The core of GIScience

a process-based approach

ITC Educational textbook series

UNIVERSITY OF TWENTE

FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Chapter 5

Pre-processing

Wan Bakx
Ben Gorte
Wim Feringa
Karl Grabmaier
Lucas Janssen
Norman Kerle
Gabriel Parodi
Anupma Prakash
Ernst Schetselaar
Klaus Tempfli
Michael Weir

5.1 Visualization and radiometric operations

This section explains the processing of raw remote sensing data and the basic principles of visualization of data. The production of images on a computer monitor or paper print-out has always been a prime concern of RS. We use images for inspecting raw data and for performing various data rectification and restoration tasks. Once data are corrected, we convert them once more to images and use these for information extraction by visual interpretation or to support digital image analysis. Many RS applications make use of multispectral data; to visualize them we have to rely on the use of colour. Section 5.1.1 explains how we perceive colour, which can help us to understand how to produce optimal images from multispectral data and how to properly interpret them.

We try to build remote sensors SO that they faithfully image a scene, and we are increasingly successful in doing so. Consider as an example a vertical photograph (or nadir view) of high resolution from a space-borne sensor: it closely resembles a map of the scene and, if the scene was a city, urban planners would be able to readily recognize objects of interest. Taking a closer look, we know that RS images will be geometrically distorted as compared to a map. The degree and type of distortion depends on the type of sensor platform used. Geometrical correction of RS images will be treated in Section 5.3.

In Chapter 2 you have learned that remote sensors measure radiances. The results of those measurements, however, are recorded as digital numbers, which have no direct physical meaning. The degree to which DNs directly correspond to radiances on the ground depends on many factors. Degradation with respect to what we would like to measure is caused, for example, by unfavourable scene illumination, atmospheric scattering and absorption, and detector-response characteristics. The need to perform radiometric correction in order to compensate for any or all types of degradation depends on the intended use of the data. Urban planners or topographic mappers do not need radiances of objects to be able to recognize them in images. Nevertheless, these images are likely to benefit from "haze correction" and contrast enhancement, to facilitate interpretation. Subsection 5.1.3 therefore briefly treats radiometric correction, only covering corrective measures that are of interest to a wider range of disciplines. (Image restoration and atmospheric correction are discussed further in Section 5.2.) A more detailed description of visualization for map production and spatial analysis is given in Chapter 10.

Elementary image processing techniques to improve the visual quality of an image—so that interpretation becomes easier—are introduced in Section 5.1.4. Image enhancement is not only useful for Earth observation: you may even find it handy for "touching up" your own digital photos.

5.1.1 Visualization

Perception of colour

The perception of colour takes place in the human eye and associated part of the brain. Colour perception concerns our ability to identify and distinguish colours, which in turn enables us to identify and distinguish entities in the real world. It is not completely known how human vision works, or what exactly happens in the eyes and brain before someone decides that an object is, for example, dark blue. Some theoretical models, supported by experimental results, are generally accepted, however. Colour perception theory is applied whenever colours are reproduced, for example in colour photography, TV broadcasting, printing and computer animation.

Tri-stimuli model

cones and rods

We experience light of different wavelengths as different colours. The retinas in our eyes have *cones* (light-sensitive receptors) that send signals to the brain when they are hit by photons that correspond to different wavelengths in the visible range of the electromagnetic spectrum. There are three different kinds of cones, responding predominantly to blue, green and red light (Figure 5.1). The signals sent to our brain by these cones give us sensations of colour. In addition to cones, we have *rods*, which sense brightness. The rods can operate with less light than the cones and do not contribute to colour vision. For this reason, objects appear less colourful in conditions of low light.

colour monitors

This knowledge of the three stimuli is important for displaying colour. Colour television screens and computer monitors are composed of a large number of small dots arranged in a regular pattern of groups of three: a red, a green and a blue dot. At a normal viewing distance from A TV screen, for example, we cannot distinguish the individual dots. We can individually trigger these dots and vary the amount of light emitted from each of them. All colours visible on such a screen are, therefore, created by mixing different amounts of red, green and blue. This mixing takes place in our brain. When we see a mixture of red (say, 700 nm) and green (530 nm) light, we get the same impression as when we see monochromatic yellow light (i.e. with a distinct wavelength of, say, 570 nm). In both cases, the cones are apparently stimulated in the same way. According to the tri-stimuli model, therefore, three different kinds of dots are necessary and sufficient to recreate the sensation of all the colours of the rainbow.

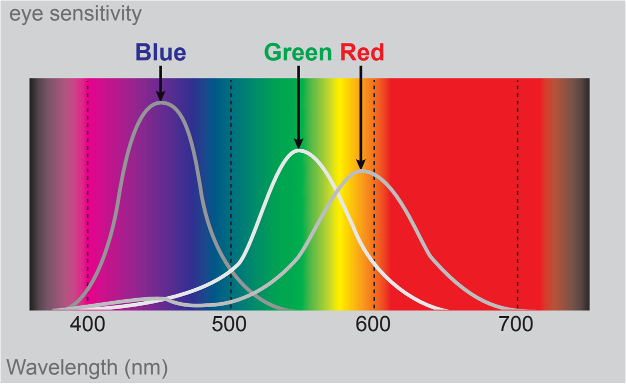


Figure 5.1
Visible range of the electromagnetic spectrum, including the sensitivity curves of cones in the human eye.

Colour spaces

The tri-stimuli model of colour perception is generally accepted. It states that there are three degrees of freedom in the description of a colour. Various three-dimensional spaces are used to describe and define colours. For our purposes, the following three are sufficient:

1. Red-Green-Blue (RGB) space, which is based on the additive mixing principle of colour;
2. Intensity-Hue-Saturation (IHS) space, which most closely resembles our intuitive perception of colour;
3. Yellow-Magenta-Cyan (YMC) space, which is based on the subtractive principle of colour.

RGB

additive colour scheme

The RGB definition of colour is directly related to the way in which computer and television screens function. Three channels directly related to the red, green and blue dots are the input to the screen. When we look at the result, our brain combines the stimuli from the red, green and blue dots and enables us to perceive all possible colours from the visible part of the spectrum. During the combination, the three colours are added. We see yellow when green dots are illuminated in addition to red ones. This principle is called the *additive colour scheme*. Figure 5.2 illustrates the additive colours caused by activating red, green and blue dots on a monitor. When only red and green light is emitted, the result is yellow. In the central area, there are equal amounts of light emitted from all three dots, so we experience *white*.

In the additive colour scheme, all visible colours can be expressed as combinations of red, green and blue, and can therefore be plotted in a three-dimensional space with R, G and B each being one of the axes. The space is bounded by minimum and maximum values for red, green and blue, thus defining what is known as the colour cube. Figure 5.3 shows the normalized colour cube; the maximum value for each colour is set to 1.

IHS

In day-to-day speech, we do not express colours using the RGB model. The IHS model more naturally reflects our perception of colour. *Intensity* in the colour space describes

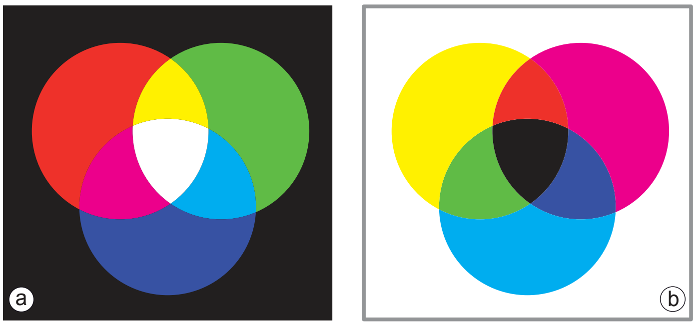


Figure 5.2
Comparison of the (a) additive and (b) subtractive colour schemes.

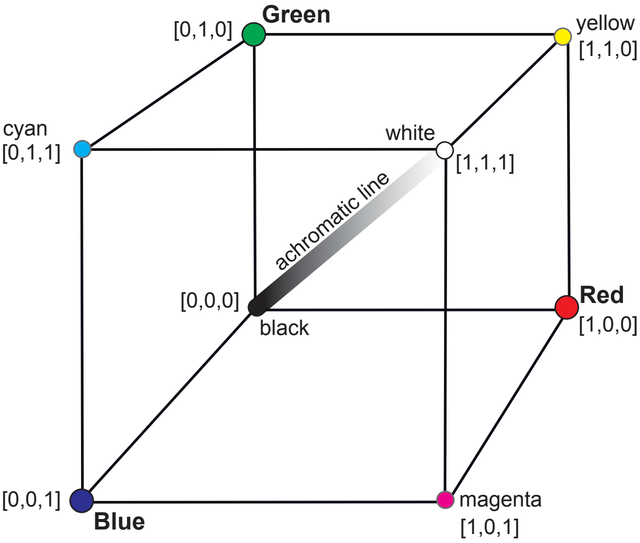


Figure 5.3
The RGB cube. Note the red, green and blue corner points.

intensity

hue

saturation

whether a colour is dark or light and we use for intensity the value range 0 to 1 (projection on the achromatic diagonal). *Hue* refers to the names that we give to colours: red, green, yellow, orange, purple, etc. We quantify hue by degrees in the range 0 to 360 around the achromatic line. *Saturation* describes a colour in terms of purity and we quantify it as the distance from the achromatic line. “Vivid” and “dull” are examples of common words in the English language that are used to describe colour of high and low saturation, respectively. A neutral grey has zero saturation. As is the case for the RGB system, again three values are sufficient to describe any colour.

Figure 5.4 illustrates the correspondence between the RGB and the IHS models. The IHS colour space cannot easily be transformed to the RGB space because they are completely different. The cube in Figure 5.3 must be converted to a double cone; the inset in Figure 5.4 illustrates this. Although the mathematical model for this description is tricky, the description itself is natural. For example, “light, pale red” is easier to imagine than “a lot of red with considerable amounts of green and blue”. The result, however, is the same. Since the IHS model deals with colour perception, which is somewhat subjective, complete agreement of the definitions does not exist. Important for image fusion is the calculation of intensity values and, luckily, this is the simplest of all the calculations. Be aware that the values in the RGB model actually range from 0 to 255, while in the IHS model, intensity ranges from 0 to 1. The formula for intensity is:

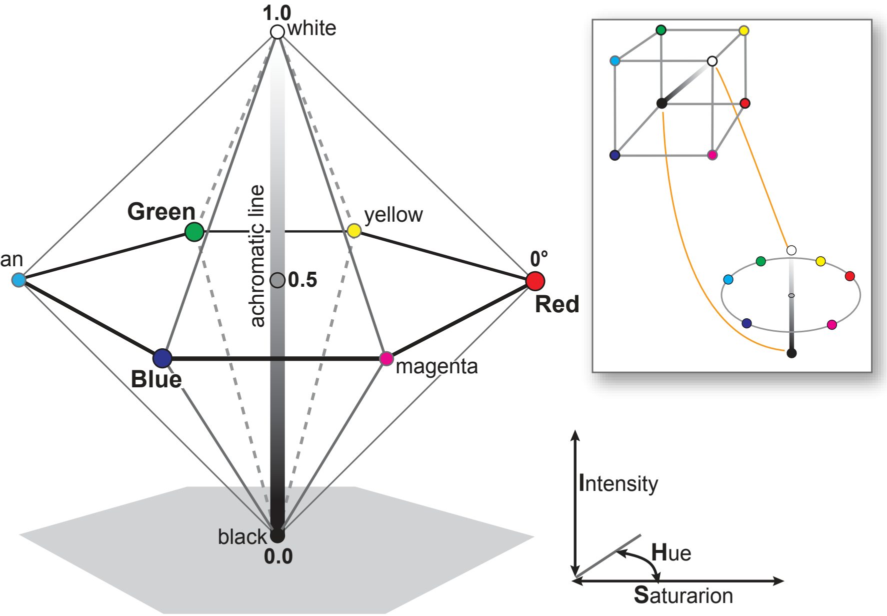


Figure 5.4
Relationship between RGB and IHS colour spaces.

$$I = \frac{R + G + B}{3 \cdot 255}, \tag{5.1}$$

an orthogonal projection on the achromatic line. For example: $(R, G, B) = (150, 200, 100)$. $I = ((150 + 200 + 100)/(3 \cdot 255)) = 0.59$.

YMC

subtractive colour scheme

Whereas RGB is used for computer and TV screens, the YMC colour model is used in colour definition for hardcopy media, such as printed pictures and photographic prints on paper. The principle of YMC colour definition is to consider each component as a coloured filter (Figure 5.2b). The filters are yellow, magenta and cyan. Each filter subtracts one primary colour from the white: the magenta filter subtracts green, so that only red and blue are left; the cyan filter subtracts red, and the yellow filter subtracts blue. Where the magenta filter overlaps the cyan filter, both green and red are subtracted and so we see blue. In the central area, all light is filtered so the result is black. Colour printing, which uses white paper and yellow, magenta and cyan ink, is based on the subtractive colour scheme. When sunlight falls on a colour-printed document, part of it is filtered out by the ink layers and the colour remaining is reflected from the underlying paper.

5.1.2 Image display

grey scale

We normally display a digital image using a grey scale. A "digital image" can be raw data such as that obtained with a panchromatic camera, or data obtained by scanning a B&W photograph, or a single band of a multi-band image. For image display, standard computer monitors support 8 bits per pixel. Thus, if we have sensor recordings of 8 bits, each DN will correspond to exactly one grey value. A pixel having the value zero will be shown as black, a pixel having the value 255 as white. Any DN in between becomes, therefore, some shade of grey. One to one correspondence between DN and grey value used to be the standard, so we still often use "grey value" as a synonym for DN. A colour monitor has three input channels, so we have to feed each of them with

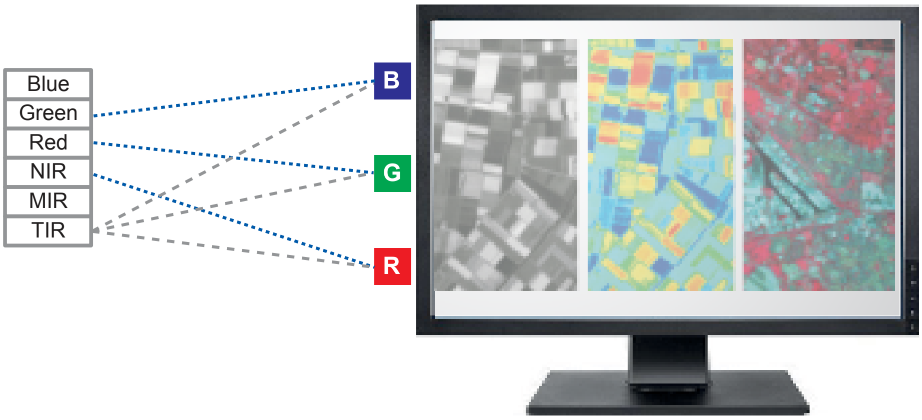


Figure 5.5
Single-band and three-band image display using the red, green and blue input channels of the monitor.

the same DN to obtain a “grey scale image” (Figure 5.5).

pseudo-colour

An alternative way of displaying single-band data is to use a colour scale to obtain a *pseudo-colour* image. We can assign colours (ranging from blue via cyan, green and yellow to red) to different portions of the DN range 0–255 (Figure 5.5). The use of pseudo-colour is especially useful for displaying data that are not reflection measurements. With thermal infrared data, for example, the association of cold *versus* warm with blue *versus* red is more intuitive than with dark *versus* bright.

