Earth Observation:
Data, Processing and Applications

Volume 2C: Processing—Image Transformations

Disclaimer:
While every effort has been made to ensure its accuracy, the CRCSI does not offer any express or implied warranties or representations as to the accuracy or completeness of the information contained herein. The CRCSI and its employees and agents accept no liability in negligence for the information (or the use of such information) provided in this report.

Background image on previous page: Principal Components Analysis of Landsat-5 image of Lake Eyre, South Australia, filling on 9 May, 2009, displayed using PCs 3, 4, 1 as RGB.
Source: Norman Mueller, Geoscience Australia

Acknowledgements

1   Harrison, B.A., and Jupp, D.L.B. (1989) *Introduction to Remotely Sensed Data: Part ONE of the microBRIAN Resource Manual.* CSIRO, Melbourne. 156pp.
Harrison, B.A., and Jupp, D.L.B. (1990) *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.
Harrison, B.A., and Jupp, D.L.B. (1992) *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual.* MPA, Melbourne.
Harrison, B.A., and Jupp, D.L.B. (1993) *Image Classification and Analysis: Part THREE of the microBRIAN Resource Manual.* MPA, Melbourne.
2   Formerly Spookfish

Table of Contents

ii

Non-Linear Operations   105

List of Figures

vi

vii

List of Tables

List of Excurses

x

Review Topics

The following two sections review some principles that were introduced in earlier sub-volumes of this series of publications and have particular relevance to EO image transformations:

- overview—a reminder of some foundations when processing EO imagery (see Section 1); and
- rescaling—changing the data range for an image channel (see Section 2).

Contents
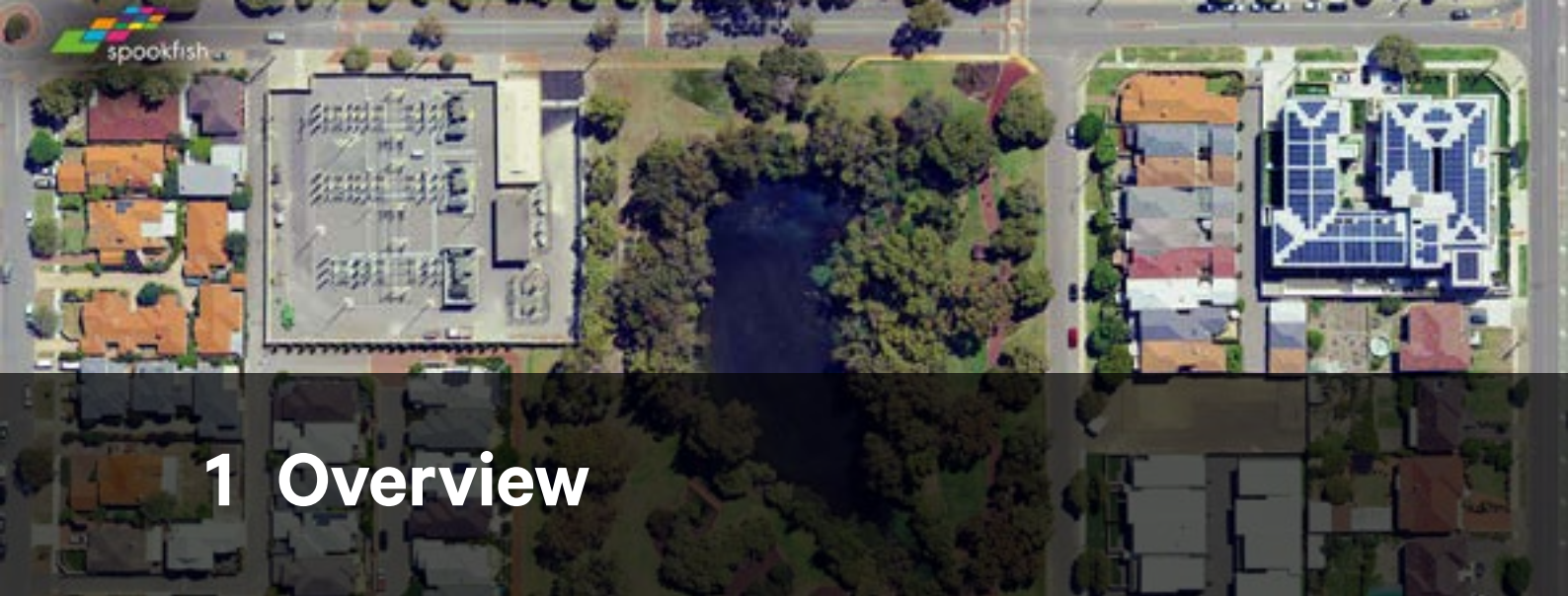
1 Overview

Basic image operations are introduced in Volume 2A of this series and specific analytical tools and methods are covered in Volumes 2B (Image Rectification), 2D (Image Integration) and 2E (Image Classification). This volume details various image transformations, which may be used for interpretation, analysis or modelling. The modelling stage relates image values to data from other sources and accesses a wide variety of transformations to determine the relationships between different data types, and is further discussed in the context of specific application areas in Volume 3: Applications.

Image transformations allow image data values to be changed by some consistent and defined process. Such processes are usually described in terms of modifying some or all of the existing pixel values in an image. While modern image processing systems allow transformations to be used without reference to the underlying algorithms, knowledge of their mathematical foundations does ensure that such transformations are applied appropriately. In the context of Earth Observation (EO), appropriate processing is more likely to result in consistent and repeatable results, and enables users to use transformed data with understanding and confidence.

In this introductory section we will review some of the basic concepts that underlie image transformations commonly applied to EO data, namely:

- the electromagnetic spectrum—wavelength regions and commonly used EO sensor bands (see Section 1.1);
- image data characteristics—structure, scales of measurement, data scales and sampling dimensions (see Section 1.2);
- spectral statistics—for single channels and pairs of channels (see Section 1.3); and
- processing operations—linear versus non-linear (see Section 1.4).

The universe is transformation: life is opinion.
(Marcus Aurelius)

Background image: High resolution aerial image of residential area in North Perth, acquired on 18 September 2017. **Source:** © EagleView

1.1 The Electromagnetic Spectrum

One of the most commonly detected energy sources for EO is electromagnetic radiation (EMR; see Volume 1). For convenience, the electromagnetic spectrum is labelled in terms of wavelength regions (see Figure 1.1).

Commonly accepted ranges for the EMR spectral regions most relevant to EO are:

- visible: 0.38–0.7 μm;
- infrared (IR): 0.7–1,000 μm; and
- microwave: > 1 mm–1 m.

Infrared wavelengths are often further sub-divided into the regions:

- near infrared (NIR): 0.7–1.3 μm;
- short wave infrared (SWIR): 1.3–3 μm;
- middle infrared (MIR) or mid-wavelength (or medium wave) infrared (MWIR): 3–8 μm;
- thermal infrared (TIR) or long wavelength infrared (LWIR): 8–15 μm; and
- far infrared (FIR): 15–1,000 μm.

These terms will be used in the following text in reference to EO image data channels. Within these broad regions, multiple EO image bands can be defined. For example, the wavelength ranges detected by five different sensors carried on recent and current EO satellites are summarised in Figure 1.2.

Figure 1.1 Electromagnetic spectrum

Note: Cosmic rays are high energy, charged particles that travel at close to the speed of light. They are not EMR but included for comparison. Scales are logarithmic
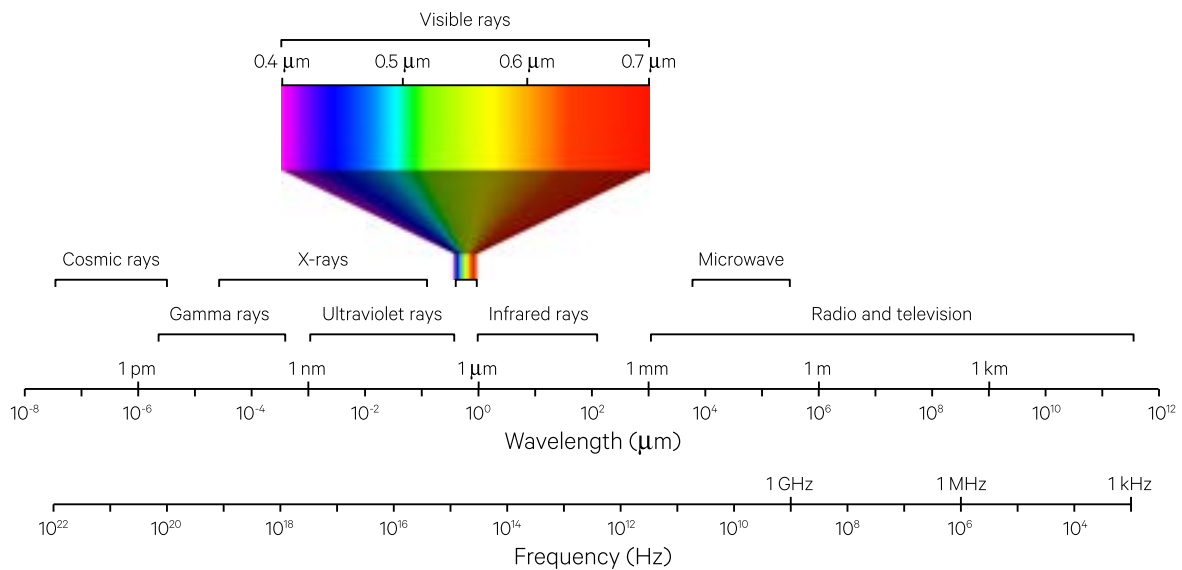
*As Gregor Samsa awoke one morning from uneasy dreams
he found himself transformed in his bed into a gigantic insect.
(Franz Kafka, The Metamorphosis)*

4

Earth Observation: Data, Processing and Applications. Volume 2: Processing

Figure 1.2 Comparison of spectral bands in different EO sensors

This diagram compares spectral bands from five commonly used, public good satellites by positioning their spectral bandwidths in the optical and thermal regions of the electromagnetic spectrum. The percentage of EMR transmission through the atmosphere in these ranges of wavelengths is shown in grey (see Volume 1B). Note that the wavelength axis is not continuous.
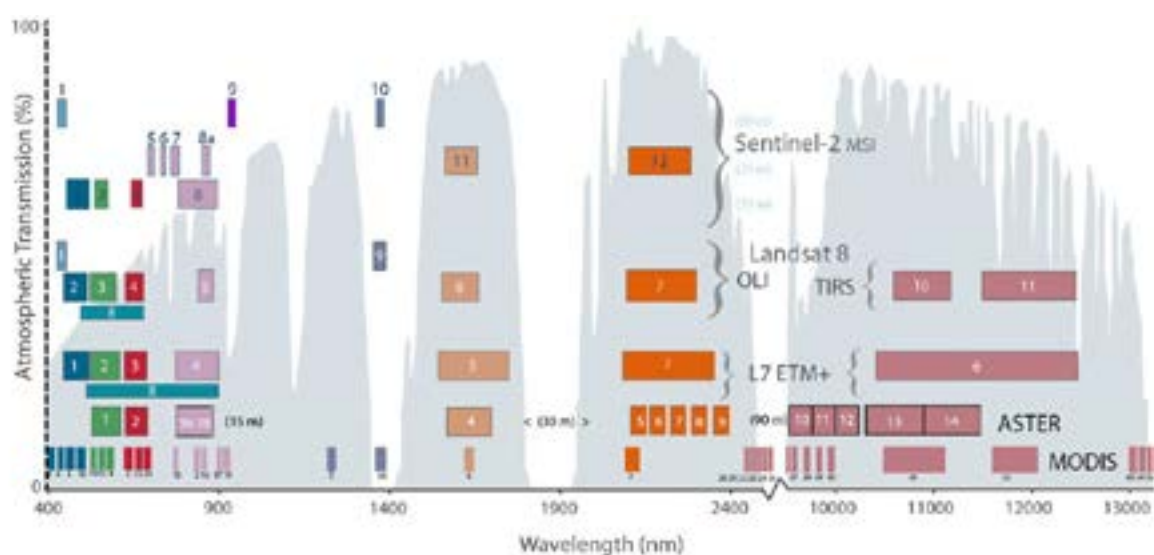
Source: NASA. Retrieved from: *https://landsat.gsfc.nasa.gov/wp-content/uploads/2016/10/all_Landsat_bands.png*

1.2 Image Data Characteristics

Characteristics of image data are reviewed below in terms of:

- structure (see Section 1.2.1);
- measurement scales (see Section 1.2.2);
- data scales (see Section 1.2.3); and
- sampling dimensions (see Section 1.2.4).

1.2.1 Structure

The structure of digital imagery is introduced in Volume 2A—Section 1. By assuming that the tessellation grid used to record EO image values is regular within and between all image channels (or bands), the resulting image dataset can be viewed as a three-dimensional structure (see Figure 1.3). This three-dimensional dataset can then be examined, analysed and transformed used a range of statistical tools. Transformed imagery may then be correlated with other datasets, such as crop yield or water depth (see Volume 3).

Figure 1.3 Structure of EO image

A multi-channel image comprises three data dimensions— pixels, lines and channels. Pixels (or 'elements' or 'samples') along the lines can be considered as columns in a matrix. Lines in the image are like rows of the matrix, with each channel being an attribute layer in the matrix.

Row in matrix = image line

Column in matrix = pixel position in image line

Layer in matrix = individual attribute = channel in image
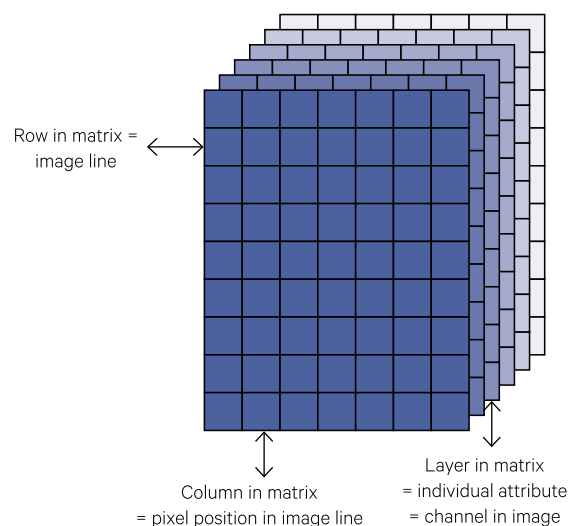
1.2.2 Scales of measurement

Statisticians differentiate four scales of measurement for defining numeric variables (Anderberg, 1973):

- nominal—refers simply to categories which define different states (such as colour names for different hues, e.g. red, yellow, green);
- ordinal—involves an implied ordering between different states (such as categories of colour brightness, or ranges in topographic elevation);
- interval—measures the extent of difference between states (such as relative illumination for a given wavelength or bandwidth); and
- ratio—denotes a scaled measure with a meaningful zero point (such as absolute illumination for a given wavelength or bandwidth). The ratio of two ratio variables could be computed without any adjustment to their values.

EO image channels would generally be considered as interval variables. With appropriate processing, however, they could be treated as ratio variables. Classified data can also become image channels, which would be considered as nominal or possibly ordinal data (see Volume 2E). Much ancillary data, such as land cover categories, would most likely be nominal or ordinal, although some, such as DEM, may be interval or even ratio (see Volume 2D). These distinctions are important when considering the types of transformations that may be applied to EO imagery. The relevance of scales of measurement is further discussed in Volume 2D (in terms of integrating ancillary data) and in Volume 2E (in terms of the image classification process).

1.2.3 Data scales

Interval and ratio variables (see Section 1.2.2) can be referenced to a linear scale or a logartihmic scale. A linear scale has equal divisions for equal values (see Figure 1.4a). In a logarithmic (or log) scale, data points are represented by the logarithm of their value (see Figure 1.4b).

A logarithmic scale is generally used when there is a wide range of values and the distribution of data is skewed. While EO image values are acquired on a linear scale, it can be advantageous to transform image data to a logarithmic scale to enhance low values in one or more channels (see Section 2.2) or apply other transformations (see Section 10.2.2).

A logarithmic function represents the equation of:

$$y = \log_{10}\left(x\right)$$

where

$$x = 10^y$$

The exponential function is expressed as:

$$y = \exp_{10}\left(x\right)$$

where

$$\exp_{10}\left(x\right) = 10^y$$

These functions are often expressed more generally as $\log_a\left(x\right)$ or a^x where the choice of a affects the scaling and properties of the functions. When a is the natural constant e, the logarithmic function, $\log_e\left(x\right)$ or ln, is referred to as the Napierian logarithm and has special properties:

$$a^x = e^{\log\left(a^x\right)} = e^{x\log\left(a\right)}$$

The base 10 number system computes these functions with the simply remembered values of:

$$\log_{10}\left(1\right) = 0 \qquad \exp_{10}\left(0\right) = 1$$
$$\log_{10}\left(10\right) = 1 \qquad \exp_{10}\left(1\right) = 10$$
$$\log_{10}\left(100\right) = 2 \quad \exp_{10}\left(2\right) = 100$$

as illustrated in the Figure 1.5.

