instrument has 256 bands in the range 8–12 µm) compared with a laboratory reference spectrum (reproduced with permission from D. Williams, Thermal multispectral detection of industrial chemicals, 2010, *personal communication*) and b ASTER multispectral thermal measurements of sand dunes compared with a laboratory sample (reproduced with permission of the IEEE from G.C. Hulley and S.J. Hook, Generating consistent land surface temperature and emissivity products between ASTER and MODIS data for earth science research, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 4, April 2011)
correction algorithms, as is the wavelength dependence of the solar curve. That allows
the surface properties to be described in terms of emissivity.

Figure 1.12 shows emissivity spectra in the thermal range for some common
substances. Also shown in figure are the locations of the wavebands for several
remote sensing instruments that take sets of measurements in the thermal region.
In Fig. 1.13 two examples are shown of identification in the thermal range, in one
case using a thermal imaging spectrometer to detect fine detail.

1.4.3 Sensing in the Microwave Range

As noted earlier, microwave, or radar, remote sensing entails measuring
the strength of the signal scattered back from each resolution element (pixel) on
the earth’s surface after irradiation by an energy source carried on the platform.
The degree of scattering is determined largely by two properties of the surface
material: its geometric shape and its moisture content. Further, because of the much
longer wavelengths used in microwave remote sensing compared with optical
imaging, some of the incident energy can penetrate beyond the outer surface of the
cover types being imaged. We will now examine some rudimentary scattering
behaviour so that a basic understanding of radar remote sensing can be obtained.

Smooth surfaces act as so-called *specular* (mirror-like) reflectors in that the
direction of scattering is predominantly away from the incident direction; as a result,
they appear dark to black in radar image data. Rough surfaces act as *diffuse*
reflectors in that they scatter the incident energy in all directions, including back towards the remote
sensing platform. Consequently they appear light in image data. Whether a surface is
regarded as rough or not depends on the wavelength of the radiation used and the angle
with which the surface is viewed (look angle). Table 1.1 shows the common fre-
quencies and wavelengths used with radar imaging. At the longer wavelengths many
surfaces appear smooth whereas the same surfaces can be diffuse shorter wavelengths,
as depicted in Fig. 1.14a. If the surface material is very dry then the incident micro-
wave radiation can penetrate, particularly at long wavelengths, as indicated in
Fig. 1.14b, making it possible to form images of objects underneath the earth’s surface.

Another surface scattering mechanism is often encountered with manufactured
features such as buildings. That is the *corner reflector effect* seen in Fig. 1.14c, which

<table>
<thead>
<tr>
<th>Band</th>
<th>Typical wavelength (cm)</th>
<th>Frequency (GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>6.67</td>
<td>0.45</td>
</tr>
<tr>
<td>L</td>
<td>23.5</td>
<td>1.28</td>
</tr>
<tr>
<td>S</td>
<td>12.6</td>
<td>2.38</td>
</tr>
<tr>
<td>C</td>
<td>5.7</td>
<td>5.3</td>
</tr>
<tr>
<td>X</td>
<td>3.1</td>
<td>9.7</td>
</tr>
<tr>
<td>Ku</td>
<td>2.16</td>
<td>13.9</td>
</tr>
</tbody>
</table>

Wavelength in metres and frequency in megahertz are related by the expression \( f(MHz) = \frac{300}{\lambda(m)} \)
results from the right angle formed between a vertical structure such as a fence, building or ship and a horizontal plane such as the surface of the earth or sea. This gives a very bright response; the response is larger at shorter wavelengths.

Media such as vegetation canopies and sea ice exhibit *volume scattering* behaviour, in that the backscattered energy emerges from many, hard-to-define sites within the volume, as illustrated for trees in Fig. 1.14d. That leads to a light tonal appearance in radar imagery, with the effect being strongest at shorter wavelengths. At long wavelengths vegetation offers little attenuation to the incident radiation so that the backscatter is often dominated by the surface underneath the vegetation canopy. Significant forward scattering can also occur from trunks when the vegetation canopy is almost transparent to the radiation at those longer wavelengths. As a consequence, the tree trunk can form a corner reflector in the nature of that shown in Fig. 1.14c.

The radar response from each of the geometric mechanisms shown in Fig. 1.14 is modulated by the moisture contents of the materials involved in the scattering process. Moisture enters through an electrical property called *complex permittivity* which determines the strength of the scattering from a given object or surface. The angle with which the landscape is viewed also has an impact on the observed level of backscatter. Scattering from relatively smooth surfaces is a strong function of look angle, while
scattering from vegetation canopies is weakly dependent on the look angle. Table 1.2 summarises the appearance of radar imagery in the different wavelength ranges.