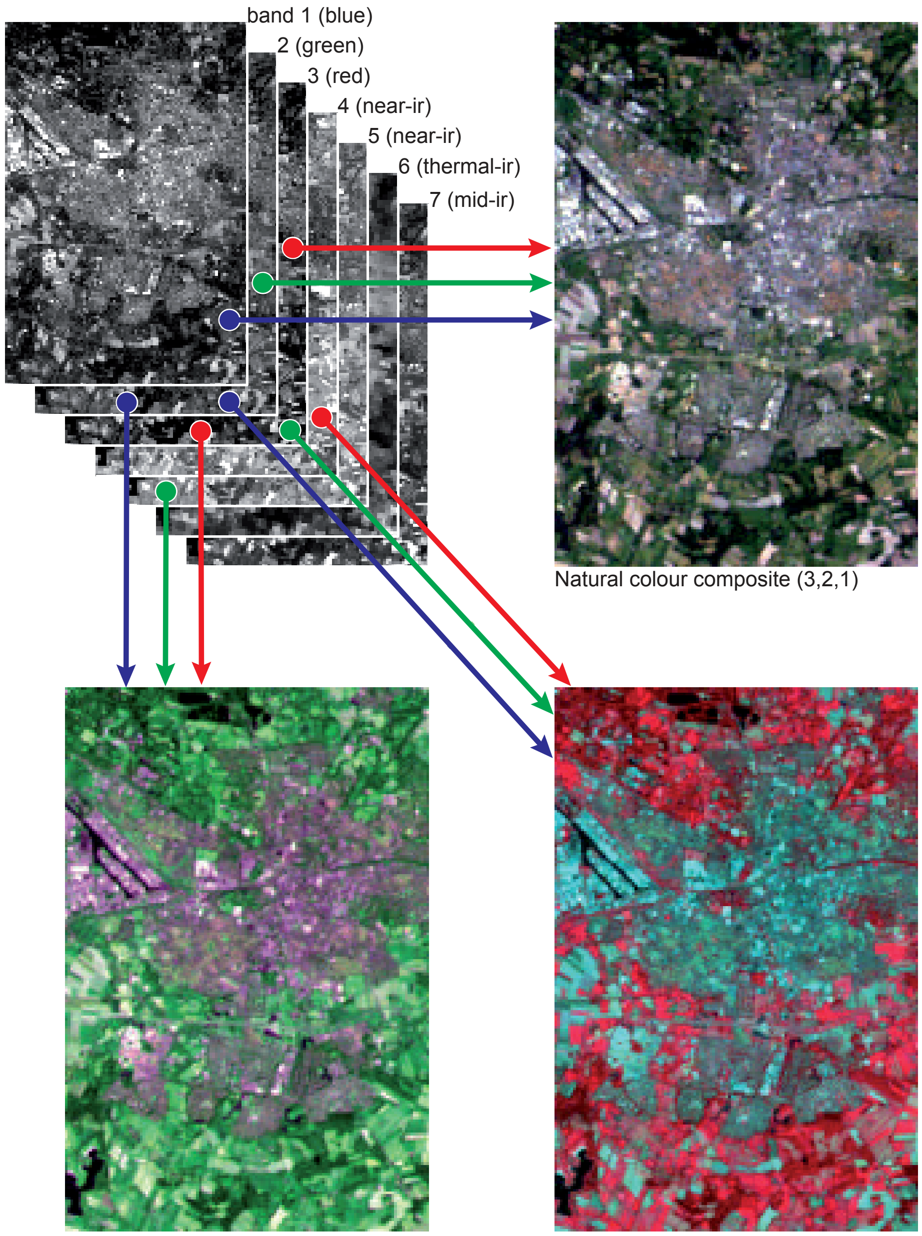
colour composites

When dealing with a multi-band image, any combination of three bands can, in principle, be used as input to the RGB channels of the monitor. The choice should be made based on the intended use of the image. Figure 5.5 indicates how we obtain a *false colour composite*.

true colour

false colour

Sometimes a *true colour composite* is made, where the RGB channels relate to the red, green and blue wavelength bands of a camera or multispectral scanner. An other popular choice is to link RGB to the near-infrared, red and green bands, respectively, to obtain a standard *false colour composite* (Figure 5.6). The most striking characteristic of false colour composites is that vegetation appears as a red-purple colour. In the visible part of the spectrum, plants reflect mostly green light, but their infrared reflection is even higher. Therefore, vegetation displays in a false colour composite as a combination of some blue and a lot of red, resulting in a reddish tint of purple.



Natural colour composite (3,2,1)

Pseudo-natural colour composite (3,5,2)

False colour composite (4,3,2)

Figure 5.6
Landsat-5 TM false colour composite of Enschede and surroundings. Three different colour composites are shown: true colour, pseudo-natural colour and false colour composites.

Depending on the application, band combinations other than true or false colour may be used. Land use categories can often be distinguished quite well by assigning a combination of Landsat-5 TM bands 5–4–3 or 4–5–3 to RGB.

pseudo-natural colour

Combinations that display NIR as green show vegetation in a green colour and are, therefore, often called *pseudo-natural colour composites* (Figure 5.6). Note that there is no common consensus on the naming of certain composites ("true" may also be referred to as "natural"; "false colour" may also be used for other band combinations than green, red and NIR, etc.). Once you have become familiar with the additive mixing of the red, green and blue primaries, you can intuitively relate the colours—which you perceive on the computer monitor—to the digital numbers of the three input bands, thereby gaining a qualitative insight into the spectral properties of an imaged scene.

anaglyph stereograph

stereoscope

To obtain a 3D visual model from a stereo-image pair on a computer screen, we must combine the images into a stereograph (Figure 5.7) and then use a device that helps us to view the left image with the left eye and the right image with the right eye. There are various technical solutions for this problem, one of which is the *anaglyph* method. The left image is displayed in red, the right one in cyan and the two images are superimposed. For viewing, you need spectacles with a red glass for the left eye and a cyan glass for the right eye. High-end digital photogrammetric systems use polarization instead of colour coding. Polarized spectacles make the images visible to the appropriate eye. The advantage of using polarized images is that we can display a full-colour stereo model and superimpose the results of measurements in any colour. Yet another approach is to use a "split screen" display and a stereoscope in front of the monitor. A stereoscope is a device consisting of a pair of binoculars and two mirrors, which allows two images positioned next to each other to be viewed simultaneously, thus achieving stereoscopic vision. Stereoscopes can also be used to view paper prints of stereo photographs.

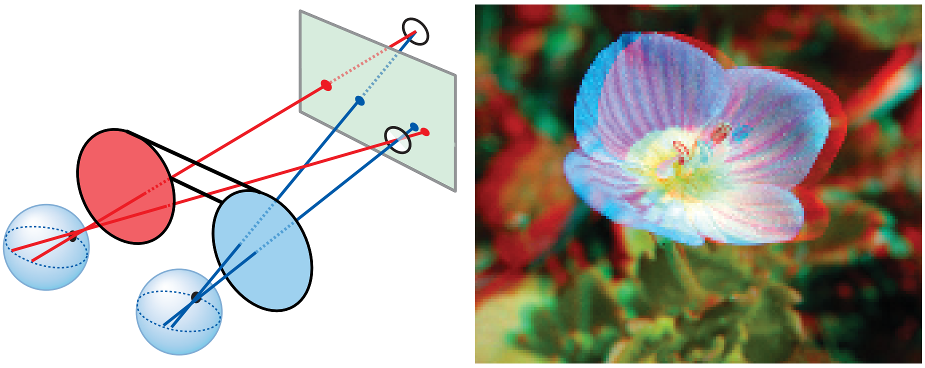


Figure 5.7
The Anaglyph principle and a stereograph.

5.1.3 Radiometric correction

Various techniques can be group under the heading radiometric correction, which aims to correct for various factors that cause degradation of raw RS data. Radiometrically correcting data should make them more suitable for information extraction. Techniques for modifying recorded DNs serve any of the three main purposes outlined below:

image enhancement

- Enhancing images so that they are better suited for visual interpretation. Image enhancement techniques are introduced in a separate section because they can be taken a step further, namely to "low-level image processing" for computer visualization.

- Correcting data for imperfections of the sensor. The detectors of a camera all have a slightly different response. We can determine the differences by radiometric calibration and, accordingly, apply radiometric correction later to the recordings of the camera. Scanners often use several detectors per channel instead of only one. Again, the detectors will each have (slightly) different radiometric responses, with the consequence that the resulting image may be striped. A *destriping* correction will normalize the detectors relatively, if calibration data is absent. A detector may also fail. We may then obtain an image in which, for example, every 10th line is black. A *line drop* correction will cosmetically fix the data. Another detector problem is random noise, which degrades radiometric information content and makes an RS image appear as if salt and pepper was sprinkled over the scene. Correcting all of these disturbances is fairly simple and explained in more detail in Subsection 5.1.8. There can be other degradations caused by the sensor-platform system that are not so easily corrected, such as compensating for image motion blur, which relies on a mathematically complex technique. We got used to referring to theses types of radiometric corrections as *image restoration*. Luckily, image restoration of new sensor data is usually done by the data providers, so you may only have to apply techniques such as destriping and dropped line correction when dealing with old data, e.g. from Landsat MSS. Image restoration should be applied before other corrections and enhancements.

image restoration

- Correcting data for scene peculiarities and atmospheric disturbances. One scene peculiarity is how the scene is illuminated. Consider an area at an appreciable latitude, such as the Netherlands. The illumination of the area will be quite different in winter than in summer (overall brightness, shadows, etc.), because of differences in Sun elevation. Normalizing images taken in different seasons to make them comparable is briefly outlined below. An atmospheric degradation effect, which is already disturbing when extracting information from one RS image, is atmospheric scattering. Sky radiance at the detector causes haze in the image and reduces contrast. Haze correction is briefly described below. - Converting DNs to radiances on the ground (Section 5.2) becomes relevant if we want to compare RS data with ground measurements, or if we want to compare data acquired at different times by different sensors to detect change.

scene normalization

atmospheric correction

5.1.4 Elementary image enhancement

There are two approaches to elementary image processing to enhance an image: histogram operations and filtering. Histogram operations aim at global contrast enhancement, in order to increase the visual distinction between features, while filter operations aim at enhancing local contrast (edge enhancement) and suppressing unwanted image detail. Histogram operations look at DN values without considering where they occur in the image and assign new values from a look-up table based on image statistics. Filtering is a "local operation" in which the new value of a pixel is computed based on the values of the pixels in the local neighbourhood. Figure 5.8 shows the effect of contrast enhancement and edge enhancement for the same input image. Subsection 5.1.5 first explains the notion of histograms.

5.1.5 Histograms

The radiometric properties of a digital image are revealed by its *histogram*, which describes the distribution of the pixel values of the image. By changing the histogram, we change the visual quality of the image. Pixel values (DNs) for 8 bit data range from 0 to 255, so a histogram shows the number of pixels having each value in this

frequency distribution of DNs

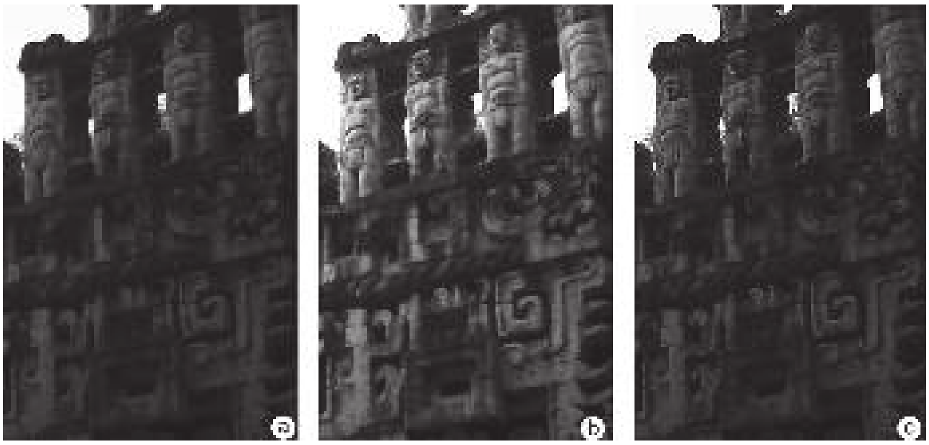


Figure 5.8
An (a) original, (b) contrast enhanced, and (c) edge enhanced image.

range, i.e. the frequency distribution of the DNs. Histogram data can be represented either in tabular form or graphically. The tabular representation (Table 5.1) usually shows five columns. From left to right these are:

- DN: Digital Numbers, in the range 0–255
- Npix: the number of pixels in the image with a particular DN (frequency)
- Perc: frequency as a percentage of the total number of image pixels
- CumNpix: cumulative number of pixels in the image with values less than or equal to a particular DN
- CumPerc: cumulative frequency as a percentage of the total number of image pixels

1% and 99% cut-off

Figure 5.9 shows a plot of the columns 3 and 5 of Table 5.1 against column 1. More commonly, the histogram is displayed as a bar graph rather than as a line graph. The graphical representation of column 5 can readily be used to find the '1% value' and the '99% value'. The 1% value is the DN, below which only 1% of all the values are found. Similarly, there are only 1% of all the DNs of the image larger than the 99% value. The 1% and 99% values are often used in histogram operations as cut-off values for display, thus classifying very small and very large DNs as noise outliers rather than signals.

A histogram can be "summarized" by descriptive statistics: mean, standard deviation, minimum and maximum, as well as the 1% and 99% values (see Table 5.2). The mean is the average of all the DNs of the image; note that it often does not coincide with the DN that appears most frequently (compare Table 5.1 and Table 5.2). The standard deviation indicates the spread of DNs around the mean.

poor contrast

A narrow histogram (thus, a small standard deviation) represents an image of low contrast, because all the DNs are very similar and are initially mapped to only a few grey values. Figure 5.11a shows the histogram of the image shown in Figure 5.8a. Notice the peak at the upper end (DN larger than 247), while most of DNs are smaller than 110. The peak for the white pixels stems from the sky. All other pixels are dark greyish; this narrow part of the histogram characterizes the poor contrast of the image (a Maya monument in Mexico). Remote sensors commonly use detectors with a

Table 5.1
Example of a histogram in a tabular format.