Figure 1.4 Data scales

a. Linear

b. Logarithmic

6
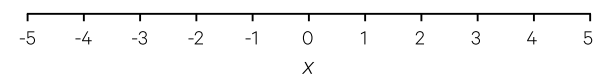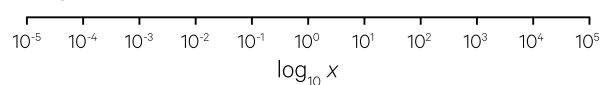
Figure 1.5 Logarithmic and exponential functions

a. Logarithmic function: $y = \log_{10}(x)$
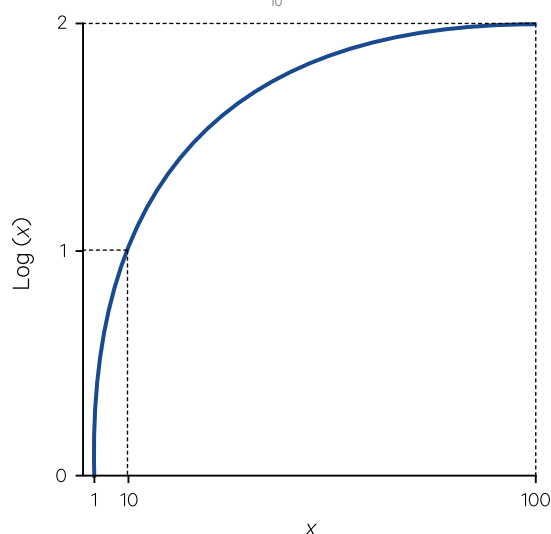
b. Exponential function: $y = \exp_{10}(x)$
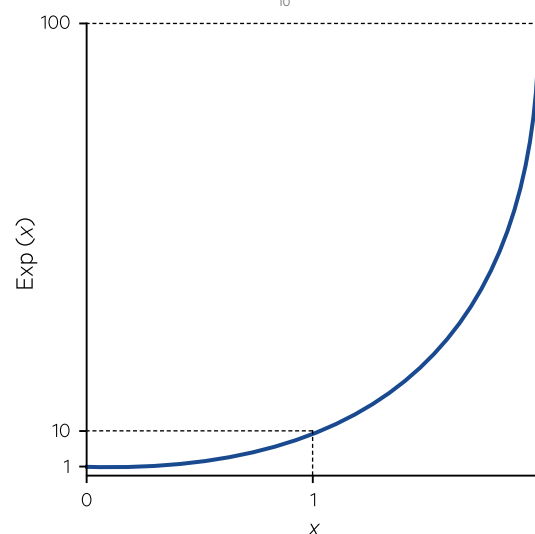
Since the exponential is the inverse function of the logarithm it can be used to convert logarithmically-scaled data to a linear scale. Similarly, the log of an exponent produces linear scaling, that is:

$$x = \log\Big(\exp\big(x\big)\Big) = \exp\Big(\log\big(x\big)\Big)$$

1.2.4 Sampling dimensions

The concepts of data resolution, density and extent were introduced in Volume 1A—Section 1 and expanded in Volume 1B—Section 1. These operate in four 'dimensions' of EO data acquisition:

- spectral—image channel values, which typically relate to spectral characteristics of the observed target;
- spatial—target area covered by image pixels;
- radiometric—increments of observed radiance represented by image pixel values; and
- temporal—time period for which an image or set of images is relevant.

The relationships between these sampling dimensions and the resolution, density and extent of EO datasets are summarised in Table 1.1.

Table 1.1 Sampling dimensions in EO imagery

Dimension	Characteristic		
	Resolution	Density	Extent
Spectral	Width of each wavelength channel	Number of channels detected by sensor	Range of wavelengths covered by all channels
Spatial	Ground area imaged per optical pixel	Number of pixels and lines in image	Area covered by image
Radiometric	Smallest change in detected energy that would be represented as a different image brightness level	Number of gradations (grey levels) used to represent full range of radiances that could be detected by sensor	Actual range of radiances detected in each channel
Temporal	Time period over which each image is acquired	Frequency of successive image acquisitions	Total time period for which this imagery is available

As introduced in Volume 2A—Section 8, a digital EO image can be analysed in terms of statistics derived from each of these four dimensions:

- spectral—techniques that effectively disregard the spatial properties of image data so that an image channel is treated as a one-dimensional array of numbers. These techniques derive statistics that describe the image in terms of the range and variation of data values they contain (see Section 1.3);

- spatial—describe the data values of an image as they relate to its spatial patterns. These statistics can be used to examine the spatial patterns within the image data and quantify the effects of changes in image spatial resolution;

- radiometric—methods to analyse components of image (or ground calibration) values that are not due to surface reflectance, such as atmospheric components, BRDF, terrain-related variations in illumination and scanner inconsistencies; and

- temporal—ways to standardise multi-date imagery to enable meaningful comparison and analysis of multiple images.

Most image transformations rely on spectral statistics but some also use spatial statistics (see Section 6). Radiometric statistics are most commonly derived during image pre-processing to correct for inconsistencies and artefacts of the imaging process (see Volume 2A—Section 3, and Sections 7.2 and 10.2.2 below). Temporal statistics are described in Volume 2D.

1.3 Spectral Statistics

Image spectral statistics summarise various aspects of the data values in each image channel and their correlation between channels (see Volume 2A—Sections 4 and 8 for details). Spectral statistics can be derived for individual channels (see Section 1.3.1) or selected pairs of channels (see Section 1.3.2). Those spectral statistics that are most commonly encountered when transforming EO imagery are summarised below.

1.3.1 Single channel statistics

For each individual image channel, the most commonly encountered spectral statistics are:

1.3.1.1 Measures of data extent

minimum: lowest value;
minimum: highest value; and
range: *minimum-maximum*.

1.3.1.2 Measures of central tendency

mode: most common value;
median: midpoint of range; and
mean: average of all pixel values.

1.3.1.3 Measures of dispersion from the mean

variance: summarises the differences between all pixel values and the mean value:

$$\sigma^2 = \frac{\sum\left(x - \mu\right)^2}{n}$$

where

x is pixel value in channel;
n is total number of pixels in channel (excluding null pixels); and
μ is true mean of the channel data.

and

standard deviation (σ): positive square root of the variance.

1.3.2 Channel histogram

A channel histogram is a statistical graph that summarises the range and distribution of values in an image channel, either in terms of the actual frequency of each pixel value (see Figure 1.6a) or the cumulative frequency of pixels with values equal to or less than a particular value (see Figure 1.6b). As transformations are applied to images to modify their values, these changes are reflected in the image histogram (Castleman, 1998; Gonzalez and Woods, 2018).

Figure 1.6 Image channel histograms

a. The frequency histogram is formed by plotting the number of pixels at each image value in a channel.
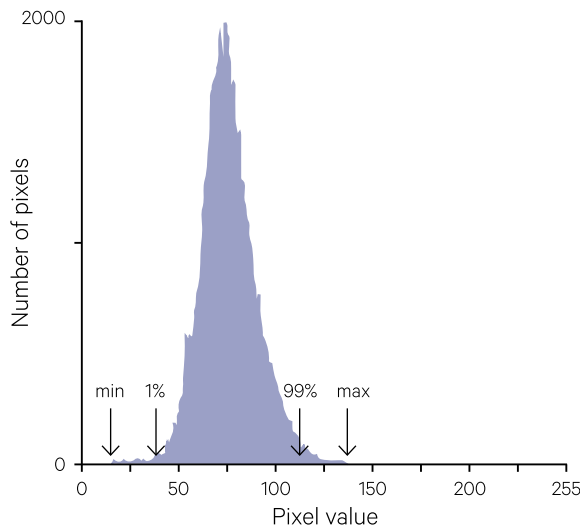
b. The (stylised) cumulative histogram for this image channel is shown in Figure 1.6a. In this example, 1% of pixels have values below or equal to 44, 20% have values below or equal to 66, while 95% of pixels have values below or equal to 89.
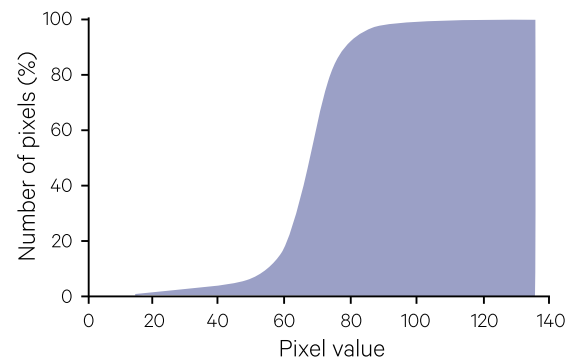
Source: Harrison and Jupp (1990) Figures 4 and 5

1.3.3 Multi-channel statistics

The degree of relationship beween the values in two image channels i and j is represented mathematically by the statistic:

covariance (v_{ij}):

$$v_{ij} = \frac{\sum_{k=1}^{n} \left(x_{ik} - \overline{x}_i \right) \times \left(x_{jk} - \overline{x}_j \right)}{n}$$

where

n is the total number of pixels in the image;
x_{ik} is value for pixel k in channel i ;
$\overline{x}_i$ is the mean of pixel values in channel i ;
x_{jk} is value for pixel k in channel j ; and
$\overline{x}_j$ is the mean of pixel values in channel j.

or its normalised form:

correlation (c_{ij}):

$$c_{ij} = \frac{v_{ij}}{\sqrt{v_{ii} \times v_{jj}}}$$

where

c_{ij} is the correlation between channels i and j;
v_{ij} is the covariance between channels i and j;
v_{ii} is the variance of channel i; and
v_{jj} is the variance of channel j.

Covariance or correlation values are typically summarised in a table for all channels in the image. This table is called a covariance or correlation matrix. The standard format of a covariance matrix is shown in Figure 1.7. Such matrices are used in a number of linear transformations as detailed in Sections 7, 8 and 9.

Figure 1.7 Format of variance/covariance matrix
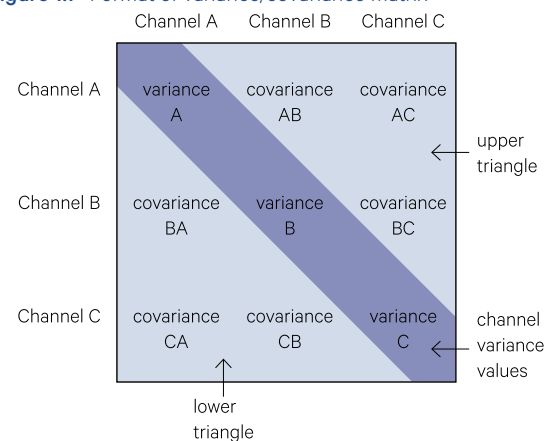
1.3.4 Channel crossplots

The relationship between values in two image channels can be shown in a crossplot. For example, using the example image channels in Figure 1.8, crossplots can be constructed for each pair of channels as shown in Figure 1.9. On each crossplot, the process is shown for the values of pixel 1, line 5 in the image. The distribution of the points on the crossplots visualises the similarity, or correlation, between the two plotted channels and can be used to plot a regression line (see Figure 1.9b).

Figure 1.8 Example image channels

Channel A

1	2	3	3	4
3	3	5	6	3
2	3	3	6	5
5	4	2	2	1
7	6	3	2	1

pixel 1, line 5 ⟶

Channel B

3	6	7	8	7
2	4	7	5	6
1	1	3	4	5
1	1	2	5	5
2	3	4	6	7

pixel 1, line 5 ⟶

Channel C

1	3	4	3	3
2	3	4	5	4
2	2	3	6	5
4	4	2	1	2
6	5	2	1	1

pixel 1, line 5 ⟶
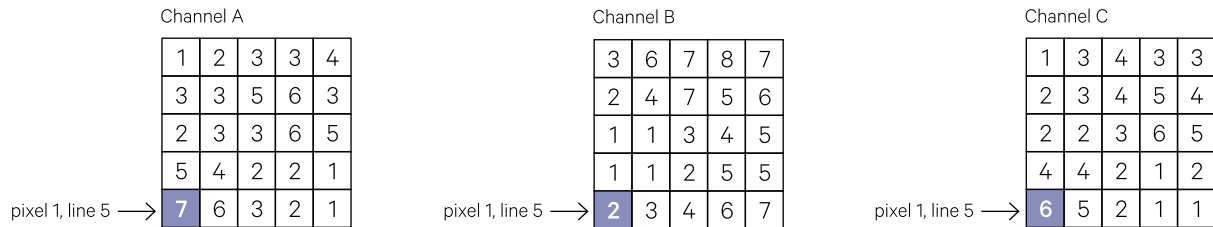
Source: Harrison and Jupp (1990) Figures 41

Figure 1.9 Example crossplots

The image channels in Figure 1.8 are crossplotted to show the extent of correlation between channels.

a. Channel B versus Channel A

This point indicates the crossplot location of pixel 1, line 5 in the image, which has a value of 7 in channel A and 2 in channel B

b. Channel C versus Channel A

Pixel 1, line 5 with value 7 in channel A and value 6 in channel B
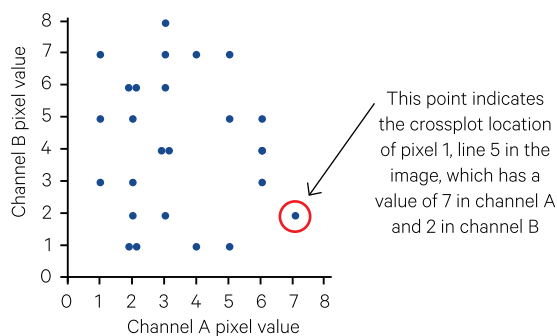
Regression line: channel C = 0.5 + 1.25 x channel A

c. Channel C versus Channel B

Pixel 1, line 5 with value 2 in channel B and value 6 in channel C
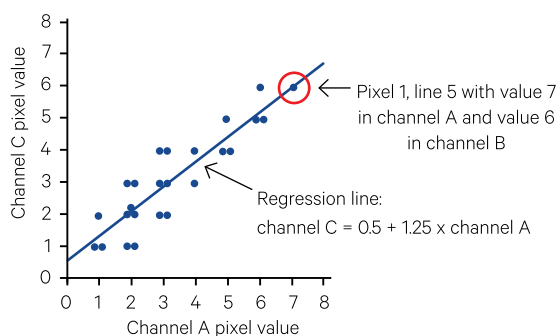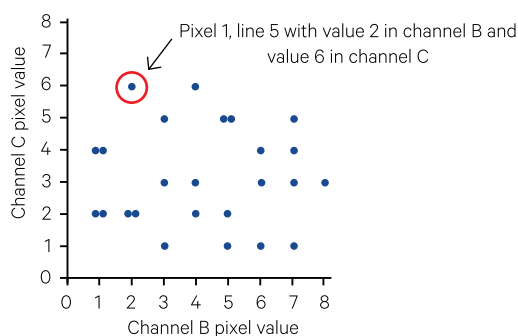
Source: Harrison and Jupp (1990) Figures 42

1.4 Processing Operations

Image processing methods are often classified as either linear or non-linear operations. Linear operations have the important property of additivity (that is, the result of a linear operation on two separate inputs is the same as that operation applied to the sum of those two inputs) whereas non-linear operations do not have this property (Gonzalez and Woods, 2018). The distinction between linear and non-linear processes is particularly relevant to the sequencing of operations in a multi-stage processing methodology (see Volume 2D).

Linear operations involve addition and subtraction. In the context of image processing, this applies to adding and subtracting pixel values from different channels, often using fractions or multiples of different channel values. In mathematical terms, the simplest way to represent linear combinations of input values is to use matrices (see Section 7 and Volume 2X—Appendix 6).

As the name implies, non-linear operations are then defined as those that are not linear. These operations include multiplication, division and the selection of order-dependent statistics such as the minimum, maximum, median, or mode.

Transformations may be applied to either individual channels or multiple channels, based on the input values in one or more channels. Single channel transformations tend to involve either:

- rescaling of the spectral value range according to a predefined relationship as introduced in Volume 2A—Section 4 and further detailed in Section 2 below; or
- modifying each pixel value by considering the values of its neighbouring pixels, usually involving the use of filters. The mathematical foundations of image filtering is introduced in Section 3 below and detailed in terms of specific uses in Sections 4, 5 and 6.

Multiple channel transformations can be broadly grouped into two categories:

- combining input channels using addition and subtraction (linear operations) as introduced in Section 7 and further discussed in terms of applications in Sections 8 and 9; or
- dividing one image channel by another (non-linear operations; see Sections 10 and 11).

While some image processing systems do not differentiate between linear and non-linear operations (and simply refer to them by a generic title such as 'band maths'), most transformations are based on some combination of the processes listed above. Processes that analyse larger datasets formed from multiple EO images are specifically considered in Volume 2D.

1.5 Further Information

Image Processing

SEOS: eLearning Tutorials: *https://www.seos-project.eu*

LEOWorks—Image Processing Tutorials: *http://leoworks.terrasigna.com/tutorials*

Image Processing Systems

ERDAS Imagine: *http://www.hexagongeospatial.com*

ERMapper: *http://www.hexagongeospatial.com*

ENVI: *http://www.harrisgeospatial.com*

TNTmips: *http://www.microimages.com*

1.6 References

Anderberg, M.A. (1973). *Cluster Analysis for Applications.* Academic Press, New York.

Castleman, K.R. (1998). *Digital Image Processing.* 2nd edn. Prentice-Hall, Inc. 667 pp.

Emelyanova, I.V., McVicar, T.R., Van Niel, T.G., Li, L.T., and van Dijk, A.I.J.M (2012) *On blending Landsat-MODIS surface reflectances in two landscapes with contrasting spectral, spatial and temporal dynamics.* WIRADA Project 3.4: Technical Report. CSIRO Water for a Healthy Country Flagship, Australia. 72pp.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing.* Pearson Educational Inc., New York.

2 Rescaling

A wide range of rescaling operations may be applied to image data for a variety of purposes. Some commonly encountered rescaling options that are available for displaying EO imagery are described in Volume 2A— Section 4. Rescaled image values can also be written to an image file to:

- match data values between different channels in an image;
- equalise values in imagery from different sources and thus allow them to be subsequently processed using similar parameters; or
- model the relationship between ground measurements and EO radiance (see Volume 2D).