We mentioned earlier that the radiation used with radar has a property known as polarisation. It is beyond the level of treatment here to go into depth on the nature of polarisation, but it is sufficient for our purposes to note that the incident energy can be called horizontally polarised or vertically polarised. Similarly, the reflected energy can also be horizontally or vertically polarised. For each transmission wavelength and each look angle, four different images can be obtained as a result of polarisation differences. If the incident energy is horizontally polarised, depending on the surface properties, the scattered energy can be either horizontally or vertically polarised or both, and so on.

Another complication with the coherent radiation used in radar is that the images exhibit a degree of “speckle”. That is the result of constructive and destructive interference of the reflections from surfaces that have random spatial variations of the order of one half a wavelength or so. Within a homogeneous region, such as a crop field, speckle shows up as a salt-and-pepper like noise that overlays the actual image data. It complicates significantly any analytical process we might devise for interpreting radar imagery that depends on the properties of single pixels.

### Table 1.2 Some characteristics of radar imagery

<table>
<thead>
<tr>
<th>Long wavelengths</th>
<th>Medium wavelengths</th>
<th>Short wavelengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Little canopy response but good tree response because of corner reflector effect involving trunks; good contrast of buildings and tree trunks against background surfaces, and ships at sea; good surface discrimination provided wavelength not too long.</td>
<td>Some canopy penetration; good canopy backscattering; fairly good discrimination of surface variations.</td>
<td>Canopy response strong, poor surface discrimination because diffuse scattering dominates; strong building response, but sometimes not well discriminated against adjacent surfaces</td>
</tr>
</tbody>
</table>

![Fig. 1.15](image-url) A typical registered spatial data set such as might be found in a GIS; some data types are inherently numerical while others are often in the form of labels.
1.5 Spatial Data Sources in General and Geographic Information Systems

Other sources of spatial data exist alongside satellite or aircraft remote sensing imagery, as outlined in Fig. 1.15. They include simple maps that show topography, land ownership, roads and the like, and more specialised sources such as geological maps and maps of geophysical measurements such as gravimetrics and magnetics. Spatial data sets like those are valuable complements to image data when seeking to understand land cover and land use. They contain information not available in remote sensing imagery and careful combinations of spatial data sources often allow inferences to be drawn about regions on the earth surface not possible when using a single source on its own.

In order to be able to process any spatial data set using the digital image processing techniques treated in this book, the data must be available in discrete form spatially and radiometrically. In other words it must consist of or be able to be converted to pixels, with each pixel describing the properties of a small region on the ground. The value ascribed to each pixel must be expressible in digital form. Also, when seeking to process several spatial data sets simultaneously they must be in correct geographic relation to each other. Desirably, the pixels in imagery and other spatial data should be referenced to the coordinates of a map grid, such as the UTM grid system. When available in this manner the data is said to be geocoded. Methods for registering and geocoding different data sets are treated in Chap. 2.

The amount and variety of data to be handled in a database that contains imagery and other spatial data sources can be enormous, particularly if it covers a large geographical region. Clearly, efficient means are required to store, retrieve, manipulate, analyse and display relevant data sets. That is the role of the geographic information system (GIS). Like its commercial counterpart, the management information system (MIS), the GIS is designed to carry out operations on the data stored in its data base according to a set of user specifications, without the user needing to be knowledgeable about how the data is stored and what data handling and processing procedures are utilised to retrieve and present the data.

Because of the nature and volume of data involved in a GIS many of the MIS concepts are not easily transferred to GIS design, although they do provide guidelines. Instead, new design concepts have evolved incorporating the sorts of operation relevant to spatial data. Attention has had to be given to efficient coding techniques to facilitate searching through the large numbers of maps and images often involved. That can be performed using procedures known collectively as data mining.\textsuperscript{7}

To understand the sorts of spatial data manipulation operations of importance in GIS one must take the view of the resource manager rather than the data analyst.

\textsuperscript{7} There is a special section on data mining in \textit{IEEE Transactions on Geoscience and Remote Sensing}, vol. 45, no. 4, April 2007. The Introduction, in particular, gives a good description of the field.
While the latter is concerned with image reconstruction, filtering, transformation and classification, the manager is interested in operations such as those listed in Table 1.3. They provide information from which management strategies and the like can be inferred. To be able to implement many, if not most, of those, a substantial amount of image processing is needed. It is expected, though, that the actual image processing being performed would be largely transparent to the resource manager; the role of the data analyst will often be in the design of the GIS system.