DN	Npix	Perc	CumNpix	CumPerc
0	0	0.00	0	0.00
13	0	0.00	0	0.00
14	1	0.00	1	0.00
15	3	0.00	4	0.01
16	2	0.00	6	0.01
51	55	0.08	627	0.86
52	59	0.08	686	0.94
53	94	0.13	780	1.07
54	138	0.19	918	1.26
102	1392	1.90	25118	34.36
103	1719	2.35	26837	36.71
104	1162	1.59	27999	38.30
105	1332	1.82	29331	40.12
106	1491	2.04	30822	42.16
107	1685	2.31	32507	44.47
108	1399	1.91	33906	46.38
109	1199	1.64	35105	48.02
110	1488	2.04	36593	50.06
111	1460	2.00	38053	52.06
163	720	0.98	71461	97.76
164	597	0.82	72058	98.57
165	416	0.57	72474	99.14
166	274	0.37	72748	99.52
173	3	0.00	73100	100.00
174	0	0.00	73100	100.00
255	0	0.00	73100	100.00

Table 5.2
Statistics summarizing the histogram of the image represented in Table 5.1.

Mean	StdDev	Min	Max	1% value	99% value
113.79	27.84	14	173	53	165

wide dynamic range, so that they can sense under a wide variety of different illumination and emission conditions. These differences, however, are not always present in one particular scene. In practice, therefore, we often obtain RS images that do not exploit the full range of DNs. A simple technique to achieve better visual quality is, then, to enhance the contrast by “grey scale transformation”, which yields a histogram stretched over the entire grey range of pixels on the computer monitor (Figure 5.11c).

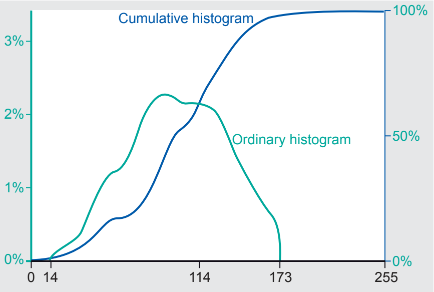


Figure 5.9
Standard histogram and cumulative histogram corresponding to Table 5.1.

5.1.6 Histogram operations

Although contrast enhancement may be done for different purposes, ultimately, in all cases, we will only do so to improve the visual interpretability of an image. In fact, there are two main purposes for applying contrast enhancement. The first is for “temporary enhancement”, where we do not want to change the original data but only want to get a better image on the monitor so that we can carry out a specific interpretation task. Image display for geometric correction is an example of this. The second purpose is to generate new data that have a higher visual quality than the original data. This would be the case if an image is to be the output of our RS activities. Orthophoto production and image mapping are examples of such output (see Section 5.3).

Two techniques of contrast enhancement will now be described: *linear contrast stretch* and *histogram equalization* (occasionally also called histogram stretch). Both are “grey scale transformations”, which convert input DNs (of raw data) to new brightness values (for a more appealing image) by a user-defined “transfer function”.

linear contrast stretch

Linear contrast stretch is a simple grey scale transformation in which the lowest input DN of interest becomes zero and the highest DN of interest becomes 255 (Figure 5.10). The monitor will display zero as black and 255 as white. In practice, we often use the 1% and 99% values as the lowest and highest input DNs. The functional relationship between the input DNs and output pixel values is linear, as shown in Figure 5.11a. The functions shown in the first row (a, d, g) of Figure 5.11 (in the background of the histogram of each input image) are called the *transfer functions*. Many image processing software packages allow users to graphically manipulate the transfer function so that they can obtain an image that appears to their liking. The actual transformation is usually based on a look-up table.

transfer function

The transfer function to be used can be chosen in a number of ways. We can use linear grey-scale transformation to correct for haze and also for other purposes than contrast stretching, for instance to “invert an image” (convert a negative image to a positive one, or vice versa) or to produce a binary image (the simplest technique of image segmentation). Inverting an image is relevant, for example, after having scanned a negative of a photograph; see the last column of Figure 5.11.

Linear stretch is a straight-forward method of contrast enhancement that gives fair results when the input image has a narrow histogram that has a distribution close to

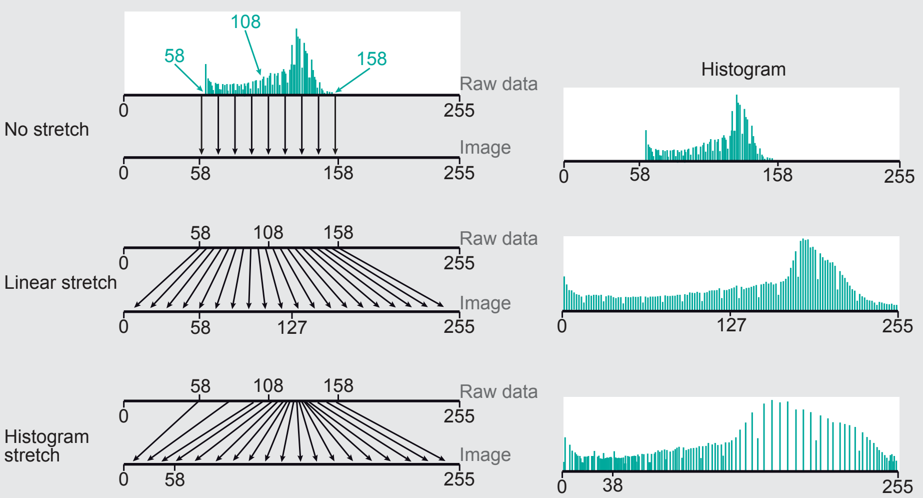


Figure 5.10
Linear contrast stretch versus histogram equalization.

being uniform. The histogram of our example image (Figure 5.8a) is asymmetric, with all DNs in a small range at the lower end, if the irrelevant sky pixels are not considered. Stretching this small range with high frequencies of occurrence (Figure 5.11d) at the expense of compressing the range with only few values (in our example, the range of high brightness) will make more detail (see Figure 5.11e) visible in the dark parts of the original.

histogram equalization

As the name suggests, *histogram equalization* aims at achieving a more uniform distribution in the histogram (see Figure 5.11f). Histogram equalization is a non-linear transformation; several image processing packages offer it as a default function. The idea can be generalized to achieve any desired target histogram.

It is important to note that contrast enhancement (by linear stretch or histogram equalization) merely amplifies small differences between DN values so that we can visually differentiate between features more easily. Contrast enhancement does not, however, increase the information content of the data and it does not consider a pixel neighbourhood. Histogram operations should be based on analysis of the shape and extent of the histogram of the raw data and understanding what is relevant for the intended interpretation. If this is not done, a decrease of information content could easily occur.

5.1.7 Filter operations

A further step in producing optimal images for interpretation is to apply filtering. Filtering is usually carried out for a single band. Filters—algorithms—can be used to enhance images by, for example, reducing noise (“smoothing an image”) or sharpening a blurred image. Filter operations are also used to extract features from images, e.g. edges and lines, and to automatically recognize patterns and detect objects. There are two broad categories of filters: linear and non-linear filters.

linear filter

Linear filters calculate the new value of a pixel as a linear combination of the given values of the pixel and those of neighbouring pixels. A simple example of the use of a linear smoothing filter is when the average of the pixel values in a 3×3 pixel neighbourhood is computed and that average is used as the new value of the central pixel in the neighbourhood (see Figure 5.12). We can conveniently define such a linear filter through a *kernel*. Figure 5.13a shows the kernel of the smoothing filter applied to the example of Figure 5.12. The kernel specifies the size of the neighbourhood that is considered (3×3, or 5×5, or 3×1, etc.) and the coefficients for the linear combination.

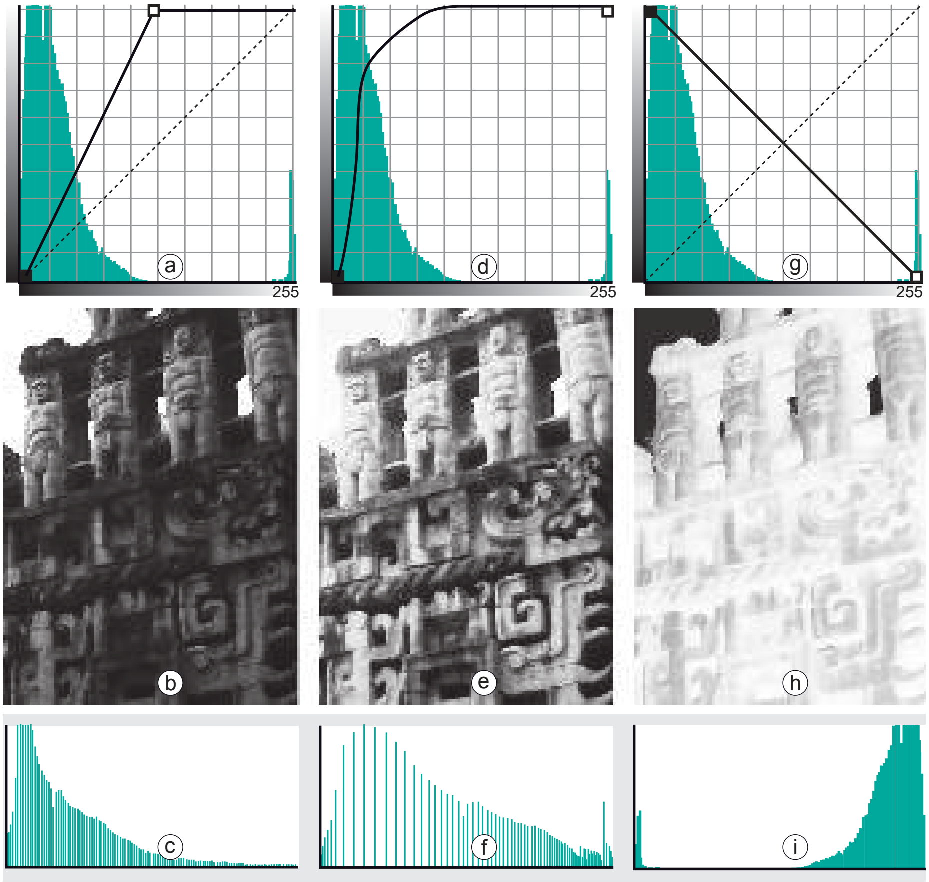


Figure 5.11
Effect of linear grey-scale transformations (b, c & h, i) and histogram equalization (e, f).

Image processing software allows us to select a kernel from a list or define our own kernel. The sum of the coefficients for a smoothing filter should be 1, otherwise an unwanted scaling of DN values will result. The filtering software will usually calculate the *gain*

$$gain = \frac{1}{\Sigma k_i} \tag{5.2}$$

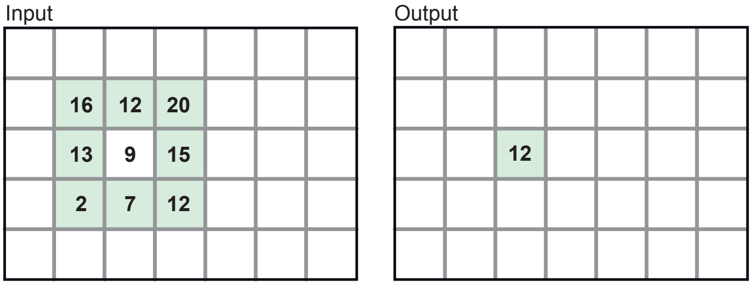
moving average

and multiply the kernel values with it. The following two subsections give examples of kernels and their gain. Since there is only one way of using the kernel of a linear filter, the kernel completely defines the filter. The actual filter operation is to "move the kernel over the input image" line by line, thus calculating for each pixel a local combination of pixel values.

The significance of the gain factor is demonstrated in the following examples. Although in the examples only small neighbourhoods of 3×3 kernels are considered, in practice other kernel dimensions may be used.

Noise reduction

Consider the kernel shown in Figure 5.13a, in which all values equal 1. This means that the values of the nine pixels in the neighbourhood are summed. Subsequently, the result is divided by 9 to ensure that the overall pixel values in the output image



Input

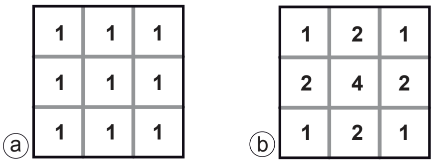
	16	12	20			
	13	9	15			
	2	7	12			

Output

		12				

Figure 5.12
Input and output result of filtering: the neighbourhood in the original image and the filter kernel determine the value of the output. In this case, a smoothing filter was applied.

are in the same range as the input image. In this situation the gain is 1/9 = 0.11. The effect of applying this *averaging filter* is that an image will become blurred or smoothed. This filter could be applied to radar images to reduce the effect of speckle.



(a)

1	1	1
1	1	1
1	1	1

(b)

1	2	1
2	4	2
1	2	1

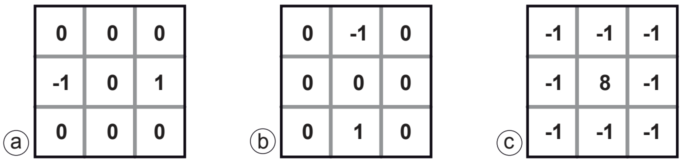
Figure 5.13
Filter kernels for smoothing: (a) equal weights, (b)distance-weighted smoothing.

In the kernel in Figure 5.13a, all pixels contribute equally to the result. It is also possible to define a distance-weighted average instead of an arithmetic mean: the larger the pixel's distance from the centre of the kernel, the smaller the weighting. As a result, less drastic blurring takes place. The resulting kernel, for which the gain is 1/16 = 0.0625, is given in Figure 5.13b.

Edge detection

Filtering can also be used to detect the edges of objects in images. Such edges correspond to local differences in DN values. This is done using a *gradient* filter, which calculates the difference between neighbour pixels in some direction. Filters presented in Figures 5.14a and b, are called x- and $y-$gradient filters; they perform detection of vertical and horizontal edges, respectively. The filter shown in Figures 5.14c detects edges in all directions. Edge detection filtering produces small values in homogeneous areas of an image, while edges are represented by large positive or negative values. Edge detection filtering can be easily recognized by examining kernel elements: their sum must be zero. This applies to all filters shown in Figure 5.14.

gradient filter



(a)

0	0	0
-1	0	1
0	0	0

(b)

0	-1	0
0	0	0
0	1	0

(c)

-1	-1	-1
-1	8	-1
-1	-1	-1

Figure 5.14
Filter kernels for edge detection: (a) x-gradient filter, (b) y-gradient filter, (c) all-directional filter.

Edge enhancement

Filtering can also be used to emphasize local differences in DN values by increasing contrast, for example for linear features such as roads, canals and geological faults. This is done using an *edge enhancing* filter, which calculates the difference between the central pixel and its neighbours. This is implemented using negative values for the non-central kernel elements. An example of an edge enhancement filter is given in Figure 5.15.