All rescaling methods change the shape of the image histogram (see Section 1.3.2). Excursus 2.1 compares the impacts of some common rescaling methods when applied to an EO image channel and its histogram. This section describes the following single channel rescaling operations in terms of process and application:

- linear (see Section 2.1);
- non-linear (see Section 2.2);
- lookup tables (see Section 2.3);
- destriping (see Section 2.4); and
- histogram manipulation (see Section 2.5).

Excursus 2.1—Rescaling results

Source: Tony Sparks, Icon Water

This example uses an image acquired by Landsat-5 TM on 20 January 2010 over southeastern NSW. The town of Cooma is near the centre of this image, with Canberra in the northwest and the coastline visible in the east. Band 1 (green; see Figure 1.2) is used to differentiate spectral features on mountainous terrain, including forest cover, cleared pasture, water bodies, and cloud.

The selected image channel is shown below with six different stretches:

- none—shows the original data range for the example image channel. This presentation does not use all available grey levels for display and results in low contrast over most of the image (see Figure 2.1a);
- linear—spreads the full data range evenly over the full extent of available grey levels for display (see Figure 2.1b);

- logarithmic—this non-linear stretch redistributes the channel data range as a logarithmic function, which increases contrast in low data values (see Figure 2.1c);
- normalised—matches the image histogram to a normalised distribution (see Figure 2.1d);
- histogram equalised—this popular stretch aims at equalising the number of image pixels associated with each data value, and thus increases overall contrast (see Figure 2.1e); and
- user-defined—this non-statistical stretch is specified by a lookup table to achieve the desired spread of grey levels in the image (see Figure 2.1f).

Each stretch presents a different balance of light and dark tones in the image channel by changing the shape of the image histogram. Selection of the most appropriate stretch will depend on the desired end result.

Background image: This Landsat-5 TM image was acquired on 6 October 2011 over the arid landscape of Munga-Thirri National Park in western Queensland. The image is displayed using bands 6, 4, 1 as RGB, contrasting active fires (red) and fire scars (dark red) with temporarily-filled waterways and parallel sand dunes. Note that the red linear streaks, oriented roughly east-west near the large active fire body, are image artefacts resulting from sensor oversaturation.
Source: Craig Shephard, DSITI

Figure 2.1 Examples of rescaling

The example image channel is Band 1 (blue) from a Landsat-5 TM image acquired on 20 January 2010 over southeastern NSW, and features mountainous terrain with forest cover, cleared pasture, water bodies and cloud.

a. Unstretched image
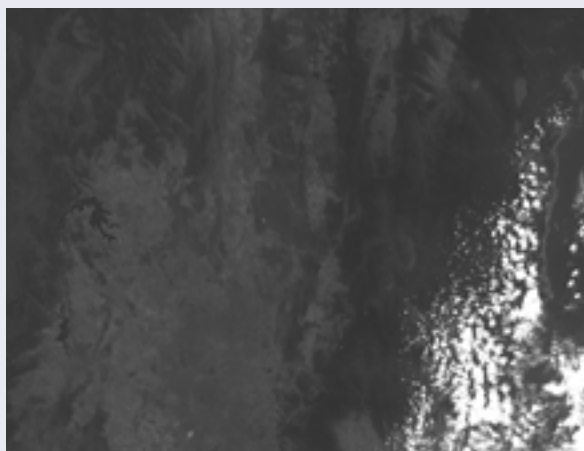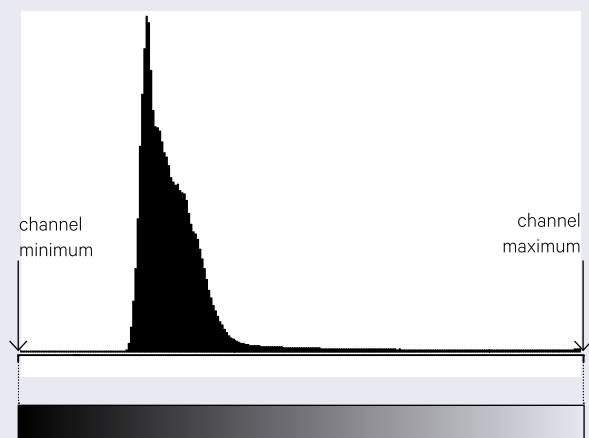
b. Linear stretch
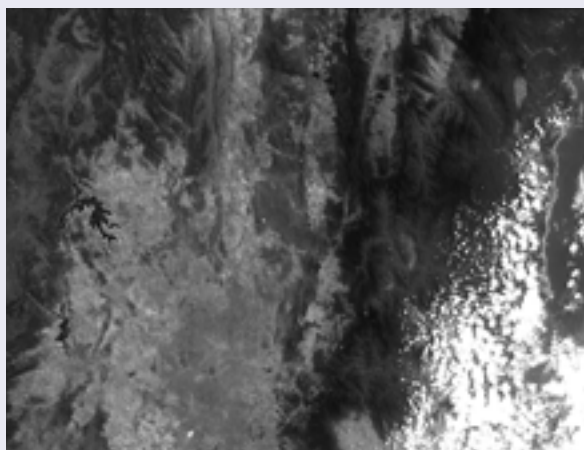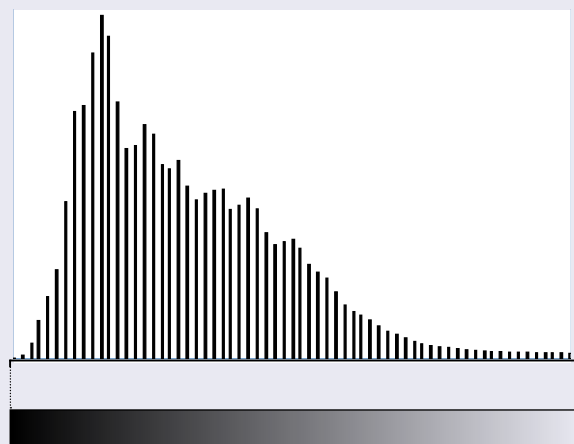
c. Logarithmic stretch
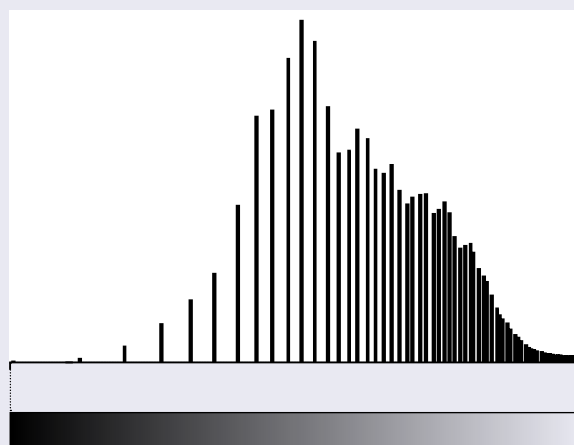
d. Normalised stretch
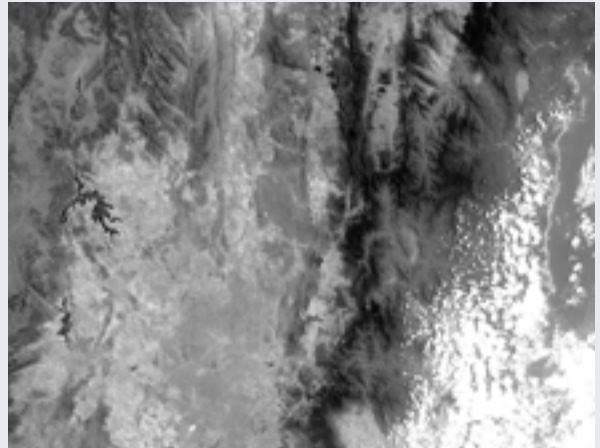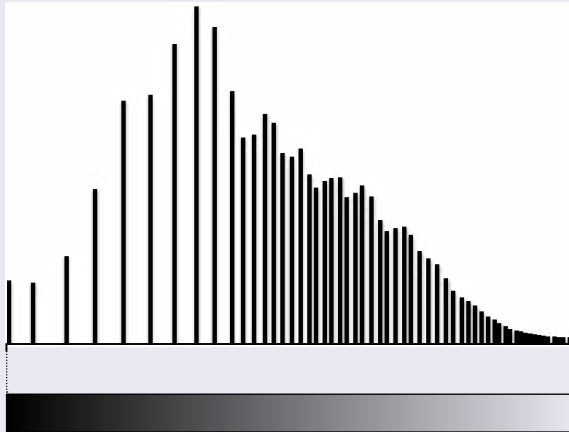
e. Histogram equalisation
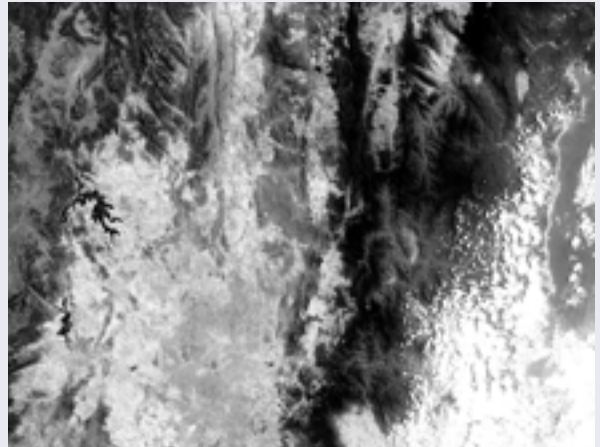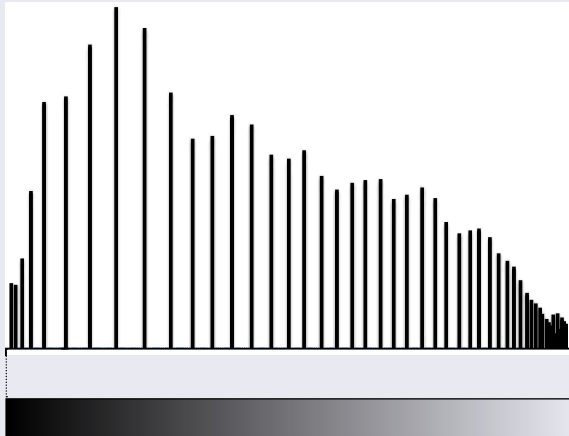
f. User-defined stretch based on lookup table (LUT)
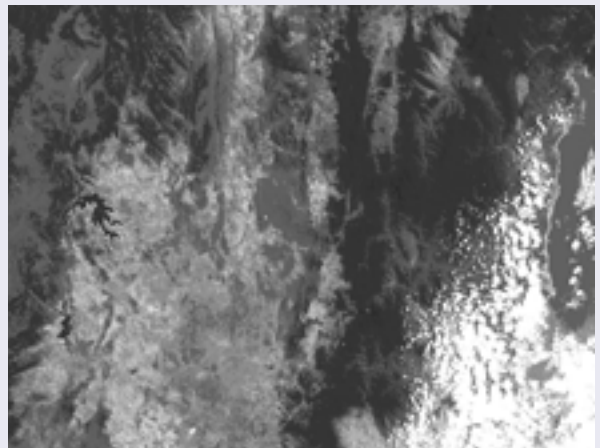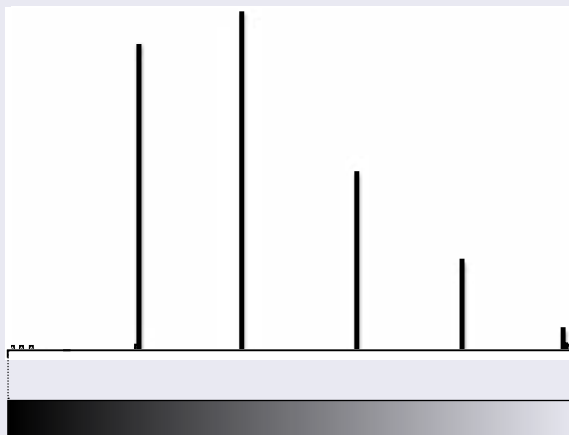
2.1 Linear Rescaling

Linear rescaling evenly matches the image data range to an output range, such as the available contrast in an image display device, while non-linear rescaling will selectively increase separation between low or high values in the image (see Volume 2A—Section 4.2). Linear rescaling is also useful for matching the contrast between images from different sources. Balancing of the data values between multi-source imagery, such as multiple images of a given ground location that were acquired on different overpass dates or by different sensors, is required before mosaicking or differencing such image data (see Volume 2D). Having the same minimum and maximum values in all channels of a single image is convenient for some operations, but such rescaling should be used with caution.

As discussed in Volume 2A—Sections 4 and 8, rescaling can produce discontinuities in the image data range (see Figure 2.1b). Given the discrete nature of image values, this discontinuity pattern can be exaggerated by rounding 'errors' during the rescaling process, which lead to a false representation of the differences between values in the output image. For example, in the rescaling illustrated in Figure 2.2, the input values 0, 1 and 2 become output values 0, 2 and 3 so that in the rescaled image the difference between the original values 0 and 1 is exaggerated relative to the original values 1 and 2. These potential problems are worth considering when selecting rescaling ranges, although in most cases the large potential range of data values in EO images means these problems do not seriously degrade image data. Similar discontinuities can occur when image data is converted to surface reflectance measurements.

The rescaling range is generally defined in terms of the minimum and maximum values in the input image, which are to be respectively mapped to the minimum and maximum values in the output image data range. When the selected rescaling range is within the absolute minimum and maximum values of the image, any differences in values greater than the rescaling maximum or less than the rescaling minimum are effectively lost in the output image and cannot be derived in any subsequent processing (unless the output data type is changed to avoid this problem).

Inversion of the data range in an image channel (that is reversing the relative order of pixel values to increase rather than decrease between the channel maximum and minimum values) can be implemented as a linear operation as described in Section 7.

Figure 2.2 Effect of rescaling operation on image data range

Original image values 0–5 are rescaled to the range 0–8. The rescaling falsely represents differences between values in the original image and the values 1, 4 and 7 do not occur in the rescaled image.

Input Value	Output Value
0	0
1	2
2	3
3	5
4	6
5	8
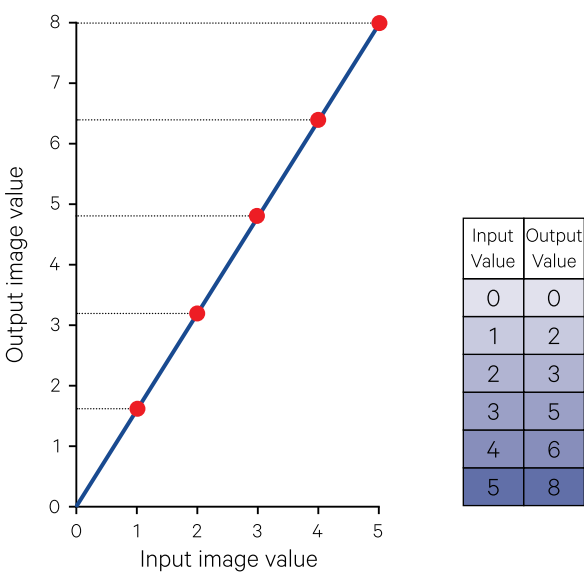
2.2 Non-linear Rescaling

Simple non-linear rescaling of image data values using logarithmic and exponential functions are introduced in Volume 2A—Section 4 and Section 1.2.3 above. These functions are often used to model the relationship between image radiance values and quantifiable ground parameters such as water depth (see Volume 3B). Transformations based on these functions then allow image values to be converted to or from a scaling which can be linearly related to ground measurements.

Some properties of logarithms make this transformation particularly useful as a pre-processing step for other transformations. For example, the product of two values in a linear data scale may be computed as the sum of their logarithmic values, that is:

$$x \times y = \exp\left(\log\left(x\right) + \log\left(y\right)\right)$$

Thus, filtering of log-transformed images effectively multiplies pixel neighbourhood values and can produce some useful enhancement effects (see Section 3.6). Similarly, the quotient of two linear values can be determined from the difference of their logarithmic values, that is:

$$\frac{x}{y} = \exp\left(\log\left(x\right) - \log\left(y\right)\right)$$

Another calculation that can be simply implemented in a logarithmic scale is exponentiation, since:

$$x^n = \exp\left(n \times \log\left(x\right)\right)$$

These relationships allow computation of complicated, non-linear combinations of channels such as:

$$\frac{ch1 \times ch2}{ch3^3 \times ch4}$$

If the logarithms of the four channels are computed, this combination can be produced from the expression:

$$\exp\left(\left(\log\left(ch1\right) + \log\left(ch2\right)\right) - \left(3 \times \log\left(ch3\right) + \log\left(ch4\right)\right)\right)$$

This computation can be easily implemented in an affine transformation as detailed in Section 7.1.

Extensive datasets are often plotted on a logarithmic scale. This scaling is effective when the full data range is large but only has significant variation near one end of the range. Logarithmic scaling does, however, affect the apparent relationship between the variables being plotted. For example, if a logarithmic function (see Figure 2.3a) is plotted using a logarithmic scale for the x-axis, the graph becomes linear as shown in Figure 2.3b.