### 1.6 Scale in Digital Image Data

Because of IFOV differences the images provided by different remote sensing sensors are confined to application at different scales. As a guide, Table 1.4 relates scale to spatial resolution. That has been derived by considering an image pixel to be too coarse if it approaches 0.1 mm in size on a photographic product at a given scale.
Landsat ETM+ data is seen to be suitable for scales smaller than about 1:250,000 whereas MODIS imagery is suitable for scales below about 1:10,000,000.

### 1.7 Digital Earth

For centuries we depended on the map sheet as the primary descriptor of the spatial properties of the earth. With the advent of satellite remote sensing in the late 1960s and early 1970s we then had available for the first time wide scale and panoramic earth views that supplemented maps as a spatial data source. Over the past four decades, with increasing geometric integrity and spatial resolution, satellite and aircraft imagery, along with other forms of spatial data, led directly to the construction of the GIS, now widely used as a decision support mechanism in many resource-related studies.

In the past decade the GIS notion has been generalised substantially through the introduction of the concept of the virtual globe.\(^8\) This allows the user of spatial data to roam over the whole of the earth’s surface and zoom in or out to capture a view at the scale of interest. Currently, there are significant technical limitations to the scientific use of the virtual globe as a primary mapping tool, largely to do with the radiometric and positional accuracy but, with further development, the current GIS model will be replaced by a virtual globe framework in which, not only positional and physical descriptor information will be available, but over which will be layers of other data providing information on social, cultural, heritage and human factors. Now known as *digital earth*, such a model puts spatial information and its manipulation in the hands of anyone with a simple home computer; in addition, it allows the non-scientific lay user the opportunity to contribute to the information estate contained in the digital earth model. Citizen contribution of spatial data goes under the name of *crowdsourcing*, or sometimes *neogeography*, and will be one of the primary data acquisition methodologies of the future.

When combined with the enormous number of ground-based and spaceborne/airborne sensors, forming the sensor web,\(^9\) the digital earth concept promises to be an enormously powerful management tool for almost all of the information of value to us both for scientific and other purposes. The idea of the digital earth formed after a seminal speech given by former US Vice-President Al Gore in 1998.\(^{10}\)

---

\(^8\) Perhaps the best-known examples are Google\(^\text{®}\) Earth and NASA’s World Wind.


The digital earth paradigm is illustrated in Fig. 1.16. To make that work many of the image processing and analysis techniques presented in later chapters need to be employed.

Fig. 1.16  Digital earth, showing the types of data gathering, the dependence on computer networks and social media, the globe as the reference framework, and the concept of inserting value-added products back into the information base, either as free goods or commercially available

1.8 How This Book is Arranged

The purpose of this chapter has been to introduce the image data sources that are used in remote sensing and which are the subject of the processing operations described in the remainder of the book. It has also introduced the essential characteristics by which digital image data is described. The remainder of the book is arranged in a sequence that starts with the recorded image data and progresses through to how it is utilised.

The first task that normally confronts the analyst, before any meaningful processing can be carried out, is to ensure as much as possible that the data is free of error, both in geometry and brightness. Chapter 2 is dedicated to that task and the associated operation of registering images together, or to a map base. At the end of that chapter we assume that the data has been corrected and is ready for analysis.

Chapter 3 then starts us on the pathway to data interpretation. It is an overview that considers the various ways that digital image data can be analysed, either manually or with the assistance of a computer. Such an overview is important because there is no single, correct method for undertaking image interpretation; it is therefore important to know the options available before moving into the rest of the book.

It is frequently important to produce an image from the recorded digital image data, either on a display screen or in hard copy format. That is essential when
analysis is to be carried out using the visual skills of a human interpreter. Even when machine analysis is to be performed the analyst will still produce image products, most likely on a screen, to assist in that task. To make visual interpretation and recognition as easy as possible it is frequently necessary to enhance the visual appeal of an image. Chapter 4 looks at methods for enhancing the radiometric (brightness and contrast) properties of an image. It also looks at how we might join images side-by-side to form a mosaic in which it is necessary to minimise any brightness differences across the join.