-1	-1	-1
-1	16	-1
-1	-1	-1

Figure 5.15
Filter kernel used for edge enhancement.

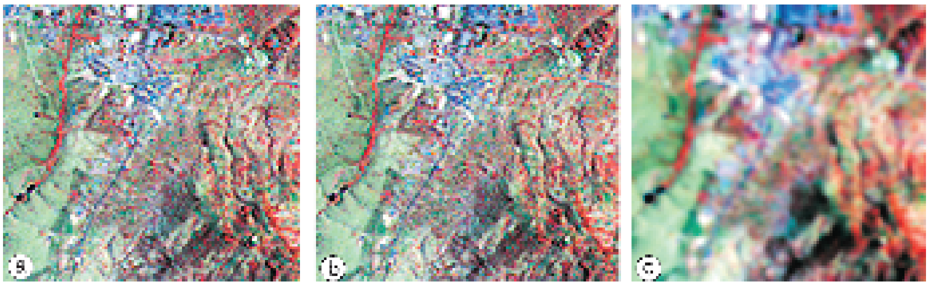


Figure 5.16
Original image (b), edge enhanced image (a) and smoothed image (c).

The gain is calculated as: $1/(16 - 8) = 1/8 = 0.125$. The sharpening effect can be made stronger by using smaller values for the centre pixel (with a minimum of 9). An example of the effect of using smoothing and edge enhancement is shown in Figure 5.16.

5.1.8 Correcting data for imperfections of the sensor

The objective of what is called here "cosmetics" is to correct visible errors and noise in the raw data. No atmospheric model of any kind is involved in these correction processes. Instead, corrections are achieved using especially designed filters and image stretching and enhancement procedures. These corrections are mostly executed (if required) at the station receiving the satellite data or at image pre-processing centres, i.e. before reaching the final user. All applications require this form of correction.

Typical problems requiring "cosmetic" corrections are:

- periodic line dropouts;
- line striping;
- random noise or spike.

These effects can be identified visually and automatically; Figure 5.17 illustrates this with Landsat-7 ETM image of Enschede.

Periodic line dropouts

Periodic line dropouts occur due to recording problems when one of the detectors of the sensor in question either gives wrong data or stops functioning. The Landsat-7 ETM, for example, has 16 detectors for each of its channels, except the thermal channel. A loss of one of the detectors would result in every sixteenth scan line being a string of zeros that would plot as a black line on the image (see Figure 5.18).

The first step in the restoration process is to calculate the average DN value per scan line for the entire scene. The average DN value for each scan line is then compared with this scene average. Any scan line deviating from the average by more than a designated threshold value is identified as defective. In regions of very diverse land cover, better results can be achieved by using the histogram for sub-scenes and processing these sub-scenes separately.

The next step is to replace the defective lines. For each pixel in a defective line, an average DN is calculated from the DNs for the corresponding pixel in the preceding and succeeding scan lines by using the principle of spatial autocorrelation. The average DN is then substituted for the defective pixel. The resulting image is a major improvement, although every sixteenth scan line (or every sixth scan line, in the case of Landsat MSS data) consists of artificial data (see Figure 5.19). This restoration program is equally effective for random line dropouts that do not follow a systematic pattern.

Line striping

Line striping is far more common than line dropouts. Line striping often occurs as a result of non-identical detector response. Although the detectors for all satellite sensors are carefully calibrated and matched before the launch of the satellite, with time the response of some detectors may drift to higher or lower levels. As a result, every scan line recorded by that detector is brighter or darker than the other lines (see Figure 5.20). It is important to understand that valid data are present in the defective lines, but these must be corrected to match the overall scene.

Though several procedures can be adopted to correct this effect, the most popular one is histogram matching. Separate histograms corresponding to each detector unit are constructed and matched. Taking one response as standard, the gain (rate of increase of DNs) and offset (relative shift of mean) for all other detector units are suitably adjusted, and new DNs are computed and assigned. This yields a destriped image in which all DN values conform to the reference level and scale.

Random noise or spike noise

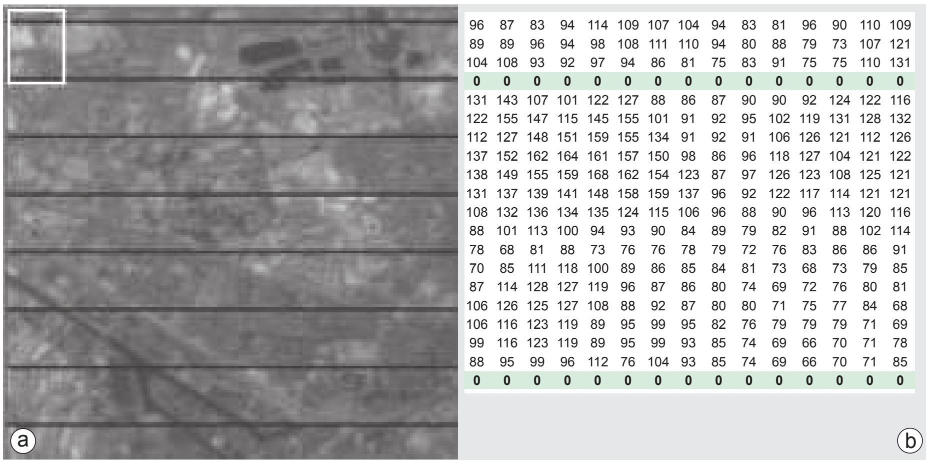
Periodic line dropouts and striping are forms of non-random noise that may be recognized and restored by simple means. Random noise, on the other hand, requires a more sophisticated restoration method, such as digital filtering.

Random noise or spike noise may be caused by errors during transmission of data or a temporary disturbance. Here, individual pixels acquire DN values that are much higher or lower than the surrounding pixels (Figure 5.21). In the image, these pixels produce bright and dark spots that interfere with information extraction procedures.

A spike noise can be detected by mutually comparing neighbouring pixel values. If neighbouring pixel values differ by more than a specific threshold margin,it is desig-

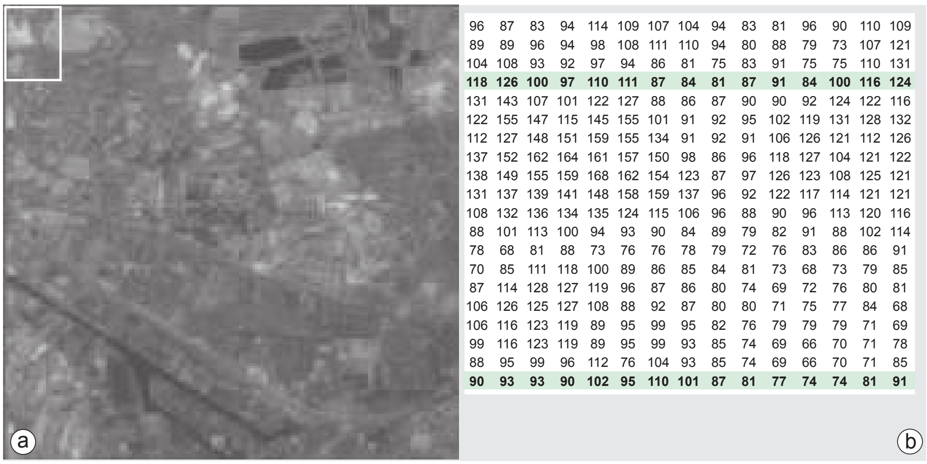
96	87	83	94	114	109	107	104	94	83	81	96	90	110	109
89	89	96	94	98	108	111	110	94	80	88	79	73	107	121
104	108	93	92	97	94	86	81	75	83	91	75	75	110	131
114	97	93	86	96	100	92	88	88	92	93	92	105	110	106
131	143	107	101	122	127	88	86	87	90	90	92	124	122	116
122	155	147	115	145	155	101	91	92	95	102	119	131	128	132
112	127	148	151	159	155	134	91	92	91	106	126	121	112	126
137	152	162	164	161	157	150	98	86	96	118	127	104	121	122
138	149	155	159	168	162	154	123	87	97	126	123	108	125	121
131	137	139	141	148	158	159	137	96	92	122	117	114	121	121
108	132	136	134	135	124	115	106	96	88	90	96	113	120	116
88	101	113	100	94	93	90	84	89	79	82	91	88	102	114
78	68	81	88	73	76	76	78	79	72	76	83	86	86	91
70	85	111	118	100	89	86	85	84	81	73	68	73	79	85
87	114	128	127	119	96	87	86	80	74	69	72	76	80	81
106	126	125	127	108	88	92	87	80	80	71	75	77	84	68
106	116	123	119	89	95	99	95	82	76	79	79	79	71	69
99	116	123	119	89	95	99	93	85	74	69	66	70	71	78
88	95	99	96	112	76	104	93	85	74	69	66	70	71	85
91	91	85	87	80	97	115	105	94	87	78	70	73	65	85

Figure 5.17
Original Landsat ETM image of Enschede and and its surroundings (a), and corresponding DNs of the the indicated subset (b).



96	87	83	94	114	109	107	104	94	83	81	96	90	110	109
89	89	96	94	98	108	111	110	94	80	88	79	73	107	121
104	108	93	92	97	94	86	81	75	83	91	75	75	110	131
0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
131	143	107	101	122	127	88	86	87	90	90	92	124	122	116
122	155	147	115	145	155	101	91	92	95	102	119	131	128	132
112	127	148	151	159	155	134	91	92	91	106	126	121	112	126
137	152	162	164	161	157	150	98	86	96	118	127	104	121	122
138	149	155	159	168	162	154	123	87	97	126	123	108	125	121
131	137	139	141	148	158	159	137	96	92	122	117	114	121	121
108	132	136	134	135	124	115	106	96	88	90	96	113	120	116
88	101	113	100	94	93	90	84	89	79	82	91	88	102	114
78	68	81	88	73	76	76	78	79	72	76	83	86	86	91
70	85	111	118	100	89	86	85	84	81	73	68	73	79	85
87	114	128	127	119	96	87	86	80	74	69	72	76	80	81
106	126	125	127	108	88	92	87	80	80	71	75	77	84	68
106	116	123	119	89	95	99	95	82	76	79	79	79	71	69
99	116	123	119	89	95	99	93	85	74	69	66	70	71	78
88	95	99	96	112	76	104	93	85	74	69	66	70	71	85
0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

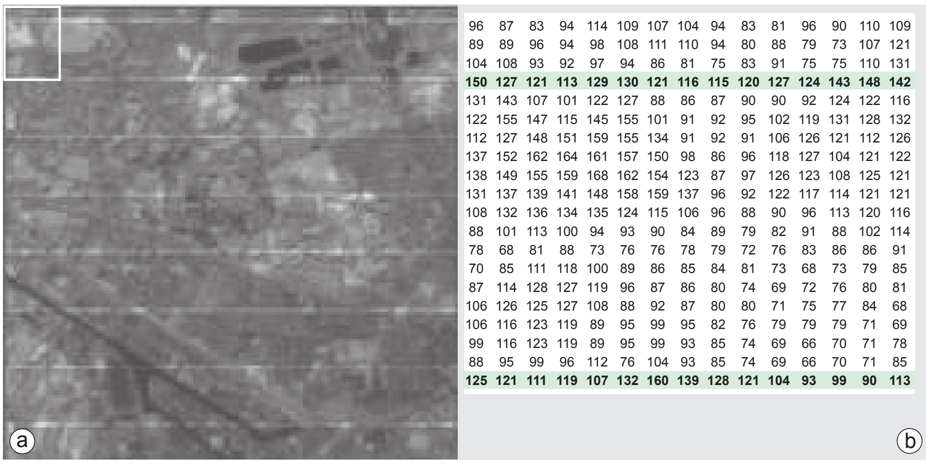
Figure 5.18
The image with periodic line dropouts (a) and the DNs (b). All erroneous DNs in these examples are shown in bold.



96	87	83	94	114	109	107	104	94	83	81	96	90	110	109
89	89	96	94	98	108	111	110	94	80	88	79	73	107	121
104	108	93	92	97	94	86	81	75	83	91	75	75	110	131
118	**126**	**100**	**97**	**110**	**111**	**87**	**84**	**81**	**87**	**91**	**84**	**100**	**116**	**124**
131	143	107	101	122	127	88	86	87	90	90	92	124	122	116
122	155	147	115	145	155	101	91	92	95	102	119	131	128	132
112	127	148	151	159	155	134	91	92	91	106	126	121	112	126
137	152	162	164	161	157	150	98	86	96	118	127	104	121	122
138	149	155	159	168	162	154	123	87	97	126	123	108	125	121
131	137	139	141	148	158	159	137	96	92	122	117	114	121	121
108	132	136	134	135	124	115	106	96	88	90	96	113	120	116
88	101	113	100	94	93	90	84	89	79	82	91	88	102	114
78	68	81	88	73	76	76	78	79	72	76	83	86	86	91
70	85	111	118	100	89	86	85	84	81	73	68	73	79	85
87	114	128	127	119	96	87	86	80	74	69	72	76	80	81
106	126	125	127	108	88	92	87	80	80	71	75	77	84	68
106	116	123	119	89	95	99	95	82	76	79	79	79	71	69
99	116	123	119	89	95	99	93	85	74	69	66	70	71	78
88	95	99	96	112	76	104	93	85	74	69	66	70	71	85
90	**93**	**93**	**90**	**102**	**95**	**110**	**101**	**87**	**81**	**77**	**74**	**74**	**81**	**91**

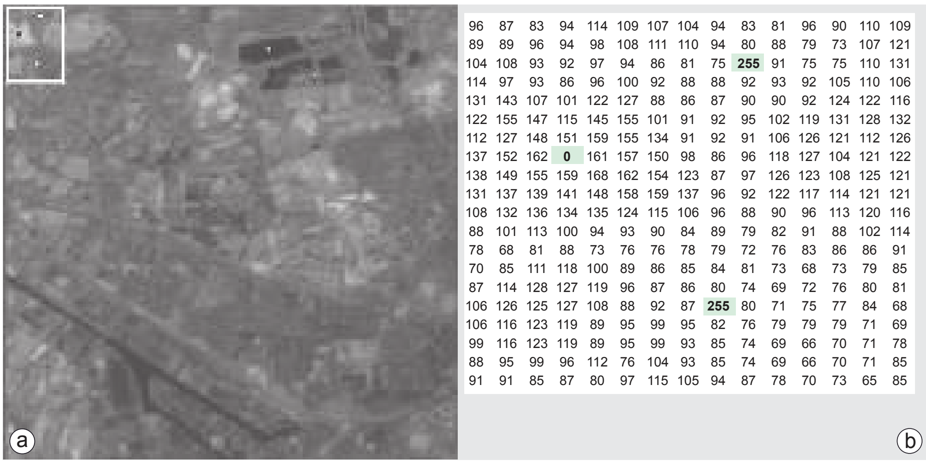
Figure 5.19
The image after correction for line dropouts (a) and the DNs (b).

nated as spike noise and the DN is replaced by an interpolated DN.