Figure 2.3 Logarithmic scaling

a. The function $y = \log_{10}(x)$ is plotted using a linear scale.
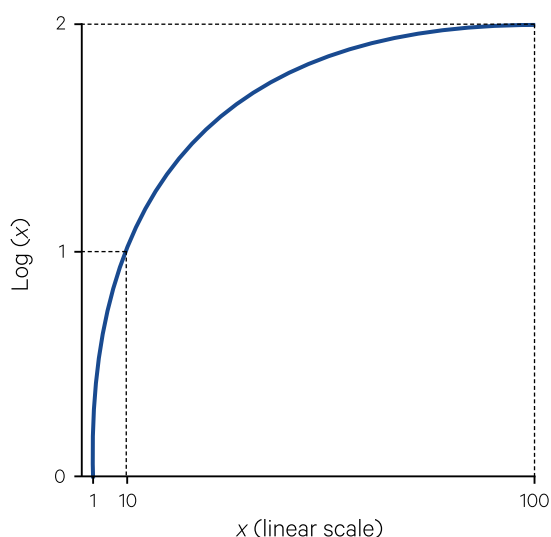
b. The function $y = \log_{10}(x)$ is plotted here using a logarithmic scale for the x axis.
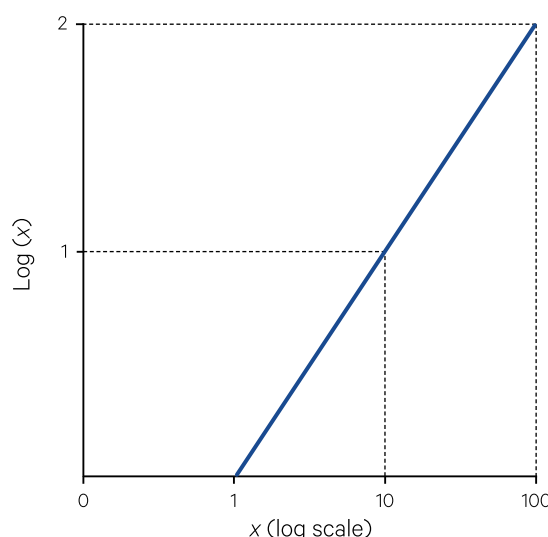
Source: Harrison and Jupp (1990) Figures 65a and 66

The logarithmic transformation has particular relevance to models relating EO image radiance to measurable ground features such as crop yield (see Volume 3A). In image processing systems, logarithmic and exponential transformations can be implemented by computing the functions:

$$\log\left(x + a\right)$$

and

$$\exp\left(\frac{x}{a}\right)$$

where

x is an image pixel value; and
a is the scaling factor.

As indicated in the graph of Figure 2.3a, the logarithm function for value 0 is undefined (said to be negative infinity). The scaling factor '*a*' allows an offset to be added to, or subtracted from, the image values during processing, that is:

$$a = \frac{\beta}{\alpha}$$

where

$$\log\left(\alpha \times x + \beta\right) = \log\left(\alpha\right) + \log\left(x + \frac{\beta}{\alpha}\right)$$

Since the $\log\left(\alpha\right)$ term is a constant offset, it can be ignored when rescaling image values. However, the term would become important for model development if image values were to be related to physical parameter measurements (see Volumes 2D and 3). When log-ratio images are being computed, dark values for the ratio can be specified as negative scaling factors during the logarithmic transformation (see Section 10.1.2).

Similarly a suitable divisor, *a*, for the exponential function avoids extremely large output values which would be clipped during the rescaling process:

$$\exp\left(\alpha \times x + \beta\right) = \exp\left(\alpha \times x\right) \times \exp\left(\beta\right)$$

In this case, the offset can again be ignored in image rescaling and:

$$a = \frac{1}{\alpha}$$

In addition to the scaling factor, a data range is specified for each image channel during rescaling so that all other values are clipped to this range and values within the range are adjusted to increment from a minimum value of zero. This sub-range has relevance when the log channels are to be used for other transformations, such as to form channel ratios (see Section 10.1.3). Generally the minimum range value should be one less than the absolute minimum of a channel. This value is scaled to the minimum value in the image data range of the output channel. Similarly, the maximum value specified in the processing range is scaled to maximum image data range value in the transformed channel.

The actual transformed minimum and maximum values are often reported during processing and are useful for subsequent rescaling operations and model fitting. The resulting image channels contain data ranges that have been redistributed using logarithmic and/or exponential functions as defined. However, in byte images, the pixel values are usually integer representations of these function values, which are subsequently rescaled to fill the maximum image data range. The final rescaling operation needs to be taken into account when using image pixels values in log (radiance) models or when computing the inverse transformation.

Nothing happens until the pain of remaining the same outweighs the pain of change.
(Arthur Burt)

2.3 Lookup Tables (LUT)

Image data values may be redefined using a Lookup Table (LUT). The use of LUT for image display enhancements is discussed in Volume 2A—Sections 4 and 8. The use of LUT for density slicing, or for aggregating basic feature classes into labelled categories in the mosaic model for image classification, is also discussed in Volume 2E.

In the context of image rescaling, an LUT performs a similar function by defining an output value for each possible input value in the image. An LUT is simply a single column table—the rows of the table represent input data values while the entries in the table indicate the output values to assign to each input value. The definitions may represent a wide range of contrast enhancements such as piecewise linear or sawtooth ('overlapping contrast enhancement') stretches as illustrated in Figure 2.4.

Other more sophisticated rescaling transformations, such as histogram equalisation (see Section 2.5), may be represented as an LUT (although this is more commonly done as an 'on the fly' statistical transformation). Indeed, for rapid computation, many other image processing functions such as non-linear functions are implemented using LUT in image processing programs. LUT are also a convenient means for representing hierarchies and other many-to-one mappings (see Volume 2E).

Figure 2.4 LUT operations

a. Piecewise linear rescaling

LUT entry	Output Value
0	0
1	1
2	2
3	2
4	4
5	5

b. Sawtooth stretch

LUT entry	Output Value
0	0
1	4
2	0
3	4
4	0
5	4
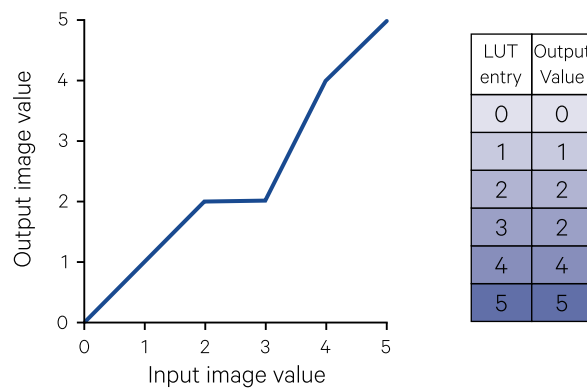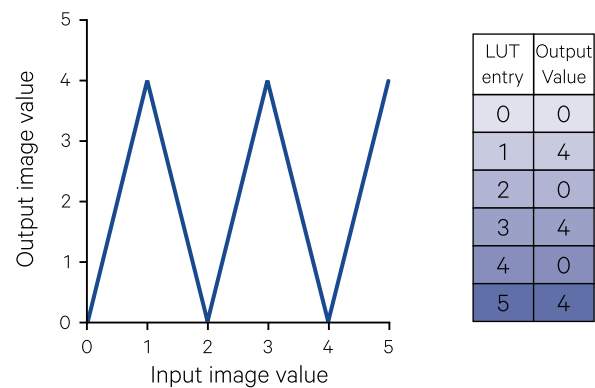
Source: Harrison and Jupp (1990) Figure 10

2.4 Destriping

EO imagery may contain spatial striping patterns due to the miscalibration between multiple detectors in the scanning system (see Volumes 1 and 2A). In scanners that record adjacent image lines using different detectors, a line striping is frequently observed in which the periodicity of the stripes matches the number of detectors in the instrument. Landsat TM, which senses 16 lines per scan, records during both forward and back scans across the swath so its imagery can have 16 or 32 line striping. When such data are geometrically corrected, the striping pattern can be observed to have 17 or 34 line periodicity. Additionally, in linear array or 'pushbroom' scanners, which operate banks of detectors across an image line, different pixel values in the line will be determined by different detectors. Imagery produced this way can show vertical striping, as was visible in some early SPOT data.

In an attempt to correct for miscalibration effects, various pre-processing operations are performed by ground receiving stations before image data are distributed. However, such variations are data-dependent so are difficult to remove effectively using whole scene statistics or instrument calibration readings. Consequently, striping patterns commonly occur in image data from EO sources and usually need to be reduced in some way before applying other image processing operations for enhancement, interpretation or classification. It must be emphasised that destriping algorithms can only be applied to imagery that has not been resampled since the resampling process generally rotates image lines, which invariably rotates the striping artefact (see Figure 2.5).

Figure 2.5 Rotated image striping pattern

This Landsat-5 TM image demonstrates a visible striping pattern in the water. As detailed in Volume 2A—Section 3, this striping results from miscalibration of the sensor's 16 detectors, which are designed to image a 16-line swath. After image resampling, however, the striping pattern has been rotated so would be difficult to remove. This image was acquired on 31 July 2013 over the northern fringe of Townsville and is displayed using bands 5, 4, 2 as RGB.

Source: Tony Sparks, Icon Water

Striping patterns are most evident in areas with relatively uniform image values, such as deep water, with different land covers showing different striping effects. In the scanning device, individual detectors are calibrated relative to a standard lamp, which radiates a standard range of wavelengths (see Volume 2A—Section 3). For this calibration target, all detectors may record, or can be 'calibrated' to record, the same radiance; however differences in response functions between the detectors can still cause miscalibration for other wavelengths. Thus, the striping pattern varies with the land cover being imaged. In destriping then, these patterns are best identified if the image is segmented into broad land cover regions (such as water, vegetation, bare soil and cloud) before applying destriping methods (see Volume 2A—Section 10 regarding image segmentation methods).

A number of destriping algorithms have been developed to try to minimise the effects of detector imbalance. These algorithms basically accumulate statistics relating to each detector. For example, for an image with six-detector striping, one set of statistics are gathered separately for lines 1, 7, 13, 19, etc., then another for lines 2, 8, 14, 20, etc., and so on. The destriping process then attempts to remove the differences by producing imagery in which selected statistics match between all detectors. A simple example of destriping based on this approach is shown in Figure 2.6. While some banding is still visible in the destriped image, partly due to the limited dynamic range of the original data (8 bit), the processed image has significantly reduced the striping pattern.

Destriping algorithms typically accumulate mean and variance statistics for a user-defined periodicity in each channel in a (stack of) image(s). Divergence and centrality values are generally reported, which indicate the extent to which each detector is similar to the other detectors in each channel (with low values indicating a close match between detectors). This algorithm is applied to the image(s) by matching all detectors to the mean detector values or to a selected reference detector. The detector reporting the smallest sum of centralities for each channel is recommended as the reference detector. This process should be applied to imagery before resampling as the original striping pattern is likely to be modified to some form of diagonal artefact in the geo-corrected image, which would be more difficult to correct (see Figure 2.5).

Figure 2.6 Destriping example

In this example, an ocean portion of band 4 (NIR) in an unresampled Landsat-5 TM image has been extracted and destriped (the resampled version of the full image is shown in Figure 2.5). While some banding is still visible in the destriped image, partly due to the limited dynamic range of the original data (8 bit), the processed image has significantly reduced the striping pattern.

a. Original image band
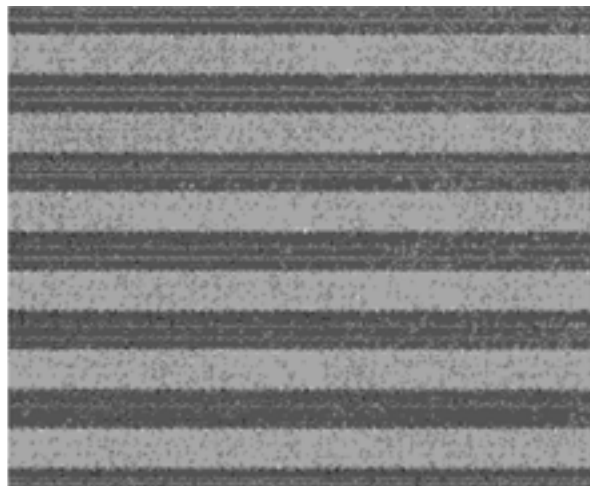
b. Destriped image band
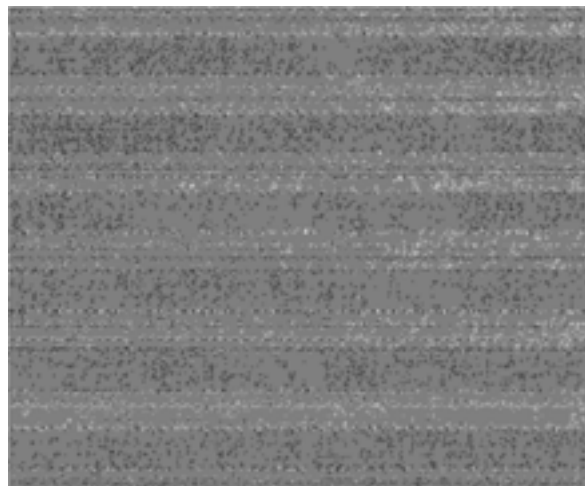
Source: Tony Sparks, Icon Water

These detector statistics are used to compute a correction equation for each detector and channel combination. This equation can be expressed as:

$$x_{ij} + a_{ij} + b_{ij} \times x_{ij}$$

where

x_{ij} is pixel i in channel j; and
a_{ij} and b_{ij} are adjustment factors computed from the detector statistics.

Some destriping algorithms allow a theme (a range of channel values; see Volume 2A—Section 9.1.1) to be defined for the destriping operation, so that the process is only applied to pixels that satisfy the theme. This option effectively allows the image to be spectrally segmented into different land cover categories during destriping. It is highly recommended that appropriate theme values be used in conjunction with image destriping to account for data-dependent patterns.

The effects of destriping may be difficult to discern visually in some imagery but can be quantified by computing the difference between the original and destriped images (see Volume 2D). In some cases, a narrow data range may make the destriping operation very difficult since the striping effectively occurs about a single grey level value. Linear rescaling of the data channel before destriping may improve the performance of the destriping algorithm.

If a significant striping pattern cannot be removed adequately by using a destriping algorithm, image filtering may be used to reduce the visual impact of the pattern. This process is described in Section 4. In this context, average filtering using filter weights for a gentle smoothing filter, with a filter size that is at least as wide as the detector striping, would be most appropriate.

Very noisy imagery should be despiked before destriping (see Section 4.5). For example, isolated noise spikes (pixels with unrealistically high values) are commonly observed in imagery of deep water due to sensor saturation from sunglint off the water surface. The despiking process replaces these values with more reasonable ones and thus avoids biasing any statistics generated for the destriping algorithm.

2.5 Histogram Manipulation

As introduced in Section 1.3.2 and illustrated in Excursus 2.1, the 'appearance' of an image determines the shape and distribution of its histogram. Accordingly, transformations that change image values almost invariably change its histogram. Just as manipulating image values changes the image histogram, manipulating the image histogram changes the image appearance (Gondalez and Woods, 2018).

Commonly encountered approaches to histogram manipulation were introduced in Volume 2A:

- histogram normalisation (see Volume 2A—Section 4.2.3.1 and Figure 2.1d); and
- histogram equalisation (see Volume 2A—Section 4.2.3.2 and Figure 2.1e).

Some applications, however, require image histograms to be matched to a tailored shape (Gonzalez and Woods, 2018). Such transformations may be appropriate to compensate for 'over exposure' of image noise resulting from histogram equalisation, typically in low image values. In practice, they are often computed via the histogram equalisation process.

Most histogram manipulation methods effectively redistribute the histogram 'bins' along the intensity axis (see Figure 2.1). Accordingly, these methods only ever approximate the shape of the target histogram. In order to precisely match a specified histogram shape, the original image values need to be modified and redistributed. This process may be relevant to EO imagery to establish a reference image for sensor calibration. One approach to exact histogram matching is based on three stages (Coltuc *et al.*, 2006):

- order pixels in image by some predefined criteria;
- divide pixels into a set number of groups so that all groups are the same size; and
- allocate an appropriate intensity value to each group.

This method can still falter on images with large homogeneous features, in which case additional processing with image masks may be appropriate to enable detail in particular portions of the image to be enhanced. Further details on histogram matching methods are provided by Gonzalez and Woods (2018).

2.6 Further Information

Jensen (2016) Section 8

Gonzalez and Woods (2018)

2.7 References

Coltuc, D., Bolon, P., and Chassery, J.-M. (2006). Exact Histogram Specification. *IEEE Transactions on Image Processing*, 15(5), 1143–1152.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Single Channel Filtering

Filtering operations are generally applied to individual image channels to modify pixel values based on the values of each pixel neighbourhood. The following sections consider spatial filtering in terms of:

- filter operation, including weights, sizes, thresholds and variations (see Section 3);

- low pass or smoothing filters to blend pixel values (see Section 4);

- differential and high pass filters to detect and enhance edges respectively (see Section 5); and

- filters that highlight variations in a surface, including texture, relief shading, directional, and curvature (see Section 6).

Contents

3 Filtering Operations

Image data can be represented in terms of the spatial arrangement of pixels (the spatial domain) or in terms of the frequencies of their spatial patterns (the frequency domain). The frequency domain represents image spatial patterns as sinusoidal waves of differing frequency, amplitude and direction. In EO imagery, large spatial patterns (such as regional landscape patches) correspond to low frequency waves in the frequency domain while noise or feature edges are represented as high frequencies.

When image data are represented in the frequency domain, specific frequencies can be modified using frequency filtering techniques such as Fourier analysis. Fuller (1966), Castleman (1998) and Gonzalez and Woods (2018) discuss methods for designing filters to suppress, enhance or isolate specific ranges of frequencies. Discrete Fourier analysis allows image data to be transformed from the spatial to the frequency domain where they can be filtered in terms of specific frequencies then transformed back to the spatial domain using the inverse Fourier transform function. Fourier analysis is appropriate where data patterns are (semi-)periodic, such as in signal processing applications, but is less effective with data containing local variations, such as EO imagery.