The visual impact of an image can be also improved through operations on image geometry. Such procedures can be used to enhance edges and lines, or to smooth noise, and are the subject of Chap. 5. In that chapter we also look at geometric processing operations that contribute to image interpretation.

In Chap. 6 we explore a number of transformations that generate new versions of images from the imagery recorded by remote sensing platforms. Chief among these is the principal components transformation, well-regarded as a fundamental operation in image processing.

Several other transformations are covered in Chap. 7. The Fourier transform and the wavelet transform are two major tools that are widely employed to process image data in a range of applications. They are used to implement more sophisticated filtering operations than are possible with the geometric procedures covered in Chap. 5, and to provide means by which imagery can be compressed into more efficient forms for storage and transmission.

At this stage the full suite of so-called image enhancement operations has been covered and the book moves its focus to automated means for image interpretation. Many of the techniques now to be covered come from the field of machine learning.

Chapter 8 is central to the book. It is a large chapter because it covers the range of machine learning algorithms commonly encountered in remote sensing image interpretation. Those techniques are used to produce maps of land cover, land type and land use from the data recorded by a remote sensing mission. At the end of this chapter the reader should understand how data, once corrected radiometrically and geometrically, can be processed into viable maps by making use of a small number of pixels for which the appropriate ground label is known. Those pixels are called training pixels because we use them to train the machine learning technique we have chosen to undertake the full mapping task. The techniques treated come under the name of supervised classification.

On occasions the user does not have available known samples of ground cover over which the satellite data is recorded—in other words there are no training pixels. Nevertheless, it is still possible to devise machine learning techniques to label satellite data into ground cover types. Chapter 9 is devoted to that task and covers what are called unsupervised classification and clustering.

Frequently we need to reduce the volume of data to be processed, generally by reducing the number of bands. That is necessary to keep processing costs in bounds, or to ensure some analysis algorithms operate effectively. Chapter 10 presents the techniques commonly used for that purpose. Two approaches are
presented: one involves selecting optimal subsets of the existing bands, while the
other entails transforming the data beforehand in an endeavour to make the task of
discarding less useful (transformed) bands easier.

Chapter 11 brings together much of the material on classification and machine
learning into a set of methodologies that are used to produce reliable classification
and mapping results. Included here are the methods used to assess the accuracy of
a classification exercise.

In Chap. 12 we look at techniques for performing a classification when several
different types of image or spatial data are available. Those procedures can be
numerical or statistical but can also be based on expert system methodologies.

A set of appendices is given to provide supplementary material, including
background mathematics and statistics.

1.9 Bibliography on Sources and Characteristics
of Remote Sensing Image Data

This book is principally about the computer processing of remote sensing image data
and is not a detailed treatment of remote sensing as a field. Should more background
in remote sensing be needed then one of the standard treatments, such as

IL, 2007,

is recommended. Highlighted below are a number of image processing and
analysis texts that will add further detail to the coverage in this book, often at a
higher mathematical level. One of the first comprehensive texts on the computer
processing of remotely sensed imagery is


Even though much of the material has now been supplemented, this standard book
still has one of the best chapters on the spectral reflectance characteristics of earth
surface cover types, information that is essential to understand when carrying out
image interpretation. For examples of thermal spectral emission properties see

D.J. Williams, A.N. Pilant, D.D. Worthy, B. Feldman, T. Williams and P. Lucey,
Detection and identification of toxic air pollutants using airborne LWIR hyperspectral
imaging, *SPIE 4th Int. Asia–Pacific Environmental Remote Sensing Symposium*, Honolu-
lulu, Hawaii, 8–11 November 2004, vol. 5655, pp. 1–8, 2005, and

G.C. Hulley and S.J. Hook, Generating consistent land surface temperature and emissivity
products between ASTER and MODIS data for earth science research, *IEEE Transactions on

A standard treatment on image enhancement procedures, both radiometric and
geometric, is
Gonzalez and Woods was first published in 1977 as Gonzalez and Wintz, and has been through several editions (and changes of author) since, each time revising and adding new material. There is a companion volume, showing how the various techniques can be implemented in Matlab®:


A simpler treatment of digital image processing techniques will be found in


Like digital image processing there are many specialised and general texts on the pattern recognition or machine learning techniques that are fundamental to remote sensing image interpretation. Perhaps the most commonly used, with a broad coverage of techniques, is


A more recent and mathematically detailed treatment is


When the number of spectral bands recorded by a sensor exceeds about 100 there are special challenges for computer image interpretation. A book devoted to that problem is


For a treatment of radar remote sensing that assumes little prior knowledge see


For information on developments in the digital earth concept see


A good mathematical companion on matrix methods is


The broad topic of image compression is not treated in this book. The following is recommended for readers interested in that material.