96	87	83	94	114	109	107	104	94	83	81	96	90	110	109
89	89	96	94	98	108	111	110	94	80	88	79	73	107	121
104	108	93	92	97	94	86	81	75	83	91	75	75	110	131
150	**127**	**121**	**113**	**129**	**130**	**121**	**116**	**115**	**120**	**127**	**124**	**143**	**148**	**142**
131	143	107	101	122	127	88	86	87	90	90	92	124	122	116
122	155	147	115	145	155	101	91	92	95	102	119	131	128	132
112	127	148	151	159	155	134	91	92	91	106	126	121	112	126
137	152	162	164	161	157	150	98	86	96	118	127	104	121	122
138	149	155	159	168	162	154	123	87	97	126	123	108	125	121
131	137	139	141	148	158	159	137	96	92	122	117	114	121	121
108	132	136	134	135	124	115	106	96	88	90	96	113	120	116
88	101	113	100	94	93	90	84	89	79	82	91	88	102	114
78	68	81	88	73	76	76	78	79	72	76	83	86	86	91
70	85	111	118	100	89	86	85	84	81	73	68	73	79	85
87	114	128	127	119	96	87	86	80	74	69	72	76	80	81
106	126	125	127	108	88	92	87	80	80	71	75	77	84	68
106	116	123	119	89	95	99	95	82	76	79	79	79	71	69
99	116	123	119	89	95	99	93	85	74	69	66	70	71	78
88	95	99	96	112	76	104	93	85	74	69	66	70	71	85
125	**121**	**111**	**119**	**107**	**132**	**160**	**139**	**128**	**121**	**104**	**93**	**99**	**90**	**113**

Figure 5.20
The image with line striping (a) and the DNs (b). Note that the destriped image would look similar to the original image.



96	87	83	94	114	109	107	104	94	83	81	96	90	110	109
89	89	96	94	98	108	111	110	94	80	88	79	73	107	121
104	108	93	92	97	94	86	81	75	**255**	91	75	75	110	131
114	97	93	86	96	100	92	88	88	92	93	92	105	110	106
131	143	107	101	122	127	88	86	87	90	90	92	124	122	116
122	155	147	115	145	155	101	91	92	95	102	119	131	128	132
112	127	148	151	159	155	134	91	92	91	106	126	121	112	126
137	152	162	**0**	161	157	150	98	86	96	118	127	104	121	122
138	149	155	159	168	162	154	123	87	97	126	123	108	125	121
131	137	139	141	148	158	159	137	96	92	122	117	114	121	121
108	132	136	134	135	124	115	106	96	88	90	96	113	120	116
88	101	113	100	94	93	90	84	89	79	82	91	88	102	114
78	68	81	88	73	76	76	78	79	72	76	83	86	86	91
70	85	111	118	100	89	86	85	84	81	73	68	73	79	85
87	114	128	127	119	96	87	86	80	74	69	72	76	80	81
106	126	125	127	108	88	92	87	**255**	80	71	75	77	84	68
106	116	123	119	89	95	99	95	82	76	79	79	79	71	69
99	116	123	119	89	95	99	93	85	74	69	66	70	71	78
88	95	99	96	112	76	104	93	85	74	69	66	70	71	85
91	91	85	87	80	97	115	105	94	87	78	70	73	65	85

Figure 5.21
Image with spike errors (a) and the DNs (b).

5.2 Correction of atmospheric disturbance

Introduction

The radiance values of reflected polychromatic solar radiation and/or the emitted thermal and microwave radiances from a certain target area on the Earth's surface are for researchers the most valuable information obtainable from a remote sensor. In the absence of an atmosphere, the radiance for any wavelength on the ground would be the same as the radiance at the sensor. No atmosphere would make RS easier—but life impossible. So we have to figure out how we can convert remotely detected radiances to radiances at ground level.

In this section we will consider relative and absolute atmospheric correction. Relative atmospheric correction is based on ground reflectance properties, while absolute atmospheric correction is based on atmospheric process information. Before we explain how to correct, Subsection 5.2.1 will review the imaging process and the occurring disturbances.

5.2.1 From satellite to ground radiances: atmospheric correction

The presence of a heterogeneous, dense and layered terrestrial atmosphere composed of water vapour, aerosols and gases disturbs the signal reaching sensors in many ways. Therefore, methods of atmospheric corrections (AC) are needed to "clean" the images from these disturbances, in order to allow the retrieval of pure ground radiances from the target. The physics behind AC techniques in visible and thermal ranges is essentially the same, meaning that the same AC procedures that are applicable to one also apply to the other. However, there are a number of reasons for making a distinction between techniques applied to visible data and thermal data:

- Incident and reflected solar radiation and terrestrial thermal emissions belong to very different parts of the spectrum.
- Solar emission and reflection depends on the position of the Sun and the satellite at the time of image acquisition. Thermal emission is theoretically less dependent on this geometry.
- Solar rays travel twice through the atmosphere before they reach the sensor (Top of the Atmosphere (TOA)–ground–sensor), whereas ground thermal emissions only pass through the atmosphere once (ground–sensor; see Figure 2.6).
- Solar reflection at the Earth's surface depends on material reflectance (ρ). Thermal emission from the Earth depends on the emissivity of the surface materials (ϵ). Since solar reflection and Earth thermal emission occur at different wavelengths, the behaviour of one is not an indication of the other.
- The processes of atmospheric attenuation, i.e. scattering and absorption, are both wavelength dependent and affect the two sectors of the spectrum differently.
- As a result of the previous point, AC techniques are applied at a monochromatic level (individual wavelengths). This means that attenuation of radiation is calculated at every individual wavelength and then integrated across the spectrum of the sensor by mathematical integration.
- Atmospheric components affect different areas of the spectrum in different ways, meaning that some components can be neglected when dealing with data belonging to the thermal or the visible part of the spectrum.

A classification of different AC methods allows us to assess what kind of effort is needed to correct raw data for the particular application at hand. Some RS applications do not require AC procedures at all, except for some "cosmetics", while others call for rigourous and complex procedures. "Intermediate" solutions are sufficient for many applications.

In general, applications for which the actual radiance at ground level is not needed do not require atmospheric correction. Some "cosmetic" and/or image enhancement procedures may suffice: for instance, mapping applications where visual interpretation and image geometry are important, but not the chemical properties of surface material.

Applications that require the quantification of radiation at ground level must include rigourous atmospheric correction procedures. Quantification of evapotranspiration or CO_2 sequestration, or surface temperature and reflectivity mapping, are examples of such applications.

Applications concerned with the evolution of certain parameters or land properties over time, rather than their absolute quantification, are "intermediate" cases. For these, knowledge of the relative trend may suffice. Such procedures apply mainly when the mapping parameters do not really have a meaningful physical value, simply because they were designed primarily for multi-temporal relative comparison. Index evolution and correlation procedures, where radiances are associated with the evolution of a certain parameter (e.g. turbidity) are examples of this category. Be aware that some indexes such as NDVI typically require some absolute atmospheric correction.

The "effort" required is commensurate with the amount of information required to describe the components of the atmosphere at different altitudes (atmospheric profiling) at the moment and position at which the image is taken, and less so with sophistication of the AC procedure itself. State-of-the-art atmospheric models allow the "cleaning" of any cloudless image regardless of sensor type, as long as atmospheric profile data are available. Unfortunately, such detailed atmospheric information can only be obtained through atmospheric sounding procedures, which use a series of instruments to sample the atmosphere at fixed intervals while being transported vertically by a balloon, or sounding sensors on board a satellite. This kind of profiling is carried out daily (at fixed times) at some atmospheric centres, regardless of satellite overpass times. However, the atmosphere is dynamic. Atmospheric processes and composition change rapidly, mainly at low altitudes (water vapour and aerosols), meaning that soundings made somewhere close to the target and near the time of a satellite overpass might not be enough to ensure an adequate atmospheric description. As as rule of thumb regarding AC techniques, first consider the objectives of the project, then identify the appropriate AC procedure, and finally establish the effort, i.e. the required information to execute the chosen correction procedure.

5.2.2 Atmospheric corrections

Haze correction

subraction per band

Equation 2.2 shows that atmospheric scattering adds a "sky radiance". Haze correction aims at removing sky radiance effects from raw data, and doing so can be beneficial to many applications of space-borne RS. In Section 2.4 you have also learned that scattering depends on wavelength: Rayleigh scattering will hardly affect recordings in the red spectral band, while DNs in the blue band may become significantly larger. Reducing haze, therefore, must be done independently for each band of an RS image. How much to subtract from every DN from a particular band? We can find out if the scene is favourable and contains areas that should have zero reflectance (a spectral-band-specific black body). Deep clear water, for example, should yield pixel values of zero in the NIR band. If not, we attribute the minimum value found for "water pixels" to sky radiance and subtract this value from all DNs in this band. The alternative is less reliable, i.e. to look at the histogram (see Subsection 5.1.5) of the band and simply take the smallest DN found there as the haze correction constant.

Sun elevation correction

normalization by sine

Illumination differences will cause problems if we want to analyse sequences of images of a particular area that were taken on different dates (or images of the same date taken at different time), or if we would like to make mosaics of such images. We can apply a simple Sun elevation correction if the images stem from the same sensor. The trick is to normalize the images as if they were taken with the Sun at its zenith. We can achieve this normalization by dividing every pixel value of an image by the sine of the Sun elevation angle at the time of data acquisition. The Sun elevation angle is usually given in the meta-data file, which is supplied with an image. Obviously this is an approximate correction as it does not take into account the effect of elevation and height differences in the scene, nor atmospheric effects.

Relative AC methods based on ground reflectance

Relative AC methods avoid the evaluation of atmospheric components of any kind. They rely on the fact that, for one sensor channel, the relation between the radiances at TOA and at ground level follows a linear trend for the variety of Earth features present in the image. This linear relation is in fact an approximation of reality, but for practical purposes it is precise enough when there are other, more important sources of error. The AC methods are:

Two reflectance measurements: The output of this method is an absolute atmospherically corrected image, so it can be used on an individual basis for multi-temporal comparison or parameter evolution and also for flux quantification. "Absolute" means that the image output has physical units and that the calculated ground radiances are compatible with the actual atmospheric constituents. The application of this method requires the use of a portable radiometer able to measure in the same wavelength range as the image band to be corrected. If many bands are to be corrected, then the radiometer should have filters that allow measurement in all these individual bands separately.

Two reference surfaces: The output of this method is an image that matches a reflectance that is compatible with the atmosphere of a similar image taken on a previous date. No absolute values of radiances are obtained in any of the two images, only allowing comparative results. This method works on an individual band/channel basis and is valid for establishing a basis for a uniform comparison to study, for example, the evolution of non-flux related parameters such as indexes, or when certain convenient land properties can be derived directly or indirectly from the normalized radiance values in a band. The method relies on the existence of at least one dark and one bright invariant area. Normally, a sizable area should avoid mixed pixels (mixed land cover). As rule of thumb it should be a minimum of 2 or 3 times larger than the image spatial resolution. Reflective invariant areas are considered to retain their reflective properties over time. Deep reservoir lakes, sandy beaches or deserts, open quarries,large asphalted areas, and large salt deposits are examples of areas that are reflectively invariant. The supposition is that, for these pixels, the reflectance should always be the same since the reflective properties of the materials of which they are composed do not vary with time. If a difference in reflectance occurs for the reflective invariant area in the two date images, it can only be attributed to the different state of the atmosphere on those dates. The atmospheric composition is unknown in the two images, but its influence is measurable by analysing the change in radiance for the reflective invariant areas for the two dates.

Absolute AC methods based on atmospheric processes

These methods require a description of the components in the atmospheric profile. The output of these methods is an image that matches the reflectance of the ground pixels with a maximum estimated error of 10%, provided that atmospheric profiling is adequate enough. This image can be used for flux quantifications, parameter evolution assessments, etc., as mentioned above. The advantage of these methods is that ground reflectance can be evaluated for any atmospheric condition, altitude and relative geometry between the Sun and satellite. The disadvantage is that the atmospheric profiling required for these methods is rarely available. To address this inconvenience, various absolute AC methods have been developed that have different requirements in relation to the atmospheric profiling data—and differences in the accuracy of the results.

Radiative transfer models Radiative transfer models (RTMs) can be used for computing radiances for a wide variety of atmospheric and surface conditions. They require full descriptions of the atmospheric components at fixed altitudes throughout the atmosphere. RTMs are relatively easy to use when the complexity of the atmospheric input is simplified by using one standard atmosphere as input.

Because of the rapid dynamics of the atmosphere in terms of the temporal and spatial variation of its constituents, researchers have found the need to define some often-observed “common profiles” that correspond to average atmospheric conditions for different parts of the Earth. Compilation of these “fixed atmospheres” has been based on actual radio soundings carried out at different research sites, resulting in what are called “standard atmospheres”, e.g. mid-latitude summer, mid-latitude winter, tropical, desert, arctic, US standard, and so on. Researchers use these well-defined standards to characterize typical on-site atmospherics. RTMs have these standards built into the system, allowing the influence of different constituents to be compared under strict simulations. For instance, the influence of water vapour in the thermal, or of aerosols and air molecules in the visible, part of the spectrum can be accurately predicted for different atmospheres, allowing sensitivity analyses for evaluating the importance of these constituents in attenuation processes at different wavelengths.

5.3 Geometric operations

Introduction

If you did not know so before, after reading Chapter 3 you will know that the Earth has a spherical shape and that several clever scientists have devised transformations to map the curved Earth's surface to a plane. Through a map projection (transformation) we can obtain an image of the Earth's surface that has convenient geometric properties. We can, for instance, measure angles on a map and use these for navigation in the real world, or for setting out a designed physical infrastructure. Or if, instead of a conformal projection such as UTM, we use an equivalent projection, we can determine the size of a parcel of land from the map—irrespective of where the parcel is on the map and at which elevation it is on the Earth. A remote sensor "images" the Earth's surface without knowledge of map projections, so we must not expect that remote sensing images have the same geometric properties as a map. Moreover, wherever a remote sensor detects, it merely records DNs. DNs do not come with a label that tell us where exactly on the Earth is the corresponding ground-resolution cell to be found (Figure 5.22).