Filtering can also be applied directly to spatial data by 'convolution'—a mathematical operation in which the value of an image pixel is weighted by the values of the surrounding pixels (see Section 3.1). Richards (2013), Castleman (1998) and Gonzalez and Woods (2018) discuss the relationship between spatial filter convolution and frequency filtering. Filters defined in the frequency domain can be represented as spatial filters and applied more efficiently (but in some cases less precisely) using convolution in the spatial domain. Conversely, simple and direct spatial filters are difficult to implement in the frequency domain. Spatial filtering is more applicable to EO data that exhibit local rather than periodic variations.

Filtering for EO imagery can involve:

- variable filter weighting values (see Section 3.2);
- different filter sizes, generally ranging from 3×3 to 11×11 (see Section 3.3);
- thresholds (defining the minimum difference between the input and filtered values of each pixel) which need to be reached before the input value is modified (see Section 3.4);
- adaptive or edge-preserving filters (see Section 3.5); and
- homomorphic filtering (see Section 3.6).

Most image processing systems offer a wide range of filter-based transformations. These can range from generic filtering to special purpose filters for particular data types, such as the computation of Sun shadowing or curvature from elevation data. The uses of specific filter-based transformations are detailed in Sections 4 to 6 below.

In the Information Age,
the first step to sanity is filtering.
Filter the information: extract for knowledge.
(Marc Stiegler)

Background image: High resolution aerial image of apartment development in South Perth, WA, acquired on 11 January 2018. **Source:** © EagleView

3.1 Operation

Spatial filter transformations described in this Volume use those image values that are covered by a filter centred on a pixel to change the values of that pixel. The simplest form of a filter is a (square) grid of pixels, which usually has an odd number of pixels along each side to ensure symmetry about the central pixel. In this context, the filter is also commonly referred to as a moving window or box car filter, or a template or a kernel.

In a discrete dataset such as an image, the convolution process at an individual pixel (x,y) is defined as:

$$\sum_{i=-m}^{m}\sum_{j=-n}^{n} f\left(x+i, y+j\right) g\left(i,j\right)$$

where

    $2m$+1 is the number of pixels across the filter;
    $2n$+1 is the number of lines down the filter;
    f is the image;
    g is the filter;
    x is the pixel position in the image; and
    y is the line position in the image.

Weighting values may also be associated with each cell of the filter.

The operation of a 3×3 averaging filter is shown in Figure 3.1a. The filter is applied to each pixel in an image by multiplying it by the filter's central value and adding the products of the other filter elements and their corresponding neighbouring pixel values. Filters can operate by computing a (weighted) mean or average value as illustrated in Figure 3.1, or by determining the median or mode value of the pixels within the filter region (see Sections 4.2 and 4.3 below). In an average filter, a cell weight of zero effectively excludes the cell from the filter region (see Figure 3.1b).

Figure 3.1 Operation of an average filter

a. This operation is applied to every pixel in an image, that is, every pixel is treated as the central value in the filtering operation to determine its output value. To do this, the filter is systematically passed over the image and moved by one pixel after each operation. In this example, the filtered value for the central input pixel (e) will be: A = (a×m + b×n + c×o + d×p + e×q + f×r + g×s + h×t +i×u) / W
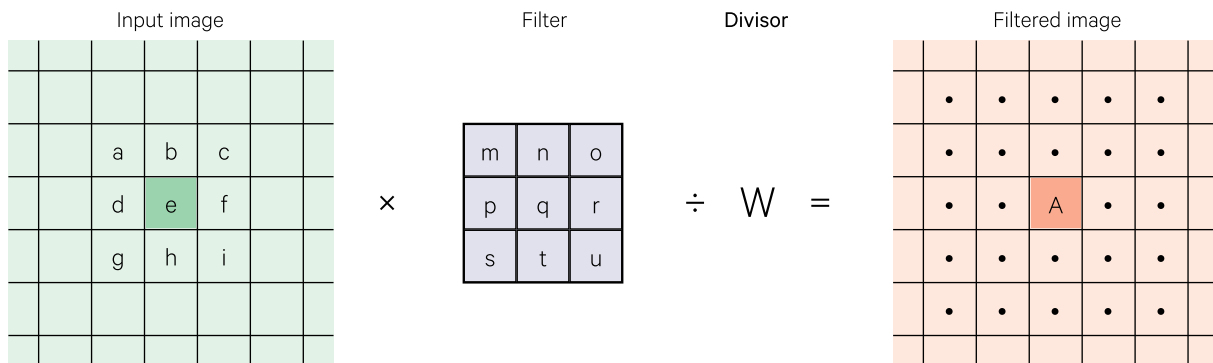
b. Numeric example. Note that edge pixels have not been filtered in this example.

Source: Harrison and Jupp (1990) Figure 67

26

Earth Observation: Data, Processing and Applications. Volume 2: Processing
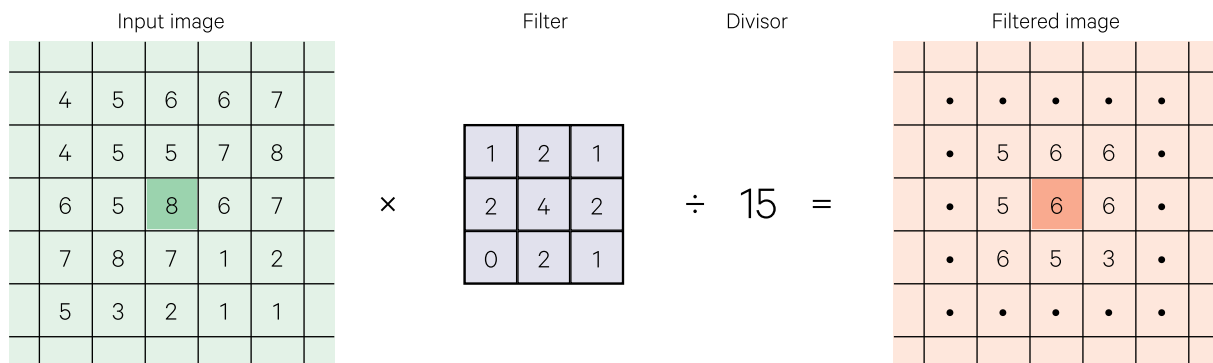
The effect of the spatial filtering process on an image largely depends on filter type (and the weighting values used in the case of an average filter), while the extent of this effect depends on filter size. Multiple applications of the same or different average filters can be computed as a larger single filter by convolution of the original filters (see Excursus 3.1).

One variation for an average filter is called the Olympic filter, which excludes the minimum and maximum values within the filter region before computing the average value. Alternatives to computing the mean, median or mode of pixel values within the filter region include replacing the central value with the minimum or maximum value (Jensen, 2016).

Excursus 3.1—Filter convolution

Two filters A and B applied sequentially to an image would have the same effect as a single application of the filter C where C is the convolution product of A and B. This convolution process is shown in Figure 3.2. The value of a cell in the combined filter is computed as the product of the weights of the overlapping cells between the two filters when one filter is centred in the new filter region and the other filter is moved around it. This process is useful for determining higher order derivative filters for edge detection or surface shape delineation (see Sections 5 or 6).

Figure 3.2 Convolution of two filters

a. To compute the weights of a 5×5 filter, C = A×B, we sum the products of overlapping cells between A and B, when one filter is centred in the 5×5 region and the other filter is moved around it. The result becomes the weight of the cell at the centre of the moving filter. In this example, for the filter weight labelled X in Filter C, the overlap between Filters A (centred) and B (moving) includes only one cell, the top left cell in Filter A (value a) and the lower right cell in Filter B (value r). These values are multiplied to create the filter value a×r for cell X in Filter C.

Filter A (3×3)

a	b	c
d	e	f
g	h	i

Filter B (3×3)

j	k	l
m	n	o
p	q	r

Filter C (5×5) is created by convolution of filters A and B
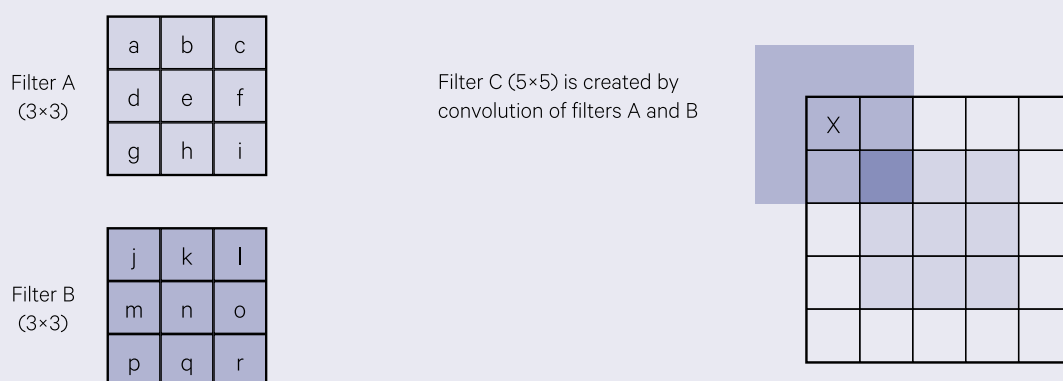
b. All resulting values for 5×5 Filter C

a×r	a×q + b×r	a×p + b×q + c×r	b×p + c×q	c×p
a×o + d×r	a×n + b×o + d×p + e×r	a×m + b×n + c×o + d×p + e×q + f×r	b×m + c×n + e×p + f×q	c×m + f×p
a×l + d×o + g×r	a×k + b×l + d×n + e×o + g×q + h×r	a×j + b×k + c×l + d×m + e×n + f×o + g×p + h×q + i×r	b×j + c×k + e×m + f×n + h×p + i×q	c×j + f×m + i×q
d×l + g×o	d×k + e×l g×n + h×o	d×j + e×k + f×l + g×m + h×n + i×o	e×j + f×k + h×n + i×n	f×i + i×m
g×l	g×k + h×l	g×p + h×q + i×r	h×k + i×k	i×j

Source: Harrison and Jupp (1990) Figure 68

Figure 3.3 Treatment of edge pixels in an image during filtering operation

The values of pixels adjacent to an image edge are mirrored or reflected about the edge to provide a filtering neighbourhood for edge pixels.

a. Input image values

b. Reflected edge values

c. A 3×3 filter centred on the pixel with value 'a' uses the image values highlighted below.

d. A 5×5 filter centred on the pixel with value 'h' uses the image values highlighted below.

Source: Harrison and Jupp (1990) Figure 70

Pixels at the edge of an image (that is, in the first or last row or column) are typically processed by 'reflecting' the pixel values of the adjacent rows or columns to 'fill' the filter being used. This process is illustrated in Figure 3.3. The edge effects are not significant with smaller filters, but should be considered with larger filters especially if image width is relatively small. When significant, they can result is anomalous pixel values around the edges of the image (see Figure 4.14).

Null values in an image channel are generally ignored by the filtering process. In such cases, the filter divisor is appropriately decremented to compensate for the missing value.

3.2 Filter Weights

Image filters can be defined to:

- reduce spatial variation in the image—such as striping patterns or noise spikes (see Section 4.5); or to
- enhance selected spatial patterns—such as linear features or edges (see Section 5).

The impact of a particular filter is determined by the weighting values used. The relationship between smoothing, edge-enhancing and edge-detecting filters is illustrated in Excursus 3.2.

Excursus 3.2—Example of Filter Operations

This example uses a scanned photograph to illustrate the differences between smoothing, edge enhancing, and edge detection filters. The original image is shown in Figure 3.4a. The equivalent filter operation matrix is shown above it, which could be considered a null filter since each pixel remains unchanged.

The impact of a strong smoothing filter applied to our example image is illustrated in Figure 3.4b. This filter replaces each pixel in the image with the average of its original value and the values of its eight neighbouring pixels. The resulting image is clearly a blurred version of the original in which edges are less distinct. Smoothing filters are discussed in more detail in Section 4.

Edge-enhancement filters, by contrast, emphasise the difference between the central pixel and its neighbours, which results in clearer transitions between different tones (see Figure 3.4c). This operation enhances minor variations in the original image, such as image artefacts, as well as actual edges.

To highlight transition zones, edge detection filters effectively subtract the smoothed image from the original as illustrated in Figure 3.4d. This operation is evident in both the filter design and the resulting image. Filters to detect and enhance edges are detailed in Section 5.

Figure 3.4 Relationship between smoothing, edge enhancement and edge detection filters

a. Original image =null filter

b. Smoothed image = original image – edges

c. Edge enhanced image = original image + edges

d. Edge detection image = original – smoothed image

Source: Harrison and Jupp (1990) Plate 1

When applied to geophysical imagery (for example, magnetics or gravity data) edge-enhancement (or high-pass) filtering can allow regional trends to be removed, which simplifies analysis of small area variations (see Section 5.2). Alternatively, local variations may be removed to explore deeper or larger geological structures using smoothing (or low-pass) filters (see Section 4). Such filters are commonly applied to frequency data but may be defined in the spatial domain using an inverse Fourier transform on the desired frequency response and shortening to a practical filter size as detailed in Fuller (1966).

Standard filtering techniques derive the filtered value of the central pixel using all (active) pixels in the filter neighbourhood. When the filter lies across the edge of a feature in the image however, smoothing operators tend to blur feature boundaries and de-emphasise small features (see Section 4). Alternative filtering approaches, such as adaptive filtering (see Section 3.5), minimise this effect.

3.3 Filter Size

Larger filters obviously involve more computations and therefore require longer processing time. While these differences may not be significant for small data volumes, they are a consideration with large datasets, even with ongoing improvements in processing speeds. Table 3.1 shows the relative processing times required for different filter sizes. As demonstrated in Excursus 3.1, larger filter sizes can be simulated by repeated filtering with smaller filters.

Table 3.1 Relative processing times for different filter sizes

Filter Size	Time Factor
3×3	1.0
5×5	1.7
7×7	2.7
9×9	4.2
11×11	6.0

3.4 Thresholding

Filtering can also be conducted in conjunction with thresholds such that the centre pixel is only replaced by a filtered value if the difference between the original and filtered values is greater than some specified magnitude. Thresholds are typically defined in terms of the image data range or as percentages of that range.

Percentage-based thresholds result in a greater filtering impact on pixels with low values and, given the skewed distribution that is typical for EO imagery, thresholds based on percentages are generally appropriate. In this case, the central pixel value is only modified if the percentage-based threshold, *x*, is greater than a user-defined value, where:

$$x = \frac{\text{original value} - \text{filtered value}}{\text{original value}} \times 100$$

Thresholding is useful for removal of noisy spikes in image data (see Section 4.5), without blurring the low frequency variations, or edge features with lower frequency than the noise. Image histogram information may be useful when selecting appropriate levels for filter thresholds (see Volume 2A—Section 8).

3.5 Adaptive Filtering

Adaptive filtering defines sub-windows within the filter region and uses some criterion, such as minimum variance, to select the 'best' sub-window to use to compute the filtered value for the central pixel. For example, four 2×2 sub-windows could be defined in a 3×3 smoothing filter as illustrated in Figure 3.5. In this case, sub-window 1 has minimum variance so the central value is replaced with the mean of that sub-window. This procedure tends to preserve, rather than blur, edges in smoothing filters and avoids emphasising edges in texture images (see Section 6.1).

Another implementation of adaptive filtering is known as the Wallis Filter (Wallis, 1976), which adjusts local image brightness to maintain or match specified local mean and standard deviation values within user-defined windows (Ballabeni *et al.*, 2015). This process results in more uniform tone, with improved local contrast in bright and dark regions (see Figure 3.6), so is particularly useful for highlighting local variations in imagery with uneven illumination. Adaptive filtering is further discussed in Section 4.4.

Figure 3.5 Adaptive filtering

A 3×3 filter can be divided into four sub-windows to allow selection of the filter region that is most compatible with the central pixel value. If the criterion of minimum variance were used to select a sub-window, the filtered value of the central pixel would be the mean of the top-left window, with value 4.0.