1.10 Problems

1.1 Suppose a given set of image data consists of just two bands, one centred on 0.65 \( \mu \)m and the other centred on 1.0 \( \mu \)m wavelength. Suppose the corresponding region on the earth’s surface consists of water, vegetation, and soil.

Construct a graph with two axes, one representing the brightness of a pixel in the 0.65 \( \mu \)m band and the other representing the brightness of the pixel in the 1.0 \( \mu \)m band. Show on this graph (which we might call the spectral space or spectral domain) where you would expect to find vegetation pixels, soil pixels and water pixels. Indicate how straight lines could, in principle, be drawn between the three groups of pixels so that, if a computer had the equations of those lines stored in its memory, it could use them to identify every pixel in the image.

Repeat the exercise for an image data set with bands centred on 0.95 \( \mu \)m and 1.05 \( \mu \)m.

1.2 Assume a frame of image data consists of a segment along the track of the satellite as long as the swath is wide. Compute the data volume of a single frame from each of the following sensors and produce a graph of average data volume per band versus pixel size.

- NOAA AVHRR
- Aqua MODIS
- Landsat ETM+
- SPOT HRG multispectral
- GeoEye multispectral

1.3 Determine a relationship between swath width and orbital repeat cycle for a polar orbiting satellite at an altitude of 800 km, assuming that adjacent swaths overlap by 10% at the equator.

1.4 A particular geosynchronous satellite is placed in orbit over the poles, rather than over the equator. How often does it appear over the same spot on the earth’s surface, and where is that?

1.5 Geostationary satellites, with time, wander from the equatorial plane and have to be repositioned; this is called station keeping. If you could see the satellite overhead (for example, through a powerful telescope) what would the satellite path look like before correction?

1.6 Reconsider Problem 1.1 but instead of drawing lines between the classes in the spectral domain consider instead how you might differentiate them by computing the mean position of each class.

1.7 Imagine a scanner of the type shown in Fig. 1.6 is carried on an aircraft. In flight the aircraft can unintentionally but slowly change altitude and can be subject to cross winds. The pilot would normally compensate for the cross wind by steering into it. Describe the effect of these two mechanisms on the geometry of the recorded image data.
1.8 Using the results in Appendix A calculate the frame acquisition time for the following satellite sensors.

- SPOT HRG
- NOAA AVHRR
- WorldView Pan

1.9 Most remote sensing satellites are in orbits that pass over the regions being imaged about mid-morning. Why is that important?

1.10 Derive a relationship between repeat cycle and swath width for a remote sensing satellite with an orbital period of 90 min, assuming a near-polar orbit. Choose swath widths between 50 and 150 km. For a swath width of 100 km how could a repeat cycle of 10 days be achieved? Would several satellites be a solution?

1.11 By examining Figs. 1.11 and 1.14 discuss how the combination of optical and radar imagery might improve the recognition of ground cover types.

1.12 Discuss the relative advantages of satellite and aircraft platforms for remote sensing image acquisition.

1.13 A particular satellite at an altitude of 800 km in near polar orbit carries a high resolution optical sensor with 1 m spatial resolution. If the orbit is arranged so that complete earth coverage is possible, how long will that take if there are 2,048 pixels per swath width? See Appendix A for the relationship between altitude and orbital period.

1.14 A particular sensor records data in two wavebands in which the radiometric resolution is just 2 bits (see Appendix B). What is the theoretical maximum number of cover types that can be differentiated with the sensor? Show that if a sensor has \( c \) channels and a radiometric resolution of \( b \) bits the total number of sites in the corresponding spectral domain (see Problem 1.1) is \( 2^{bc} \). How many different sites are there for the following sensors?

- SPOT HRV
- Landsat ETM+
- EO-1 Hyperion
- Ikonos

For an image of 512 × 512 pixels how many sites, on the average, will be occupied for each of these sensors?

1.15 Why is earth imaging from satellites not carried out at wavelengths of about 1 mm?

1.16 What imaging wavelengths would you use to map a fire burning on the earth’s surface, superimposed on general landscape features?

1.17 Discuss the concept of map scale in the context of a virtual globe in the digital earth paradigm.

1.18 Many general purpose radar satellites operate at C band. Why?