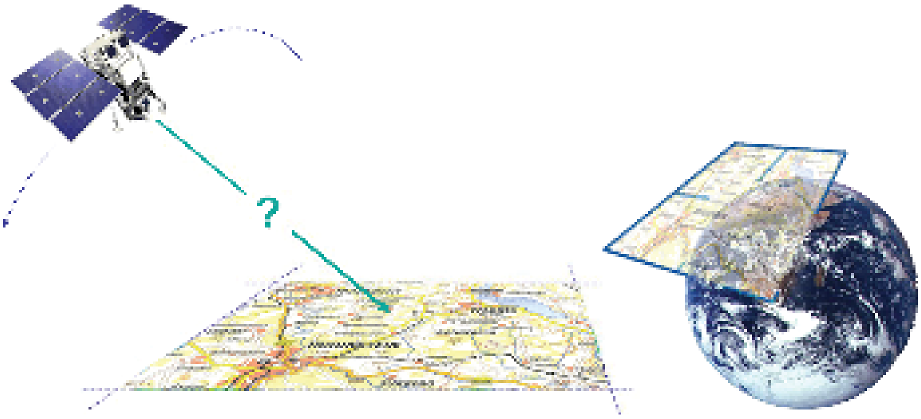


Figure 5.22
The problem of georeferencing an RS image.

Luckily, we know that the DNs are delivered in an orderly fashion, neatly arranged in rows and columns. The position of a point in the image is uniquely defined by the row and column numbers of the pixel that represents the point. Relating a pixel position in the image to the corresponding position on the ground is the purpose of *georeferencing the image*. Once we have figured out what the geometric relationship is between points on the ground and the corresponding point in the RS image, we can transform any pixel position to a position on the ground. "Position on the ground" can be defined either in a 3D terrain coordinate system or through a map projection in a 2D map coordinate system. By georeferencing an image we solve two problems at the same time: (1) we can get map coordinates of features that we identify in the image, and (2) we implicitly correct for geometric distortions of the RS image if we compute correct map coordinates for any pixel. It takes georeferencing to turn RS data into geospatial data. After georeferencing (or, speaking more generally, sensor orientation), we can:

- make measurements within images to obtain 2D and 3D object descriptions. Mapping the world in which we live has been man's concern for thousands of years, with navigation and military activities being the main triggers for topographic mapping. RS—photogrammetry, more specifically—has made that mapping much more efficient. Mapping is still the prime application of RS, although environmental monitoring is catching up quickly because of the exponential damage we are causing to our environment. We are interested in mapping our environment at a variety of spatial and thematic resolutions and accuracies. For many applications, 2D representations (by points, lines and areas) of objects suffice. As long as certain conditions are met, we can obtain these representations from a single image and simple georeferencing, which directly relates the RS image to the digital map. We need stereo images or multiple images for applications requiring 3D coordinates, or for better 2D coordinates, such as for mapping scenes of large elevation differences or objects with large height differences. Sensor orientation must then use more rigourous approaches than for 2D georeferencing.

- combine an image with other images or vector (digital map) data (see also Chapter 11). Assume you would like to see how land property units relate to land cover. If you had a digital cadastral map, you could readily overlay the parcel boundaries on a RS image, e.g. a scanned photograph, which shows land cover nicely. Combining different RS images and/or map data can be done conveniently if all the data sets do not differ in their geometry, if they are all transformed to the same geometric reference system. From a computational perspective, producing a new image from an RS image such that it fits a specific map projection is often referred to as *geocoding*.

5.3.1 Elementary image distortions

Each sensor–platform combination is likely to have its own type of geometric image distortion. Here we only examine three very common types: (1) the effect of oblique viewing, (2) the effect of Earth rotation, and (3) "relief displacement". Tilting a camera (see Figure 4.27) leads to images of non-uniform scale (Figure 5.23). Objects in the foreground appear larger than those farther away from nadir. Earth rotation affects space-borne scanners and line camera images that have a large spatial coverage. The resulting geometric image distortion (Figure 5.23) can easily be corrected by 2D georeferencing. Relief displacement shows up specifically in large-scale camera images if there is significant terrain relief or if there are high, protruding objects.

oblique view
Earth rotation

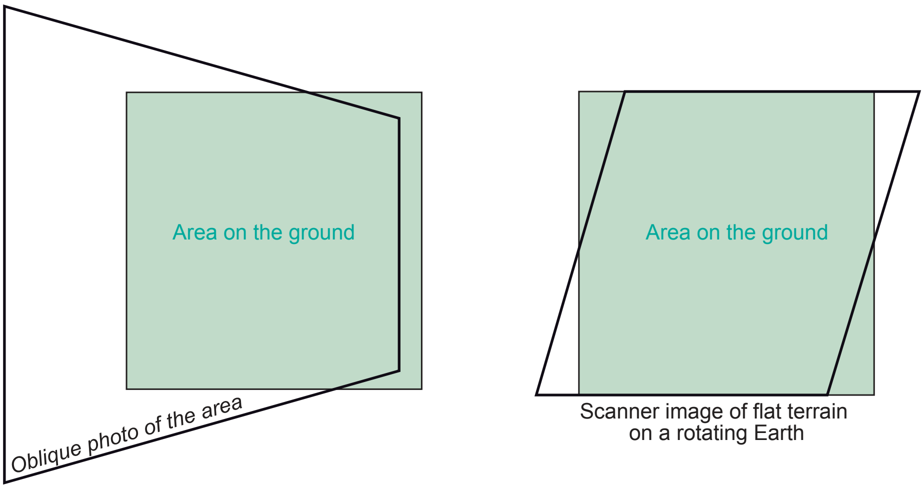


Figure 5.23
Examples of geometric image distortion.

Relief displacement

A characteristic of most sensor systems is the occurrence of distortion of the geometric relationship between the image and a conventional map of the terrain due to elevation differences. This effect is most apparent in aerial photographs but also occurs in images from space-borne line cameras. The effect of relief displacement is illustrated in Figure 5.24 for a line camera. Consider the situation on the left (a), in which a true vertical image is taken of flat terrain. The distances $A - B$ and $a - b$ are proportional to the total width of the scene and its image size, respectively. In the situation on the left, by using the scale factor we can compute $A - B$ from a measurement of $a - b$ in the image. In the situation on the right (b), there is a significant difference in terrain elevation. As you can now observe, the distance between a and b in the image has become larger, although when measured in the terrain system, it is still the same as in the situation on the left. This phenomenon does not occur in the centre of a central projection image but becomes increasingly prominent towards the edges of a camera image. This effect is called *relief displacement*: terrain points whose elevation is above or below the reference elevation are displaced, respectively, away from or towards the nadir point, A, the point on the ground directly beneath the sensor. The magnitude of displacement, δr (in mm), is approximated by:

shifts due to height and elevation

$$\delta r = \frac{rh}{H}. \tag{5.3}$$

In this equation, r is the radial distance (in mm) from nadir, h (in m) is the terrain elevation above the reference plane, and H (in m) is the flying height above the reference plane (where nadir intersects the terrain). The equation shows that the amount of relief displacement is zero at nadir (r =0) and largest at the edges of a line camera image and the corners of a frame camera image. Relief displacement is inversely proportional to the flying height.

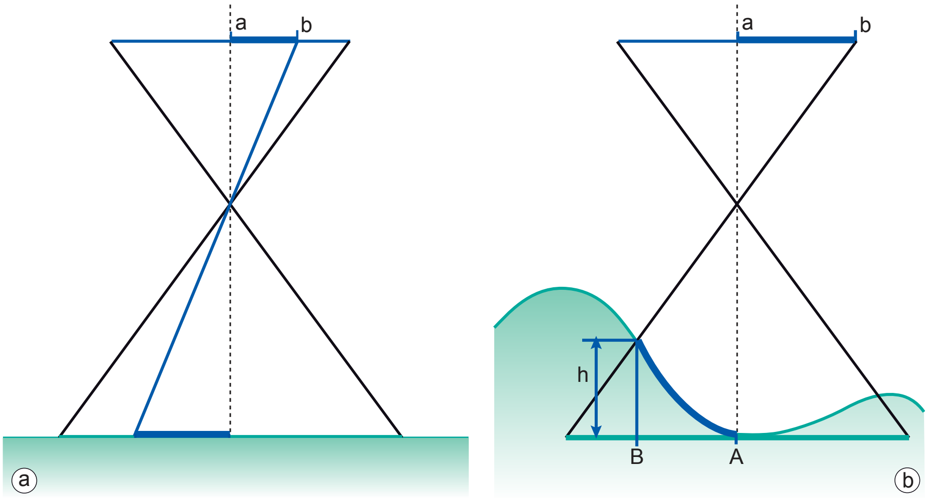


Figure 5.24
Illustration of the effect of terrain topography on the relationship between $A - B$ on the ground and $a - b$ in the image: (a) flat terrain, (b) significant elevation difference.

If the scene is just barren land, we cannot see any relief displacement. However, we can see relief displacement if there are protruding objects in the scene (occasionally referred to as height displacement). On large-scale photographs or very high-resolution space-borne images, buildings and trees appear to lean outwards, away from the nadir point (Figure 5.25).

The main effect of relief displacement is that inaccurate or wrong map coordinates will

Figure 5.25
Fragment of a large-scale aerial photograph of the centre of Enschede. Note the effect of height displacement on the higher buildings.

be obtained when, for example, digitizing images without further correction. Whether relief displacement should be considered in the geometric processing of RS data depends on its impact on the accuracy of the geometric information derived from the images. Relief displacement can be corrected for if information about terrain relief is available (in the form of a DTM); see Subsection 5.3.4 for more details. It is also useful to remember that it is relief displacement that allows us to perceive depth when looking at a stereograph and to extract 3D information from such images.

effect of relief displacement

Two-dimensional approaches

This subsection deals with the geometric processing of RS images in situations where relief displacement can be neglected, for example for a scanned aerial photograph of flat terrain. For practical purposes,“flat” may be considered as $h/H < 1/1000$, though this also depends on project accuracy requirements; h stands for relative terrain elevation, H for flying height. For space-borne images of medium spatial resolution, relief displacement is usually less than a few pixels in magnitude and thus less important, as long as near-vertical images are acquired. The objective of 2D georeferencing is to relate the image coordinate system to a specific map coordinate system (Figure 5.26).

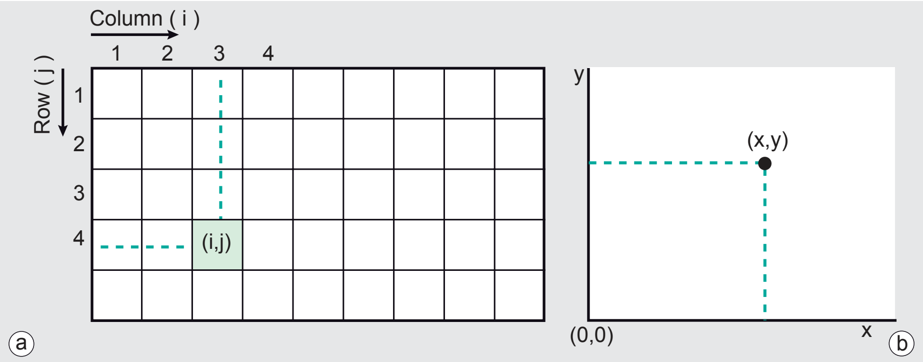


Figure 5.26
Coordinate system of (a) the image defined by rows and columns, and (b) the map with x- and y-axes.

5.3.2 Georeferencing

transformation

The simplest way to link image coordinates to map coordinates is to use a transformation formula. A *geometric transformation* is a function that relates the coordinates of two systems. A transformation relating (x, y) to (i, j) is commonly defined by linear equations, such as: $x = 3 + 5i$, and $y = -2 + 2.5j$.

Using the above transformation, for example, the image position $(i = 3, j = 4)$ corresponds to map coordinates $(x = 18, y = 8)$. Once the transformation parameters have been determined, the map coordinates for each pixel can be calculated. This implies that we can superimpose data that are given in the map coordinate system on the image vector, or that we can store features by map coordinates when applying on-screen digitizing. Note that the image in the case of georeferencing remains stored in the original (i, j) raster structure and that its geometry is not altered. As we will see in Subsection 5.3.3, transformations can also be used to change the actual shape of an image and thus make it geometrically equivalent to the map.

type of transformation

The process of georeferencing involves two steps: (1) selection of the appropriate type of transformation, and (2) determination of the transformation parameters. The type of transformation depends mainly on the sensor–platform system used. For aerial photographs (of flat terrain) what is known as "projective transformation" models well the effect of pitch and roll (see Figures 4.27, 5.23, and 5.28). Polynomial transformation, which enables 1st, 2nd to nth order transformations, is a more general type of transformation. In many situations a 1st order transformation is adequate. Such transformation relates map coordinates (x, y) with image coordinates (i, j) as follows:

$$x = a + bi + cj \tag{5.4}$$

$$y = d + ei + fj \tag{5.5}$$

ground control points

Equations 5.4 and 5.5 require that six parameters (a to f) be determined. The transformation parameters can be determined by means of *ground control points* (GCPs). GCPs are points that can be clearly identified in the image and on the target map. The target map could be a topographic map or another image that has been transformed beforehand to the desired map projection system. The operator then needs to identify corresponding points on both images. The image and map scale determine which points are suitable. Typical examples of suitable points are road crossings, crossings of waterways and salient morphological structures. Another possibility is to identify points in the image and to measure the coordinates of these points in the field, for example by GPS, and then transform those to map coordinates. It is important to note that it can be quite difficult to identify good GCPs in an image, especially in lower-resolution space-borne images. Once a sufficient number of GCPs have been specified, software is used to determine the parameters a to f of the Equations 5.4 and 5.5 and quality indications.

number of GCPs

To solve the 1st order polynomial equations, only three GCPs are required; nevertheless, you should use more points than the strict minimum. Using merely the minimum number of points for solving the system of equations would obviously lead to a wrong transformation if you made an error in one of the measurements, whereas including more points for calculating the transformation parameters enables software to also compute the error of the transformation. Table 5.3 gives an example of the input and output of a georeferencing computation in which five GCPs have been used. Each GCP is listed with its image coordinates (i, j) and its map coordinates (x, y).