Sub-window 1
$\mu = 4.0$
$\sigma^2 = 0.5$

3	4	8
4	5	9
8	10	9

Sub-window 2
$\mu = 6.5$
$\sigma^2 = 4.25$

3	4	8
4	5	9
8	10	9

Sub-window 3
$\mu = 6.75$
$\sigma^2 = 5.69$

3	4	8
4	5	9
8	10	9

Sub-window 4
$\mu = 8.25$
$\sigma^2 = 3.69$

3	4	8
4	5	9
8	10	9
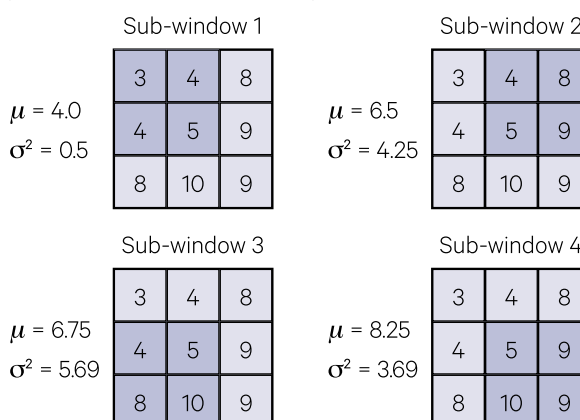
Source: Harrison and Jupp (1990) Figure 69

Figure 3.6 Wallis filter

This example, based on Pleiades pan-enhanced blue band, is centred over part of Fyshwick, ACT, and was acquired on 15 August 2016. The Wallis filter visibly enhances variations within features, especially the settlement ponds of the Fyshwick Sewerage Treatment Plant (located in central and lower portions of image) and the Canturf irrigated lawns (in top of image).

a. Pan-enhanced blue band

b. Result of Wallis filtering
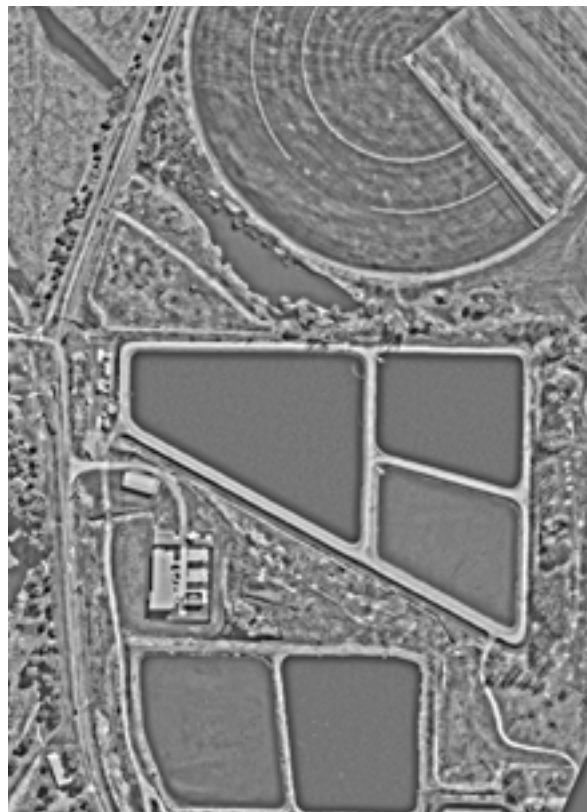
Source: Tony Sparks, Icon Water

3.6 Homomorphic Filtering

The application of filtering techniques to non-linearly scaled image data can produce some interesting results. 'Homomorphic' filtering generally describes the application of a linear, high-pass filtering operation to a logarithmically-scaled image, then exponentially-transforms the result back to a linear scale (Oppenheim and Schafer, 2004; Pitas and Venetsanopoulos, 1990). The additive operation of the filter on the logarithmically-scaled image has a multiplicative effect on the original image values. This form of filtering has traditionally been used to remove multiplicative noise from an image, especially when due to non-uniform illumination (Moik, 1980).

As an image enhancement tool, the transformation sequence reduces the overall image illumination component and enhances local contrast related to feature-specific reflectance. For example, this type of processing would enhance the subtle variations in image radiance values of shallow water features. With EO imagery, homomorphic filtering has also been used to remove cloud cover (Mitchell *et al*., 1977).

3.7 Further Information

Castleman (1998) Sections 9, 11 and 12

Gonzalez and Woods (2018) Section 3

Jensen (2016) Section 8

3.8 References

Ballabeni, A., Apollonio, F.I., Gaiani, M., and Remondino, F. (2015). Advances in Image Pre-Processing to Improve Automated 3D Reconstruction. Int. Arch. Photogram. Remote Sensing and Spatial Information Sciences, Volume XL-5/W4. *Proc. 3D Virtual Reconstruction and Visualization of Complex Architectures.* 25-27 February 2015, Avila, Spain.

Castleman, K.R. (1998). *Digital Image Processing.* 2nd edn. Prentice-Hall, Inc. 667 pp.

Fuller, B.D. (1966). Two-dimensional frequency analysis and design of Grid Operators. In *Society of Exploration Geophysicists Mining Geophysics. Vol II. Theory.* (Ed: D.A. Hansen).

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing.* Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990*). Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective.* 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Mitchell, O.R., Delp, E.J., and Chen, P.L. (1977). Filtering to remove cloud cover in satellite imagery. *IEEE Transactions on Geoscience and Remote Sens*ing, 15, 137–141.

Moik, J.G. (1980). *Digital Processing of Remotely Sensed Images.* NASA SP–431. Washington, DC. USA.

Oppenheim, A.V., and Schafer, R.W. (2004). From frequency to quefrency: A history of the cepstrum. *IEEE Signal Processing Magazine,* 21(5), 95–106.

Pitas, I., and Venetsanopoulos, A.N. (1990). Homomorphic Filters. Ch 7 in *Nonlinear Digital Filters.* The Springer International Series in Engineering and Computer Science (VLSI, Computer Architecture and Digital Signal Processing), Vol 84. Springer, Boston, MA

Richards, J.A. (2013). *Remote Sensing Digital Image Analysis: An Introduction.* 5th edn. Springer-Verlag, Berlin. ISBN 978-3-642-30061-5

Wallis, R. (1976). An approach to the space variant restoration and enhancement of images. *Proc. Symp. Current Mathematical Problems in Image Science,* Monterey, Naval Postgraduate School, pp. 329–340.

32

Earth Observation: Data, Processing and Applications. Volume 2: Processing

4 Smoothing (Low Pass)

The operation of spatial filters is introduced in Section 3.1. Various forms of smoothing filters are available, which replace the value of the central pixel in a pre-defined filter region with another value determined from the values of pixels within that filter region. Commonly available filters for image smoothing include:

- average or mean, which compute the (weighted) average of all pixels in the filter region (see Section 4.1);
- median, which select the median value of pixels within the filter region (see Section 4.2); and
- mode, which select the most common pixel value in the filter region (see Section 4.3).

The effects of these three types of filtering operation are illustrated for different filter sizes in Figure 4.1.

Other forms of smoothing filters include:

- edge-preserving filters (see Section 4.4), which are based on adaptive windows (see Section 3.5);
- despiking, which replace noise 'spikes' (pixels with markedly different values to their neighbours) with values derived from neighbouring pixels (see Section 4.5); and
- interpolation, which fill missing pixels with values derived from neighbouring pixels (see Section 4.6).

He was always smoothing and polishing himself, and in the end he became blunt before he was sharp.
(Georg Christoph Lichtenburg)

Figure 4.1 Effect of different filter sizes and operations

a. Average filtering (using weights of 1 with a divisor equal to the number of pixels in the filter)

Original Image   Filter size: 3×3   Filter size: 7×7   Filter size: 11×11

b. Median filtering

Original Image   Filter size: 3×3   Filter size: 7×7   Filter size: 11×11

c. Modal filtering

Original Image   Filter size: 3×3   Filter size: 7×7   Filter size: 11×11

Source: Harrison and Jupp (1990) Plates 2 and 3

4.1 Average Filters

Image smoothing is best applied to EO data using a mean or average filter. In most image processing systems, the calculation of this filter may be performed in integer or real arithmetic, the latter being slower but more precise. Similarly, average filters can often be defined as integer or real weights, with a divisor and offset being optionally applied to the filter product. The use of divisors simplifies the specification of the filter and enables scaling into the image data range. The filtering operation can generally be applied to a selection of, or all, channels in an image.

The operation of a simple average filter is introduced in Section 3.1 (see Figure 3.1). When all filter weights are defined using positive values, this process has a blurring or smoothing effect on image features with abrupt changes in value in the input image being softened into more gradual transitions in the filtered image (see Figure 4.2). Such filters are also referred to as low pass or regional variation filters as they effectively enhance the low frequency or regional information in an image by reducing higher frequencies.

Both the weighting values used in an average filter and the filter size combine to determine the type and extent of effect the filtering process will produce on an image. Smoothing average filters usually contain all positive weights with the central value being greater than or equal to other weights. The effect of a filter can be reduced by increasing its central value relative to the other values in the filter region. The filter size obviously determines the number of pixels that are involved in calculating the transformed value for each pixel, with larger filters producing a greater smoothing effect as illustrated in Figure 4.3.

Most implementations of image filtering allow both filter size and weighting values to be varied (see Figure 4.1 and Figure 4.3). The size and weights of the selected filter will depend on the application of the operation. In many image processing systems the default filter is a 3×3 average filter, with filter weights similar to those shown in Figure 4.4a. This filter is formed from the outer product of the two one-dimensional smoothing filters shown in Figure 4.4b.

Figure 4.2 Effect of smoothing filter on image data

a. Image noise is reduced with a smoothing filter.
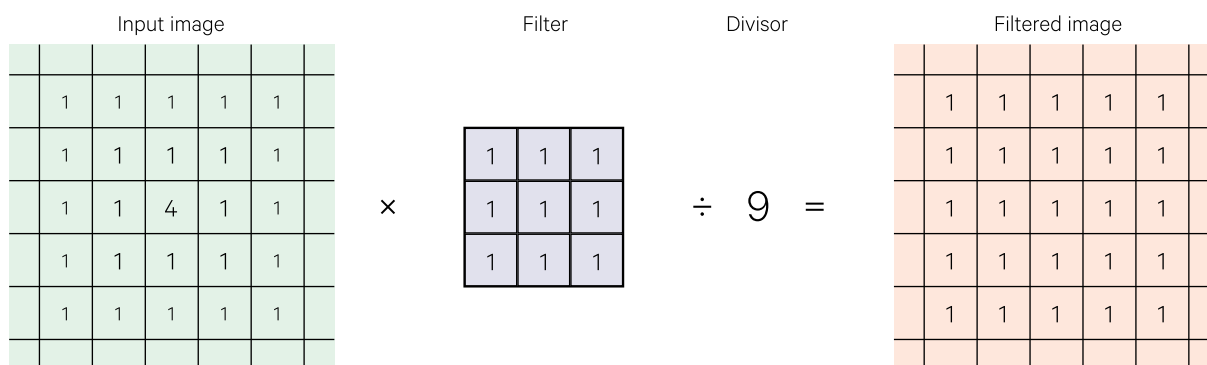
b. Linear features are broadened and reduced in intensity by smoothing. Note: in this example, the filtered values for edge pixels have been computed assuming that the pattern of the sub-image is continuous.
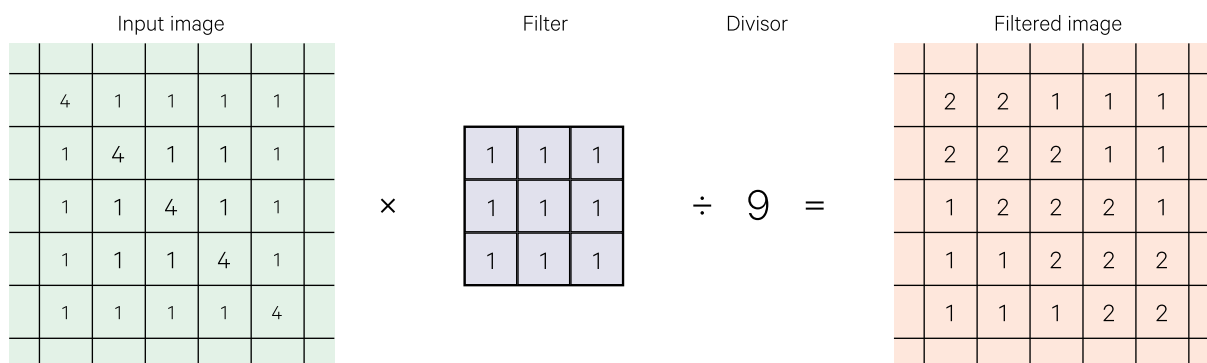
Source: Harrison and Jupp (1990) Figure 71

Figure 4.3 Effect of different filter sizes and weights

a. Weighted average filtering (using filter weights shown in Figure 4.4a)

Original Image	Filter size: 3×3	Filter size: 7×7	Filter size: 11×11

b. Full average filtering (using filter weights shown in Figure 4.5a)

Original Image	Filter size: 3×3	Filter size: 7×7	Filter size: 11×11

Source: Harrison and Jupp (1990) Plate 2

This is a relatively 'gentle' filter, that is, the value of the central pixel is largely retained, with the surrounding pixels contributing to the filtered value in proportion to their distance from the filter centre (the centres of pixels in adjacent rows and columns being closer to the centre of the filter than the centres of the four diagonal pixels). The default filter weights for larger filter sizes are derived by convolution of two smaller filters as detailed in Volume 2X—Appendix 5. The impact of these filter weights on image values is illustrated in Figure 4.3a using the default filter defined in Figure 4.4a.

Filtering operations such as smoothing generally do not rescale the data range of an image channel. Non-rescaling is ensured in an average filter if the sum of filter weights divided by the divisor equals 1. Figure 4.5 shows two simple full average filters with Figure 4.5b being a shaped window version of Figure 4.5a. Such filters would have a strong blurring effect on the image data values (see Figure 4.3b).

Figure 4.4 Weights for gentle smoothing filter

a. This gentle smoothing filter could be used to soften EO imagery.

1/36	4/36	1/36
4/36	16/36	4/36
1/36	4/36	1/36

or

1	4	1
4	16	4
1	4	1

with divisor = 16

b. This filter is formed from the outer products of two one-dimensional filters.

1
4
1

and

1	4	1

36

Earth Observation: Data, Processing and Applications. Volume 2: Processing

Figure 4.5 Full average filters with differing weights

a. A square filter, with all cells active, produces heavily smoothed image values (see Figure 4.1a and Figure 4.3a).

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

or

1	1	1
1	1	1
1	1	1

with divisor = 9

b. A filter with inactive corner cells does not consider the diagonally adjacent pixels when computing the filtered value. Smoothing filter weights are typically graduated to have higher values for those pixels that are closest to the central value. This graduation results in a 'softer' smoothing effect than can be achieved using a filter with equal weights in all cells.

0	1/5	0
1/5	1/5	1/5
0	1/5	0

or

0	1	0
1	1	1
0	1	0

with divisor = 5

Source: Harrison and Jupp (1990) Figure 72

In some implementations of average filtering, the effect of the overall filtering operation may also be varied by specifying a weighting value (λ), with values between 0 and 1. The weighting value is used to compute a weighted average of the filtered result and the original value for each pixel and is defined as:

$$\text{transformed value} = \lambda \times \text{original value}$$
$$+ \left(1 - \lambda\right) \times \text{filtered value}$$

so that a λ value of 0 results in full smoothing and a value of 1 retains the original pixel value.

In some image processing systems, threshold values may also be specified for each image channel to limit the smoothing operation to pixels where the percentage difference between the smoothed and original values (that is, *difference / original value* × 100) is greater than the threshold percentage value (see Section 3.4). Specification of the threshold as a percentage, rather than as a digital value, results in a greater filtering effect on low image values. This is usually desirable for EO data given its typically skewed distribution. Image histogram information may be useful when selecting appropriate levels for filter thresholds (see Volume 2A—Section 8.1.1).

The average filter operation can be used to reduce image noise but also blurs edge features. Within image patches such as vegetated paddocks, this process has an equalising effect so that value variations are reduced. However, definition of linear features such as edges or roads is also reduced by average filtering, although this can be advantageous for image destriping as discussed in Section 2.4. Noisy imagery often benefits from a smoothing operation before edge detection or enhancement filters are used since these operations will enhance both noise and edges.

An unbalanced filter design may be selected in an average filter to destripe an image. For example, to destripe imagery with 6 line (horizontal) striping, we could use a 7×7 filter (that is, greater than the six line striping pattern) with all values except those in the central column of the filter being zero so that only horizontal patterns in the image are modified (see Figure 4.6). Such weights should minimise the blurring effect of an average filter.

Image average smoothing effectively represents a process of removing edges from features within an image. The impact of smoothing can be seen by differencing the original and filtered images as shown in Figure 3.4 (see Volume 2D regarding image differencing). As such, this transformation performs a complementary process to edge enhancement (see Section 5.2).

Figure 4.6 Example of destriping filter

The divisor equals the sum of the coefficients to retain the same range of values in the image.

0	0	0	1	0	0	0
0	0	0	4	0	0	0
0	0	0	16	0	0	0
0	0	0	64	0	0	0
0	0	0	16	0	0	0
0	0	0	4	0	0	0
0	0	0	1	0	0	0

with divisor = 106

37

4.2 Median Filters

Image smoothing can also be effected by selecting the median of the image values within the filter region as the replacement value of the centre pixel. The median of a set of values is the middle value by count when the values are ordered (see Figure 4.7). Smoothing on this basis allows single pixels with anomalous values to be totally disregarded. By definition, the actual value of the median for the filter region already occurs in the image so it could be considered a 'reasonable' replacement value for the central pixel.

The difference between average and median filtering is demonstrated for a one-dimensional function (such as a line of image values in one channel) in Figure 4.8. The mean or average, a_m, of a set of points, x_i, would minimise the expression:

$$\sum \left(x_i - a_m \right)^2$$

whereas the median, a_n, minimises the expression:

$$\sum \left| x_i - a_n \right|$$

An example of results from moving average and moving median is given in Figure 4.9.