Software performs a "least-squares adjustment" to determine the transformation pa-

GCP	i	j	x	y	x_c	y_c	d_x	d_y
1	254	68	958	155	958.552	154.935	0.552	-0.065
2	149	22	936	151	934.576	150.401	-1.424	-0.599
3	40	132	916	176	917.732	177.087	1.732	1.087
4	26	269	923	206	921.835	204.966	-1.165	-1.034
5	193	228	954	189	954.146	189.459	0.146	0.459

Table 5.3
A set of five ground control points, which are used to determine a 1st order transformation. x_c and y_c are calculated using the transformation, d_x and d_y are the residual errors.

rameters. The least squares adjustment ensures an overall best fit of the GCPs. We then use the computed parameter values to calculate coordinates (x_c, y_c) for any image point (pixel) of interest:

$$x_c = 902.76 + 0.206i + 0.051j,$$

and

$$y_c = 152.579 - 0.044i + 0.199j.$$

For example, for the pixel corresponding to GCP 1 ($i = 254$, $j = 68$) we can calculate the transformed image coordinates x_c and y_c as 958.552 and 154.935, respectively. These values deviate slightly from the input map coordinates (as measured on the map). Discrepancies between measured and transformed coordinates of GCPs are called residual errors (*residuals* for short). The residuals are listed in the table as d_x and d_y. Their magnitude is an indicator of the quality of the transformation. Residual errors can be used to analyse whether all GCPs have been correctly determined.

residuals

The overall accuracy of a transformation is either stated in the accuracy reportusually provided by software in terms of variances or as *Root Mean Square Error* (RMSE), which calculates a mean value from the residuals (at check points). The RMSE in the x-direction, m_x, is calculated using the following equation:

$$m_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \delta x_i{}^2}. \tag{5.6}$$

For the y-direction, a similar equation can be used to calculate m_y. The overall error, m_p, is calculated by:

$$m_p = \sqrt{m_x{}^2 + m_y{}^2}. \tag{5.7}$$

For the example data set given in Table 5.3, the residuals m_x, m_y and m_p are 1.159, 0.752 and 1.381, respectively. The RMSE is a convenient measure of overall accuracy, but it does not tell us which parts of the image are accurately transformed and which parts are not. Note also that the RMSE is only valid for the area that is bounded by the GCPs. In the selection of GCPs, therefore, points should be well distributed and include locations near the edges of the image.

5.3.3 Geocoding

The previous subsection explained that two-dimensional coordinate systems, e.g. an image system and a map system, can be related using geometric transformations. This georeferencing approach is useful in many situations. However, some situations a *geocoding* approach, in which the row–column structure of the image is also transformed, is required. Geocoding is required when different images need to be combined or when the images are used in a GIS environment that requires all data to be

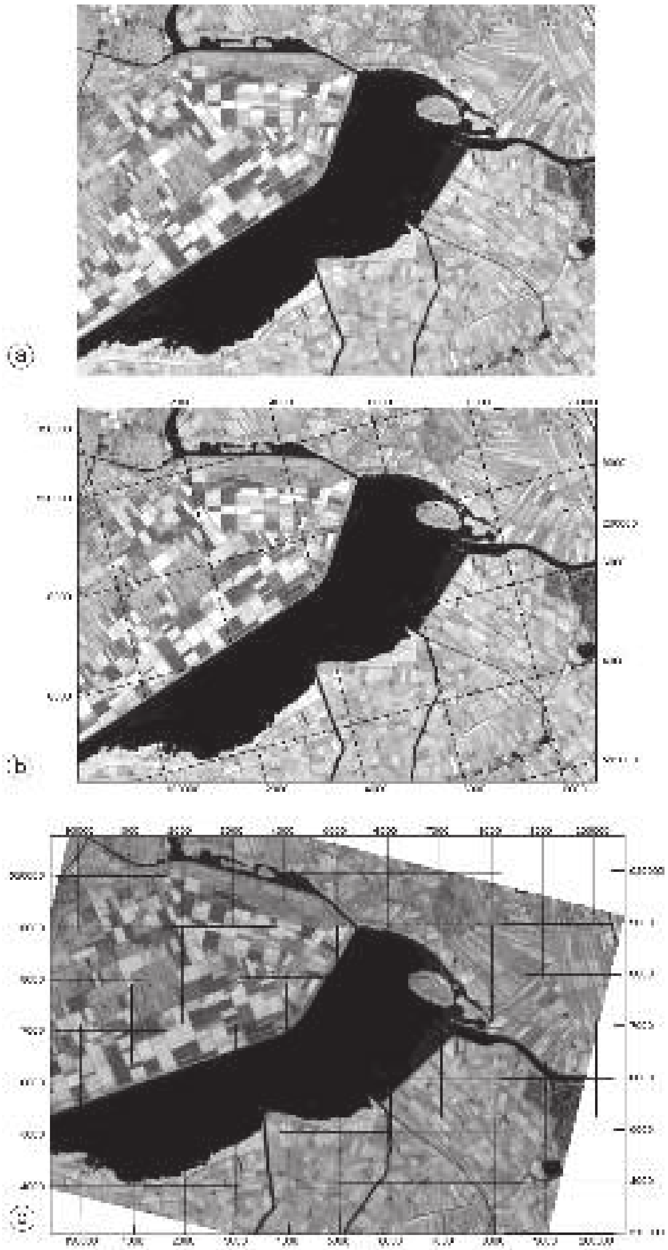


Figure 5.27
Original (top), georeferenced (middle) and geocoded (bottom) image of a part of Flevoland.

stored in the same map projection. The effect of georeferencing and geocoding is illustrated by Figure 5.27. Distinguishing georeferencing and geocoding is conceptually useful, but to the casual software user it is often not apparent whether only georeferencing has been applied in certain image manipulations or also geocoding.

Geocoding is georeferencing with subsequent *resampling* of the image raster. This means that a new image raster is defined along the $x-$ and y-axes of the selected map projection. The geocoding process comprises three main steps: (1) selection of a new grid spacing; (2) projection (using the transformation parameters) of each new raster element onto the original image; and (3) determination and storage of a DN for the new pixel.

Figure 5.28 shows four transformation types that are frequently used in RS. The types shown increase from left to right in complexity and number of parameters required. In a conformal transformation, the image shape, including right angles, are retained. Therefore, only four parameters are needed to describe a shift along the x- and y-axes, a scale change, and the rotation. However, if you want to geocode an image to make it fit with another image or map, a higher-order transformation may be required, such as a projective or polynomial transformation.

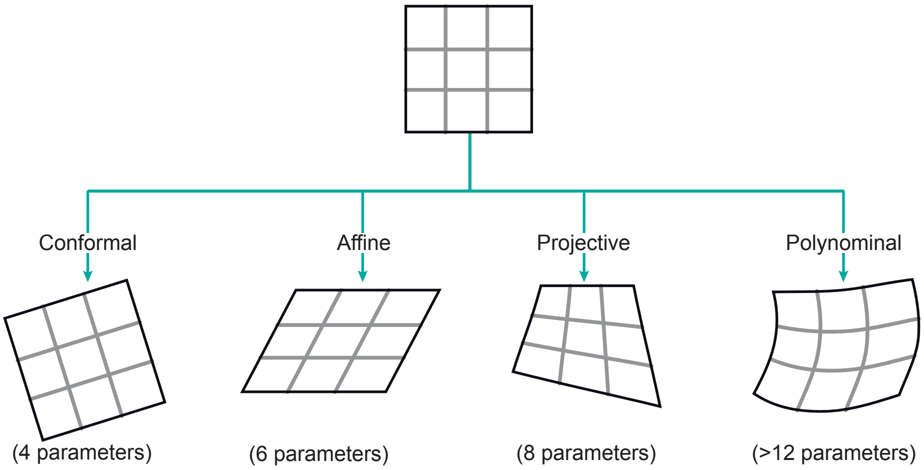


Figure 5.28
Illustration of different image transformation types and the number of required parameters.

Most of the projected raster elements of the new image will not match precisely with raster elements in the original image, as Figure 5.29 illustrates. Since raster data are stored in a regular row–column pattern, we need to calculate DNs for the pixel pattern of the corrected image intended. This calculation is performed by interpolation. This is called *resampling* of the original image.

resampling

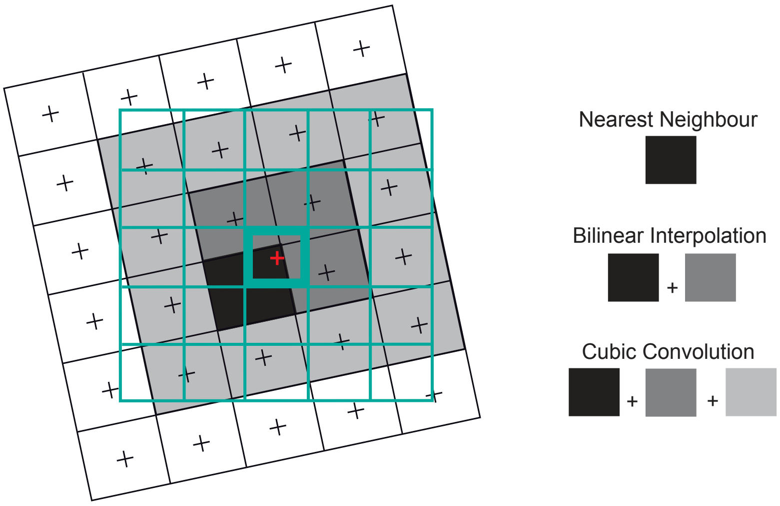


Figure 5.29
Principle of resampling using nearest neighbour, bilinear, and bicubic interpolation.

For resampling, we usually use very simple interpolation methods, the main ones being nearest neighbour, bilinear, and bicubic interpolation (Figure 5.29). Consider the green grid to be the output image to be created. To determine the value of the ceninterpolation

tre pixel (bold), in *nearest neighbour interpolation* the value of the nearest original pixel is assigned, i.e. the value of the black pixel in this example. Note that the respective pixel centres, marked by small crosses, are always used for this process. In *bilinear interpolation*, a linear weighted average is calculated for the four nearest pixels in the original image (dark grey and black pixels). In *bicubic interpolation* a cubic weighted average of the values of 16 surrounding pixels (the black and all grey pixels) is calculated. Note that some software uses the terms "bilinear convolution" and "cubic convolution" instead of the terms introduced above.

choice of method

The choice of the resampling algorithm depends, among other things, on the ratio between input and output pixel size and the intended use of the resampled image. Nearest neighbour resampling can lead to the edges of features being offset in a step-like pattern. However, since the value of the original cell is assigned to a new cell without being changed, all spectral information is retained, which means that the resampled image is still useful in applications such as digital image classification (see Section 6.2). The spatial information, on the other hand, may be altered in this process, since some original pixels may be omitted from the output image, or appear twice. Bilinear and bicubic interpolation reduce this effect and lead to smoother images. However, because the values of a number of pixels are averaged, radiometric information is changed (Figure 5.30).

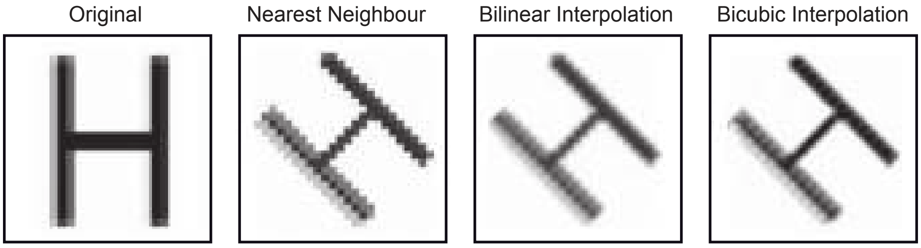


Figure 5.30
The effect of nearest neighbour and bilinear and bicubic resampling of the original data.

5.3.4 Three-dimensional approaches

In mapping terrain, we have to consider its vertical extent (elevation and height) in two types of situations:

- We want 2D geospatial data that describes the horizontal position of terrain features, but the terrain under consideration has large elevation differences. Elevation differences in the scene cause relief displacement in the image. Digitizing in the image without taking into account relief displacement causes errors in computed map coordinates. If the positional errors are larger than would be tolerated by the application (or map specifications), we should not use simple georeferencing and geocoding.
- We want 3D data.

terrain relief

When wishing to map terrain with an increasing degree of refinement, we have to first clarify what we mean by *terrain*. *Terrain* as described by a topographic map has two very different aspects: (1) there are agricultural fields and roads, forests and waterways, buildings and swamps, barren land and lakes; and (2) there is elevation changing with position—at least in most regions on Earth. We refer to land cover, topographic objects, etc. as *terrain features*; we show them on a (2D) map as areas, lines and point symbols. We refer to the shape of the ground surface as *terrain relief*, which we show on a topographic map by contour lines and/or relief shading. A *contour* *line* on a topographic map is a line of constant elevation. Given contour lines, we can determine the elevation at any arbitrary point by interpolating between contour lines.

We could digitize the contour lines of a topographic map. The outcome would be a data set consisting of (X, Y, Z) coordinates of many points. Such a data set is the basis of a *digital terrain relief model* (DTM). We then need a computer program to utilize such a list of coordinate of points of the ground surface, to compute elevation at any horizontal position of interest and derive other terrain relief information such as slope and aspect. The idea of a DTM was developed in the late 1950s at MIT for computerizing highway design. Fifty years later, we use DTMs in all geosciences for a wide range of applications and we have remote sensors specifically built to supply us with data for DTMs (SRTM, SPOT-5 HRS, etc.). One of the applications of a DTM is to accomplish *digital monoplotting* and *orthoimage* production, as outlined later in this subsection.

A *DTM* is a digital representation of terrain relief, i.e. a model of the shape of the ground. We have a variety of sensors at our disposal that can provide us with 3D data: line cameras, frame cameras, laser scanners and microwave radar instruments. They can all produce (X, Y, Z) coordinates of terrain points, but not all the terrain points will be points on the ground surface. Consider a stereo pair of photographs, or a stereo pair from SPOT-5, of a tropical rainforest. Will you be able to see the ground? Coordinates obtained will pertain to points of the terrain relief. Since a model based on such data is not a DTM, we refer to it as digital surface model (DSM). The difference between a DTM and a DSM is illustrated by Figure 5.31; we need to filter DSM data to obtain DTM data.

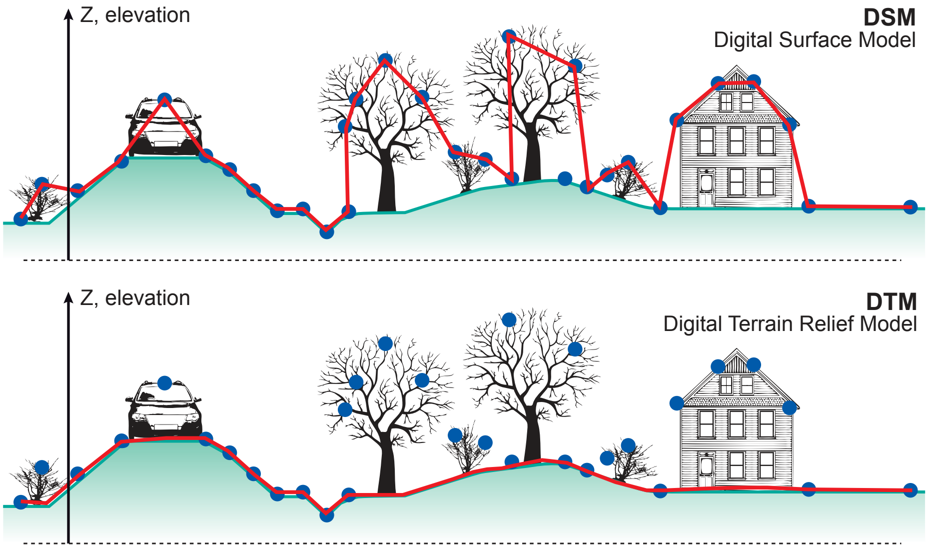


Figure 5.31
The difference between a DTM and a DSM.