Figure 4.7 Operation of a median filter

a. The filtered value for the central pixel with value a in the input image is the median (or middle ranking value) of the values in the filter region—in this case, value d.
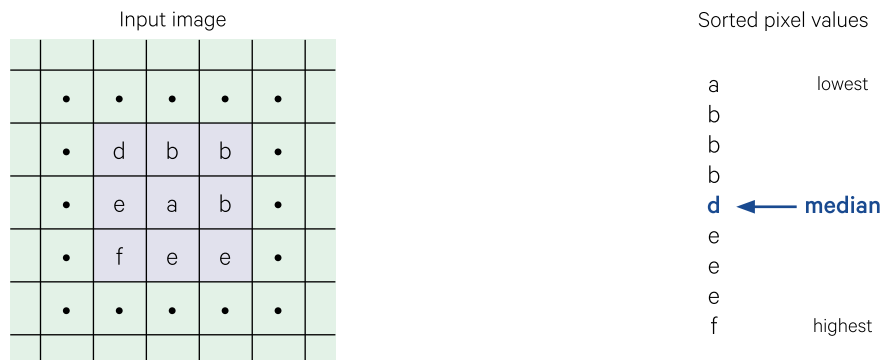
b. Numeric example. If an odd number of null values exist in the filter region, then two median values can occur. In this case, the two medians are generally averaged to generate the output value.
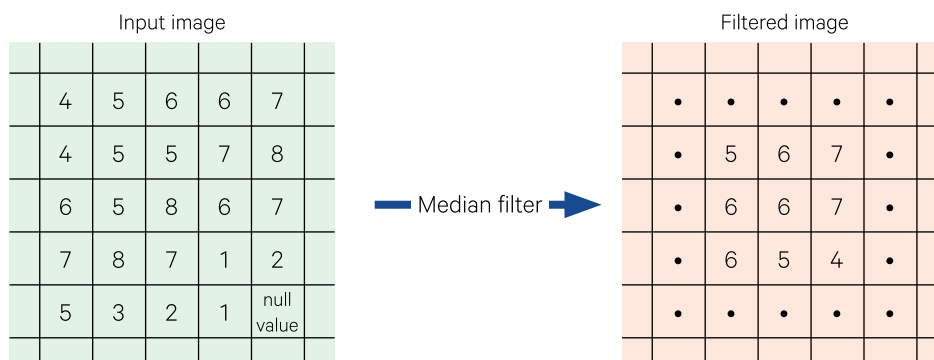
Source: Harrison and Jupp (1990) Figure 73

Figure 4.8 Average versus median smoothing

The effect of mean and median filtering is shown for a one-dimensional function. The median filter removes outlier values and preserves edges more effectively than the mean filter.
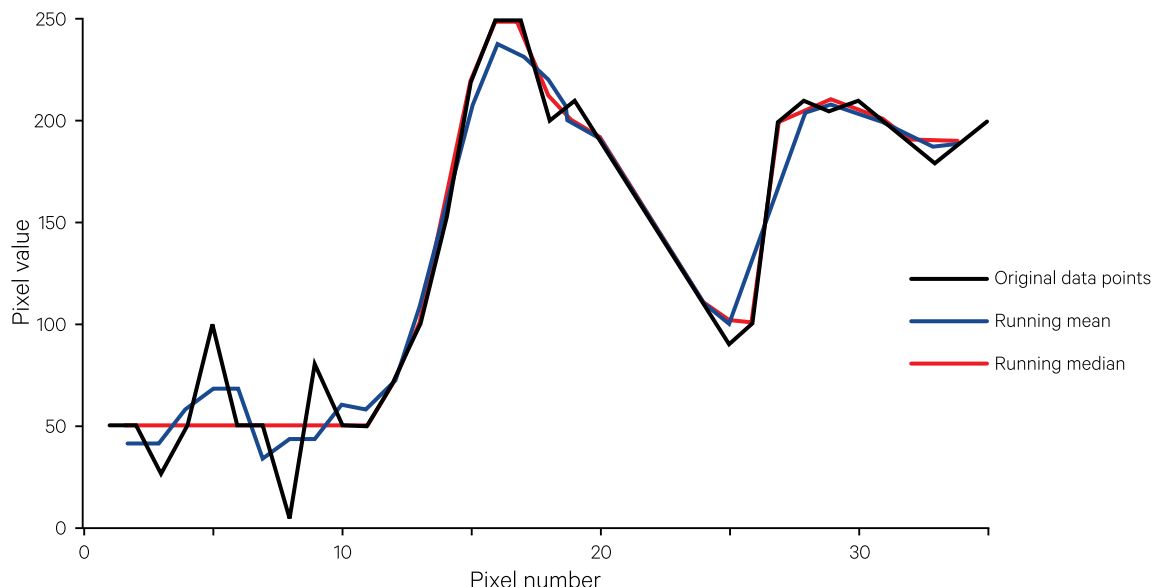
Source: Harrison and Jupp (1990) Figure 74

Figure 4.8 and Figure 4.9 demonstrate the way a median filter removes outlier points whereas averaging merely spreads their effect into adjacent data points. Sudden changes in data values however, such as feature boundaries, are also better preserved by median filtering as illustrated in Figure 4.1.

Repeated median filter operations tend to quickly reach a stable end point (Tukey, 1977) whereas average smoothing can cycle through a series of minor variations in pixel values. This form of smoothing provides a powerful tool for noise reduction in image processing with less loss of edge information. This filtering mode is commonly used in conjunction with thresholds for despiking an image as described in Section 4.5.

Median filtering can be applied in most image processing systems. The effect of filter size with a median filtering is illustrated in Figure 4.1 above. If an even number of non-null pixels occurs in a filter region (that is, some pixels may be equal to the null value so are excluded from the filtering calculation) the average of the two mid-range values is generally computed as the output median value.

Figure 4.9 Example of moving average and moving median

a. For the dataset below, the mean value is 15 / 6 = 2.5 while the median values is 1.0.

1.0	0.5	1.0	1.5	9.0	1.0	1.0

b. The moving average values (that is the average of each triplet of values) would be:

•	0.83	1.0	3.8	3.8	3.7	•

c. The moving median values are:

•	1.0	1.0	1.5	1.5	1.0	•

4.3 Modal Filters

Imagery can also be smoothed by selecting the mode, that is, the most 'popular' value, within the filter region as shown in Figure 4.10. This approach is particularly useful for nominal (categorical) data where the ordering of values does not imply a relationship between them since their average or median would be a meaningless result. The effect of variations in filter size in the modal filtering operation is illustrated in Figure 4.1 above.

This operation performs a type of 'logical' smoothing by assuming that the central pixel should be related to its surrounding pixels. Typically, in an image classification of EO data for example, patch size will be larger than a single pixel so that patches consisting of only one pixel are likely to be noise. When the classification is represented as a single image channel, modal filtering can be used to replace isolated pixel patches with the most commonly occurring class in the filter region. Modal filtering can often be conducted in conjunction with thresholds,

which only permit the centre pixel to be replaced by a filtered value if the mode occurs in a minimum number of pixels within the filter region (see Section 3.4).

This filtering operation is particularly useful for nominal data, such as an image classification channel, where the original image values can be reorganised but not interpolated (see Figure 4.11). When a filter region is bi-modal, that is there are two modes, a predefined rule is needed to determine whether the one with the lowest or highest image value is used. Any automatic selection of the lowest or highest value in bi-modal filter regions assumes that an implicit priority order exists in the nominal data, which should be considered when the data are coded.

Nominal data may be derived from EO imagery, such as a classification channel, or created to represent ancillary data from map sources. The modal filter is a very effective image 'cleaning' tool after classification to remove small, disconnected clusters (see Volume 2E).

Figure 4.10 Operation of a modal filter

a. The filtered value for the central pixel (with value c in the input image) becomes the most commonly occurring value, or mode, in the filter region—in this case, value b.
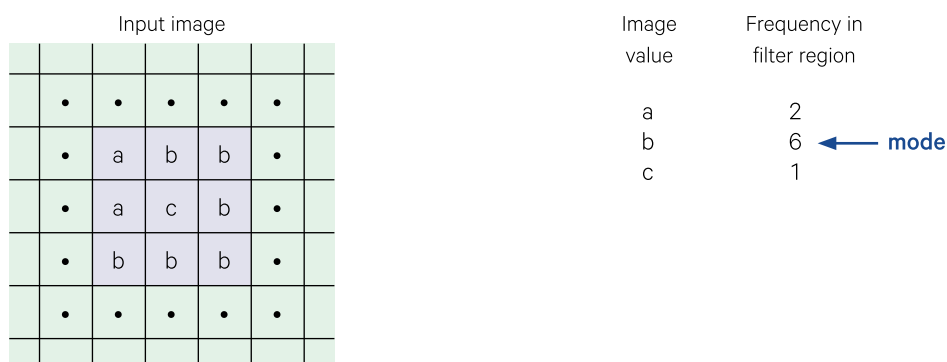
b. Numeric example. Where two or more values are equally 'popular' in the filter region, a predefined rule determines which value is used as the model. In this example, the lowest value is selected for the output pixel value.
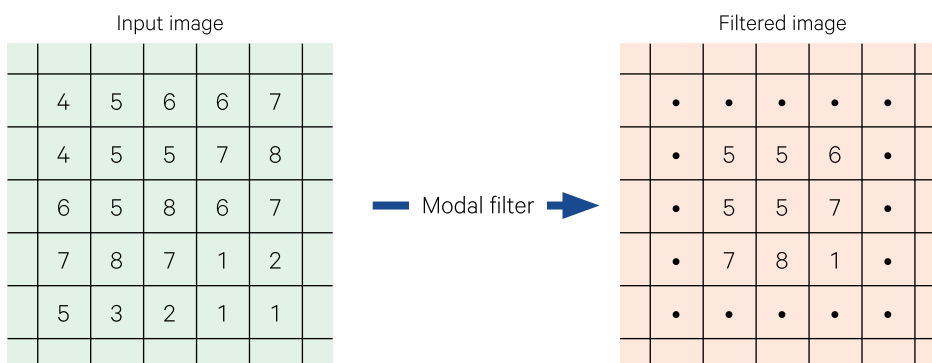
Source: Harrison and Jupp (1990) Figure 75

Figure 4.11 Modal filtering of classified image

a. Original classified image
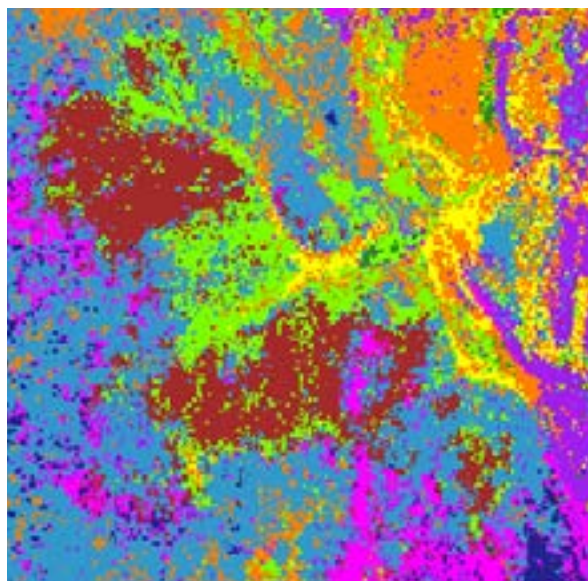
b. Classified image after three passes of 3x3 modal filter
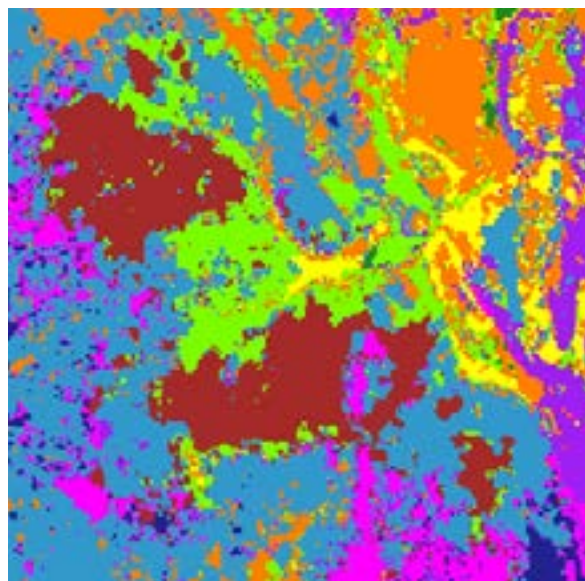
Source: Megan Lewis, University of Adelaide

4.4 Edge-Preserving Filters

As the name implies, edge-preserving smoothing filters are designed to smooth large, homogeneous patches in an image while preserving edges. These filters are particularly appropriate for pre-processing of noisy images (Cheng, 1989), including SAR data (Argenti *et al.*, 2013; Kupidura, 2016). Various algorithms have been proposed for edge-preserving filtering, not all of which are appropriate for EO imagery.

One type of edge-preserving filters, originally designed for medical imaging (Kuwahara *et al.*, 1976) uses adaptive windows within the filter (see Section 3.5). This technique involves selecting the best sub-window to use for the filtering process. The selection of sub-windows is based on some relevant criterion such as minimum variance (see Figure 3.5) or local standard deviation. This process removes patches smaller than the filter size and results in an image in which the boundaries of major patches are preserved rather than blurred, as typically occurs with smoothing operations. Other categories of edge-preserving filters include bilateral filters (which use differences between pixel values to determine the contribution of each pixel to the local average) and anisotropic diffusion filters (which iteratively smooth the image using the image gradient to prevent local 'diffusion' at edges; Perona and Malik, 1990; Burger and Burge, 2016).

The effect of filter size on edge-preserving filtering is illustrated in Figure 4.12. Two average filters are compared with an edge-preserving filter for a scanned photograph in Figure 4.13. The differences between the filters are evident in this example.

Alternative approaches to adaptive filtering rely on computing the difference between the central pixel and the mean of the filter region. For example, Eliason and McEwen (1990) use this difference in two ways:

- to remove random errors—pixel values are only replaced with the local mean when they differ from it by more than 1–2 standard deviations; and
- to smooth image noise—when computing the value for the central pixel, this approach only averages those pixels within the filter whose values are within 1–2 standard deviations of it (see Figure 4.14). This approach is similar to the sigma filter proposed by Lee (1983) and is often called a 'local sigma' filter.

41

Figure 4.12 Effect of filter size on edge-preserving filters

a. Original Image
b. Filter size: 3×3
c. Filter size: 7×7
d. Filter size: 11×11

Source: Harrison and Jupp (1990) Plate 4

Figure 4.13 Comparison of average and edge-preserving filters

A scanned photograph has been filtered using three different 3×3 filters.

a. Original image
b. Full average filter
c. Weighted filter
d. Edge-preserving filter

Source: Harrison and Jupp (1990) Plates 2 and 4

Figure 4.14 Edge-preserving filter applied to EO image

A Landsat-8 OLI image of Canberra, acquired on 10 August 2017 and shown with bands 6, 5, 3 as RGB, has been filtered in two ways. Large homogeneous patches are more even in both smoothed images but edges are visibly clearer after the edge-preserved filtering (Eliason and McEwen, 1990) compared with full average filtering. In this example, the active region of both filters was 3×3 pixels and the difference between them highlights edge features. The difference channel is based on Euclidean distance with brighter pixels indicating larger differences. Note that different treatment of edge pixels by the two filters are visible along the top and left edges of the difference image[3].

a. Original image

b. Edge-preserving filter (5x5)

c. Full average filter (3x3)

d. Difference between filtered images
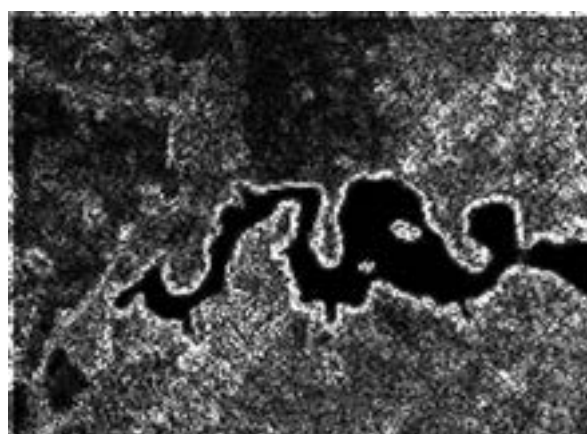
Source: Norman Mueller, Geoscience Australia

4.5 Despiking (Noise Reduction)

A specialised use of image smoothing, called despiking, may be applied to imagery to remove outlier pixels, or noise 'spikes'. These pixels can occur in features of relatively uniform 'colour' and have unrealistically high or low values. For example, image spikes commonly occur in water features, so despiking is a standard pre-processing operation in various shallow water mapping analyses.

A measure of similarity between the central pixel and its neighbours can be used to identify image spikes. Then for every central pixel which is not 'similar' to its neighbours, its value can be replaced by a new estimate based on the neighbourhood values, or by a null value. This general function can be implemented to allow:

- selection of filter neighbourhood size;
- different measures to determine similarity;
- threshold values to identify spikes;
- replacement values based on neighbourhood values; and
- iteration of the despiking process.

---

3   In this example filtering was applied to a larger image so that the edge effect is only visible on two edges of the sub-image.

Measures of similarity can be defined to filter either by:

- outlier—based on some 'distance' from neighbouring values; or by
- rarity—based on the frequency of values occurring in the neighbourhood.

These two basic types of similarity generally apply to numerical and nominal data respectively (see Section 1.2.2 for definition of these terms).

A suitable measure for an outlier would be the numerical 'distance' between the central value and the values of the pixels in the selected neighbourhood (see Volume 2E). The distance function could be the sum of squares:

$$d_2\left(x_{ij},N\right)=\sum_{k,l\in N}\left(x_{ij}-x_{kl}\right)^2$$

or the sum of absolute values:

$$d_1\left(x_{ij},N\right)=\sum_{k,l\in N}\left|x_{ij}-x_{kl}\right|$$

where $d_p(x_{ij},N)$ (with p = 1, 2) represents the distance between the central pixel's values (x_{ij}) and those of the pixels in the filter region, or neighbourhood N. The central values that would minimise these distance functions are respectively the mean and median of the pixels in the filter region excluding the central pixel. That is, the neighbourhood mean and median values are most similar to the neighbourhood values as measured by the two distance functions where the neighbourhood can be selected from (usually square) patches of size 3×3 pixels and larger.