In terrain modelling, it is handy to choose the coordinate system such that Z is the variable for elevation. If we model a surface digitally by nothing else than elevation values Z at horizontal positions (X, Y), why not call such a model a *digital elevation model* (DEM)? The term DEM was introduced in the 1970s with the purpose of distinguishing the simplest form of terrain relief modelling from more complex types of digital surface representation. Originally the term DEM was exclusively used for raster representations (thus elevation values given at the intersection nodes of a regular grid). Note that both a DTM and DSM can be a DEM and, moreover, “elevation” would not have to relate to terrain but could relate to some subsurface layer such as groundwater layers, soil layers or the ocean floor. Unfortunately, you will find in the literature variations in the use of the terms introduced above. This is particularly so for DEM, which is often used carelessly. In this context, it is also worth mentioning the misuse of “topography” as synonym for terrain relief.

5.3.5 Orientation

The purpose of *camera orientation* is to obtain the parameter values for transforming terrain coordinates (X, Y, Z) to image coordinates, and vice versa. In solving the orientation problem we assume that any terrain point and its image lie on a straight line that passes through the projection centre (i.e. the lens). This assumption about the imaging geometry of a camera is called the *collinearity* relationship and it does not take into consideration that atmospheric refraction has the effect of slightly bending light rays. We solve the problem of orienting a single image with respect to the terrain in two steps: *interior orientation* and *exterior orientation*.

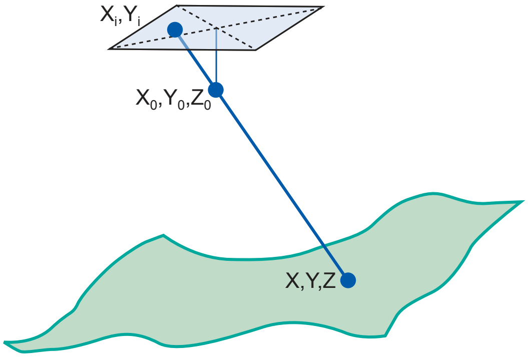


Figure 5.32
Illustration of the collinearity concept, where image point, lens centre and terrain point all lie on one straight line.

Orienting a single image as obtained by a line or frame camera

interior orientation

principal distance

principal point

Interior orientation determines the position of the projection centre with respect to the image. The problem to solve is different for digital cameras and film cameras. In the case of a digital camera (line or frame camera, space-borne or aerial camera), the position of the projection centre with respect to the CCD array does not change from image to image, unless there are extreme temperature or pressure changes. For standard applications, we can assume that the position of the projection centre does not change when defined in the row–column system of the digital image. Two parameters are needed, the *principal distance* and the *principal point* (Figure 5.33); they are both determined by camera calibration. The *principal distance* is the mathematical abstraction of the focal length (which is a physical property of a lens). The *principal point* is the point of intersection of the perpendicular from the projection centre point with the image plane.

The camera calibration report states the row (and column) number of the principal point and the principal distance. When using digital photogrammetric software for working with images from digital aerial cameras, the user only has to enter the principal distance, the position of the principal point and the pixel size to define the interior orientation. In the case of a film camera, the position of the principal point is only fixed in the camera and determined by camera calibration with respect to the fiducial marks. When scanning photographs, the principal point will have different positions in the row–column system of each image, because each image will be placed slightly differently in the scanner. Therefore, you have to measure for every image the imaged fiducial marks and calculate the transformation onto the fiducial marks as given by the calibration report. Digital photogrammetric software can then relate row–column coordinates of any image point to image coordinates (x, y) at the time of exposure (Figure 5.33).

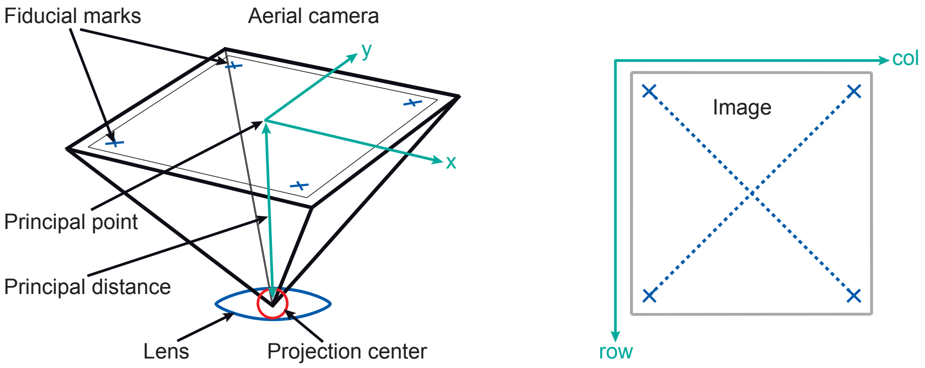


Figure 5.33
Inner geometry of a camera and the associated image.

exterior orientation

sensor position and attitude

The *exterior orientation* determines the position of the projection centre with respect to the terrain and also the attitude of the sensor. Frame cameras (film as well as digital ones) acquire the entire image at once, thus three coordinates for the projection centre (X, Y, Z) and three angles for the attitude of the sensor (angles relating to roll, pitch, and yaw) are sufficient to define the exterior orientation of an entire image. Line cameras yield images for which each image line has its own projection centre and the sensor attitude may change from one line to the next. Satellites have a very smooth motion and airborne line cameras are mounted on a stabilized platform so that any attitude change is gradual and small. Mathematically, we can model the variation of an attitude angle through a scene with a low-degree polynomial.

RPC

Images from modern, high-resolution line cameras on satellites come with *rational polynomial coefficients* (RPCs). The rational polynomials define by good approximation the relationship between the image coordinates of an entire frame (in terms of row–column pixel positions) and terrain coordinates. The nice thing is that RPCs are understood by RS software such as ERDAS and that they take care of interior and exterior orientation. For cases in which RPCs are not given or are considered not accurate enough, the exterior orientation needs to be solved in one of the following ways:

- *Indirect camera orientation*: identify GCPs in the image and measure the row and column coordinates; acquire (X, Y, Z) coordinates for these points, e.g. by GPS or a sufficiently accurate topographic map; use adequate software to calculate the exterior orientation parameters, after having completed the interior orientation.
- *Direct camera orientation*: during image acquisition, make use of GPS and IMU recordings by employing digital photogrammetric software.
- *Integrated camera orientation*, which is a combination of (a) and (b).

For high resolution satellite images such as Ikonos or QuickBird, adding one GCP can already considerably improve the exterior orientation as defined by the RPC. For Cartosat images, it is advisable to improve the exterior orientation by at least five GCPs. For orienting a frame camera image you need at least three GCPs (unless you also have GPS and IMU data). After orientation, you can use the terrain coordinates of any reference point and calculate its position in the image. The differences between measured and calculated image coordinates (the residuals) allow you to estimate the accuracy of orientation. As you may guess, advanced camera/sensor/image orientation is a topic for further study.

Orienting a stereo-image pair obtained by a line or frame camera

relative orientation

absolute orientation

The standard procedure is to individually orient each image. In the case of images originating from a frame camera, we can, however, readily make use of the fact that both images partly cover the same area. Instead of doing two independent exterior orientations, we can better first do a *relative orientation* of the two images, followed by an *absolute orientation* of the pair to the terrain coordinate system. The *relative orientation* will cause the imaging rays of corresponding points to intersect more accurately than if you orient one image without the knowledge of the other. You do not have to measure points with known terrain coordinates to solve the relative orientation problem; you only need to measure image coordinates of corresponding points in the two images (after the individual interior orientations have been established). Measuring of corresponding points can even be done automatically (by *image matching*, see below). For absolute orientation, at least three GCPs are needed. The idea of splitting up the exterior orientation into a relative orientation and an absolute orientation is also used for orienting a whole block of overlapping images, not just two. The advantage is that we need a only a few GCPs (which are usually expensive to acquire) and we can still obtain accurate transformation parameters for each image. The method is known as *aerial triangulation*.

5.3.6 Monoplotting

Suppose you need to derive accurate planimetric coordinates of features expressed in a specific map projection from a single aerial photograph. For flat terrain, this can be achieved using a vertical photograph and a georeferencing approach. Recall from the earlier discussion on relief displacement (Figure 5.24) how elevation differences lead to distortions in the image, preventing the use of such data for direct measurements. Therefore, if there is significant terrain relief, the resulting relief displacement has to be corrected. The method of *monoplotting* was developed (with major research input from ITC) for just this purpose.

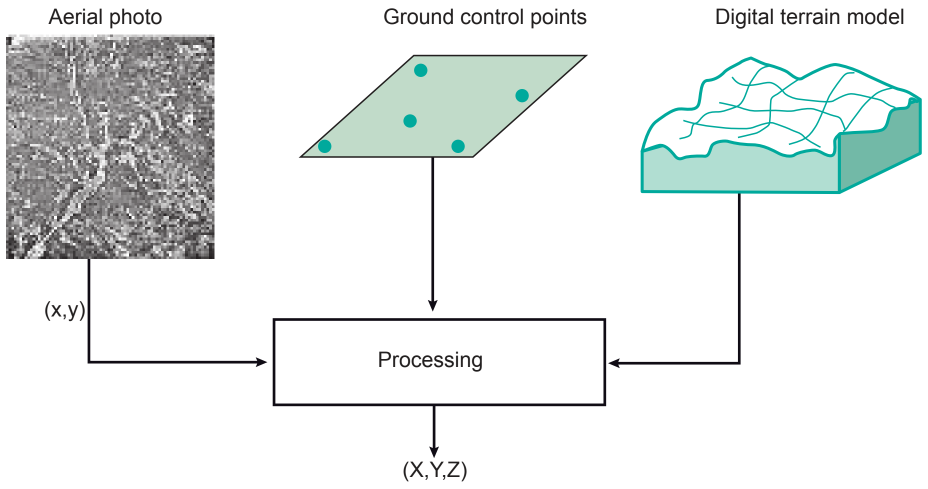


Figure 5.34
The process of digital monoplotting enables accurate determination of terrain coordinates from a single aerial photograph.

3D data from a single image

Monoplotting is based on the reconstruction of the position of the camera at the moment of image exposure with respect to the GCPs, i.e. the terrain. This is achieved by identifying several (at least four) GCPs for which both the photo and map coordinates are known. Information about the terrain relief is supplied by a DTM of adequate accuracy. The DTM should be given in the required map projection system and the elevations should be expressed in an adequate vertical reference system. When digitizing features from the photograph, the computer uses the DTM to calculate the relief displacement for every point and corrects for it (Figure 5.34). A monoplotting approach is possible by using a hardcopy image on a digitizer tablet or by on-screen digitizing on a computer monitor. In the latter case, vector information can be superimposed on the image to update the changed features. Note that monoplotting is a (real-time) correction procedure and does not yield a new image, i.e. no resampling is carried out.

5.3.7 Orthoimage production

geocoding an image of rough terrain

Monoplotting can be considered a georeferencing procedure that incorporates corrections for relief displacement without involving any resampling. For some applications, however, it is useful to actually correct the photograph or RS image, taking into account the effect of terrain relief. In such cases, the image should be transformed and resampled (making use of a DTM) into a product with the geometric properties of a specific map projection. Such an image is called an *orthophoto*.

The production of orthophotos is quite similar to the process of monoplotting. Consider a scanned aerial photograph. First, the photo is oriented using ground control points. The terrain elevation differences are modelled by a DTM. The computer then calculates the position in the original photo for each output pixel. Using one of the resampling algorithms, the output value is determined and stored in the required raster. The result is geometrically equivalent to a map, i.e. direct distance or area measurements on the orthoimage can be carried out.

5.3.8 Stereo restitution

3D data from a stereo pair

After relative orientation of a stereo pair, we can exploit the 3D impression gained from the stereo model to make measurements in 3D. The measurements made in a stereo model make use of a phenomenon known as *parallax* (Figure 5.35). Parallax refers to the fact that an object photographed from different camera locations (e.g. from a moving aircraft) has different relative positions in the two images. In other words, there is an apparent displacement of an object as it is observed from different locations. Figure 5.35 illustrates that points at two different elevations, regardless of whether it is the top and bottom of a hill or of a building, experience a relative shift.

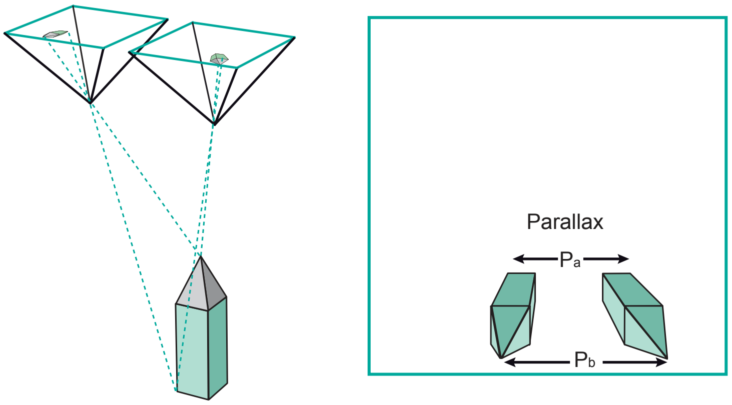


Figure 5.35
The same building is observed from two different positions. Because of the height of the building, the positions of the building top and base relative to the photo centres are different. This difference (parallax) can be used to calculate its height.

The measurement of the difference in position is basic input for elevation calculations. We could use stereo restitution to measure (X, Y, Z) coordinates of many points, in this way obtaining data for a DTM. This is both a boring and an error-prone process. Digital photogrammetric software can do this job automatically with a high degree of success—after some 30 years of research and development—using *image matching*. Manual measurements are then only needed as a supplement for difficult areas. The main purpose of stereo restitution is to collect 3D vector data of objects. Unlike monoplotting, elevation is measured directly and not interpolated from a DTM, hence the coordinate accuracy can be higher. Another main advantage of stereo restitution is the better image interpretability obtained, because one can see and interpret in 3D.

image matching

A stereo model enables parallax measurement using a special 3D cursor. If the stereo model is appropriately oriented, the parallax measurements yield (X, Y, Z) coordinates. *Analogue* systems use hardcopy images and perform the computation by mechanical, optical or electrical means. *Analytical* systems also use hardcopy images, but do the computation digitally, while in modern *digital* systems, both the images and the computation are digital. By using digital instruments, we cannot only make a few spot elevation measurements, but also generate automatically a complete DSM for the overlapping part of the two images. Recall, however, that reliable elevation values can only be extracted if the orientation steps were carried out accurately, using reliable ground control points.

analogue,
analytical,
digital systems