An outlier is defined as a pixel for which the distance function is greater than a selected threshold, which is commonly expressed as a percentage of the neighbourhood mean or median. Outliers are typically replaced by the mean, median or mode of the neighbourhood values or a null value. In this way, images can be despiked to remove isolated extreme values very efficiently. This process also alters the neighbourhoods so that a pixel that is not an outlier in the original image may become one if the process is iterated. Thus, iterative despiking can be used to remove higher densities of extreme values.

Rarity allows filtering with a similarity measure that is appropriate to the mode (or most common value) of the neighbourhood. A rarity filtering operation could be described as the number of pixels in the neighbourhood that are NOT the same as the central pixel in value:

$$d_R\left(x_{ij},N\right)=\sum_{k,l\in N}\left(1-\delta_{x_{ij}x_{kl}}\right)$$

where δ_{xy} is 1 if x=y, and 0 otherwise. The value of x_{ij} that minimises this distance measure is clearly the mode of the values in the filter region. This distance measure and the mode could be evaluated for any data type, including classified pixels and even null values.

A pixel is considered 'rare' if $d_R(x_{ij},N)$ is greater than some selected numerical threshold. If a pixel is rare, it can be replaced by the mean, median or mode of the values in the filter region or with a null value. The threshold could be any value between 1 and the number of pixel neighbours in the filter region. If the filter region has a radius k the threshold value could be between 1 and $((2k+1)^2-1)$. Using this definition, a large threshold value would indicate that the value of the central pixel is more rare, that is, its value differs from the values of most of its neighbouring pixels.

The key feature of such a rarity option is that it can be applied to nominal data and to nulls with very convenient results for image cleaning. Rare null values can be 'filled' by local means, medians or modes. If null values are scattered through a classification channel, for example, they can be 'filled' using the mode of the data within neighbourhood N. Alternatively, if null values are scattered through numerical image data, they may be 'filled' by the mean or median of the (non-null) data in the neighbourhood.

Despiking can be simply applied to imagery in two ways:

- removing the spike pixels with specialised algorithms then filling them with the neighbourhood mode value; or
- using a median filter with 70% thresholds.

Some specialised algorithms identify spikes using a median filter with selected thresholds then set those pixels to the null value in the output image. The use of 70% threshold values during median filtering ensures that a pixel is only altered if its value differs from the median of the filter by 70% or more. Using the median value to replace noisy pixels enforces the use of an existing image value, which is more reliable than the neighbourhood mean value for spikes covering more than one pixel.

Appropriate thresholds for filtering can be determined either by trial and error, or statistically (such as three standard deviation units; see Volume 2A—see Section 8.1.2). The null pixels can be subsequently filled using the mean, median or mode value of the surrounding pixels. This sequence is the preferred method for despiking since it uses the median to detect outliers then allows them to be replaced using some other criteria. The mode value is recommended in this sequence for classified channels.

4.6 Filling Missing Values (Interpolation)

Filtering techniques may be used to interpolate missing values in image data. Such pixels are usually given the image null value. Scanner or reception malfunctions can occasionally cause lines or partial lines in EO images to be 'lost' (see Volume 1B—Section 2). During processing, pixels may be deemed 'missing' after certain transformations (such as smoothed ratios with themes—see Section 10.1.3) or left unclassified after image analysis (see Volume 2E). The boundaries of segmented imagery may also be filled lightly to reduce jagged edges (see Volume 2A—Section 9).

Alternatively, data from non-EO sources, such as water depth transects or terrain elevation, may be registered with an image base but form a sparse dataset relative to the image grid. Provided such data does not contain large contiguous areas of missing values, filter interpolation techniques may be used to fill the missing values (see Volume 2D).

Filter-based interpolation techniques operate in a similar fashion to image smoothing with only the values of the pixels in a 3×3 filter region surrounding the central 'missing' pixel being used to compute the filled value. This process is illustrated in Figure 4.15 using the filter shown in Figure 4.16 for the weighted mean.

Figure 4.15 Filter-based interpolation

Whenever the central pixel in a 3×3 neighbourhood is equal to a specified mask value (such as the null value), it can be replaced by the median, mode or weighted mean (as selected) of the surrounding 'active' pixels (that is, values not equal to the specified value). A minimum number of active pixels can be specified before the central pixel is filled. In this case, the output value for the central mask value could be the:

- median of (a, b, c, d, e, f, g); or
- mode of (a, b, c, d, e, f, g); or
- weighted mean of $\dfrac{a+4b+c+4d+4e+f+4g+h}{20}$

Input image

Source: Harrison and Jupp (1990) Figure 76

Filter interpolation is generally based on a 3×3 mean, median or modal filter to fill missing pixels. An average filter would use predefined weights, such as shown in Figure 4.16. The missing pixel value may generally be selected as any valid pixel value in the image data range (for example 0–255 for byte format images). To control the effect of this transformation, a minimum number of neighbouring non-missing pixels could be specified to ensure that the operation is only applied to isolated missing pixels. If more than one pixel is missing in the neighbourhood, the filter weights would need to be adjusted accordingly to preserve image scaling.

An example of average filter interpolation is shown in Figure 4.17. This example is based on an image acquired by Landsat-7 ETM+ after the Scan Line Corrector (SLC) had stopped operating (see Figure 4.17a). The SLC accounted for forward motion of the satellite and ensured that image lines represented contiguous ground pixels of constant width, without any duplication (USGS, 2018). Geometric pixels that do not have valid image values are indicated as null (missing) pixels in this imagery and shown as black in Figure 4.17a. When appropriate interpolation filtering is applied to such imagery, a continuous image can be created as shown in Figure 4.17b.

With median filtering, if an even number of 'active' (non-missing) pixels occurs in the filter region, a decision needs to be made as to which value to use for the replacement value in the output image. In some image processing systems, the lower of the two mid-range values is used as the replacement value. Modal interpolation filtering of a labelled classification channel provides a convenient method for assigning spatial labels to any unclassified pixels (see Volume 2E).

While large targets may be filled iteratively using this process, care must be taken to ensure that the results represent a reasonable interpolation of the original data (see Volume 2D). Filter-based interpolation is not appropriate for large area interpolation; it simply provides a mechanism for tidying spotty image data or creating a continuous data channel from thickly sampled ancillary data (see Figure 6.9).

Figure 4.16 Example of average filter

1/20	4/20	1/20
4/20	•	4/20
1/20	4/20	1/20

Figure 4.17 Average filter interpolation

This Landsat-7 ETM+ image was acquired over the Black Saturday fires in central Victoria in 2009.

a. Original image showing missing pixels due to SLC failure.

b. Interpolation filtering creates a continuous image.
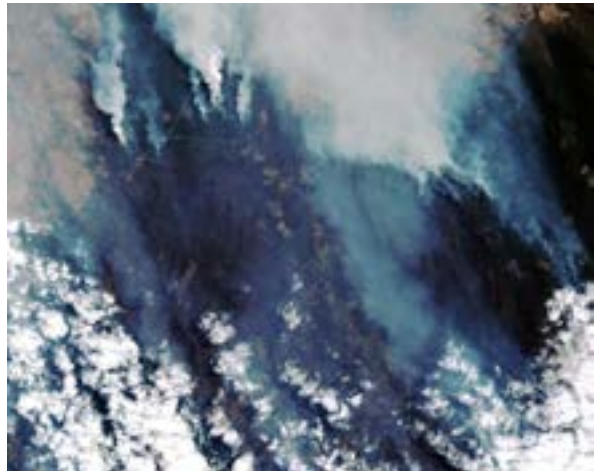
Source: Norman Mueller, Geoscience Australia

4.7 Further Information

Castleman (1988) Sections 9, 11 and 12

Gonzalez and Woods (2018) Section 3

Jensen (2016) Section 8

4.8 References

Argenti, F., Lapini, A., Bianchi, T., Alparone, L. (2013). A Tutorial on Speckle Reduction in Synthetic Aperture Radar Images. *IEEE Geoscience and Remote Sensing Magazine*, 1(3), 6–35.

Burger, W., and Burge, M.J. (2016). *Digital Image Processing*. Springer, London.

Castleman, K.R. (1998). *Digital Image Processing*. 2nd edn. Prentice-Hall, Inc. 667 pp.

Cheng, S. (1989). An adaptive noise smoothing filter for remotely sensed images with microphonic noise. *International Journal of Remote Sensing*, 10, 1015–1034.

Eliason, E.M., and McEwen, A.S. (1990). Adaptive Box Filters for Removal of Random Noise from Digital Images. *Photogrammetric Engineering and Remote Sensing*, 56(4), 453.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Kupidura, P. (2016). Comparison of filters dedicated to speckle suppression in SAR images. *International Archives of Photogrammetric, Remote Sensing and Spatial Information Sciences* XLI-B7. 12–19 July, Prague, Czech Republic. pp. 269–276. doi:*10.5194/isprsarchives-XLI-B7-269-2016*

Kuwahara, M. Hachimura, K., Eiho, S., and Kinoshita, M. (1976). Processing of RI-angiocardiographic images. Ch 13 in *Digital Processing of Biomedical Images*. (Eds: K. Preston Jr. and M. Onoe). Springer, Boston. p. 187–202. ISBN 978-1-4684-0771-6

Lee, J.S. (1983). Digital image smoothing and the sigma filter. *Computer Vision, Graphics and Image Processing*, 24, 255–269.

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7),* 629–639.

Tukey, J.W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, Mass.

USGS (2018). *SLC-off Products*: *Background* webpage: *https://landsat.usgs.gov/slc-products-background*

46

Earth Observation: Data, Processing and Applications. Volume 2: Processing

5 Highlighting Edges

Edge features in EO imagery can be highlighted using differential filters. The following sub-sections consider filters that:

- detect edges (see Section 5.1); and
- enhance edges (see Section 5.2).

5.1 Edge Detection (Differential)

The filters described for image smoothing manipulate image values using balanced filter weights around the central pixel value (see Section 4). The effect of smoothing operations is to effectively reduce the strength of the central pixel value relative to those other pixels surrounding it.

In a 'balanced' filter, the same weighting values are used for all adjacent pixels at the same linear distance from the central pixel. Filters can also be designed as 'unbalanced' to highlight or reduce the effect of image features that have a directional component such as roads, lineaments or detector striping. Such filters basically rely on differential calculus, or derivatives, to achieve their result. Before considering the design of differential filters, we will review the meaning of the derivative in Excursus 5.1.

Excursus 5.1—Differential calculus review

From high school mathematics you may remember that the first derivative of a one-dimensional function, say

$$f(x)$$

is written as

$$f'(x) \text{ or } \frac{df}{dx}$$

In terms of image values, x could represent a row or column of pixels in an image that have the values $f(x)$ where these values can somehow be modelled in terms of pixel position.

For example, if the image values across an image line are as shown above the graph in Figure 5.1a, we could represent these as a graph (or a 'spectral transect'), then plot a continuous function line to describe the

changes in image value in terms of position along the line. If we know the equation for this line, we could use differential calculus to determine the derivative equation, then compute a derivative value for each pixel along the line. For non-continuous functions, such as image data, these calculations can be approximated using filter-based processes

The initial function value, $f(x)$, describes the relationship between two variables—in the case of Figure 5.1a these are the pixel value and its position along a line. To determine the way values change along the line, we can look at the slope of the function line and identify where the values are maximum, minimum, increasing or decreasing. What we are really extracting from the graph is a visual measure of the change in the $f(x)$ value caused by a change in the x values. These observations may also be performed analytically using differential calculus.

Mathematically these changes can be expressed as the function:

$$\frac{df}{dx} = f'(x) = \lim_{\delta x \to 0} \frac{\delta f}{\delta x}$$

where

$\dfrac{df}{dx}$ is referred to as the derivative of x;

δx is the change in x values; and

δf is the change in $f(x)$ corresponding to δx.

Figure 5.1 First and second derivative for spectral transect

a. Pixel values across an image line are shown above the graph, in which the X axis represents pixel positions along an image line and the Y axis shows the pixel values.

1	4	6	7	8	9	11	9	8	7	6	4	1	0	1	2	4	7
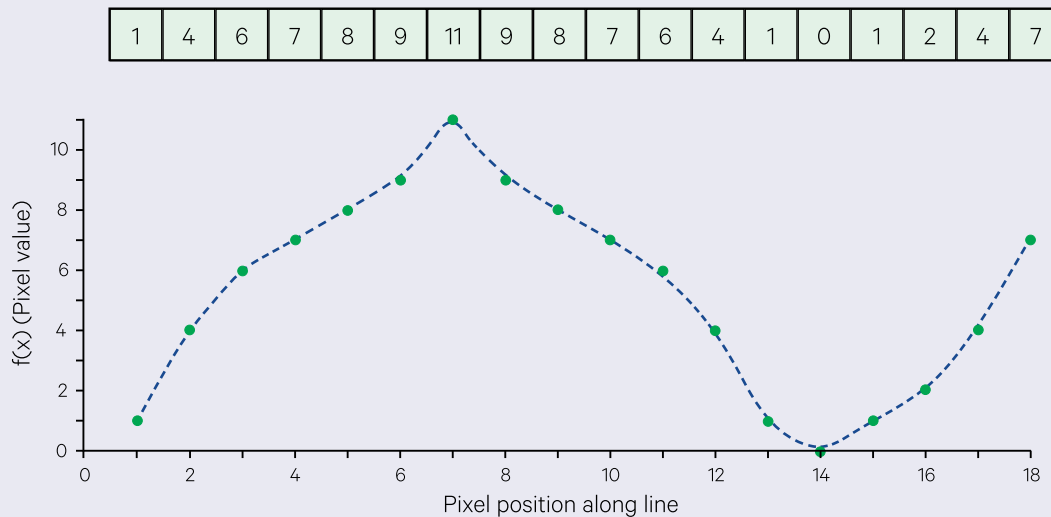
b. Central difference values are shown and graphed below. The first derivative indicates the changes in slope of the original function.

•	2.5	1.5	1.0	1.0	1.5	0	-1.5	-1.0	-1.0	-1.5	-2.5	-2.0	0	1.0	1.5	2.5	•
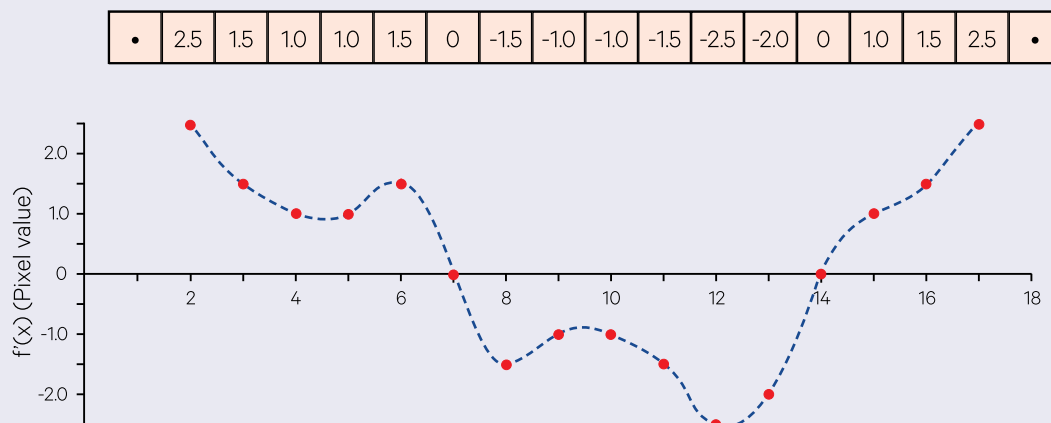
c. Second derivative values indicate the direction and extent of curvature in the original function.

•	•	-0.75	-0.25	0.25	-0.5	-1.5	-0.5	0.25	-0.25	-0.75	-0.25	1.25	1.5	0.75	0.75	•	•
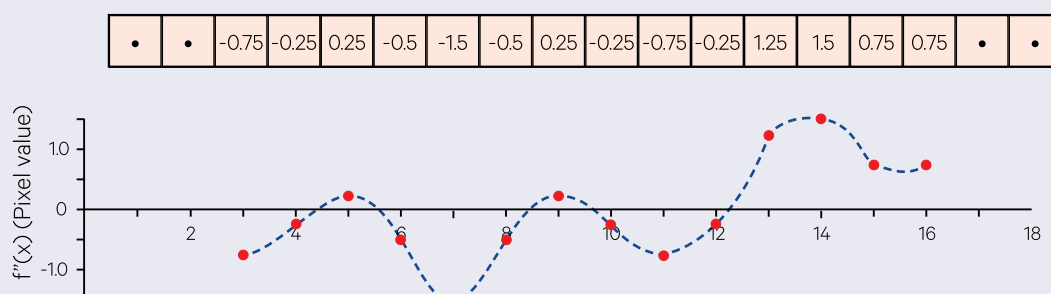
Source: Harrison and Jupp (1990) Figure 77

When the function is defined at points that are a constant distance apart, a number of operators may be used to measure its rate of change. For a forward difference, that is, the difference between values as x increases:

$$\Delta_F f\left(x\right) = f\left(x+h\right) - f\left(x\right)$$

where h is the δx, that is the number of pixel positions, or lag, between values. For a backward difference:

$$\Delta_B f\left(x\right) = f\left(x\right) - f\left(x-h\right)$$

or for the central difference (the average of the forward and backward differences):

$$\Delta_C f\left(x\right) = \frac{f\left(x+h\right) - f\left(x\right)}{2} + \frac{f\left(x\right) - f\left(x-h\right)}{2}$$
$$= \frac{f\left(x+h\right) - f\left(x-h\right)}{2}$$

These difference formulae give different results where h is not equal to, or close to, 0. The derivative function defines the way the function changes as h approaches 0. Based on forward difference, the derivative function becomes:

$$f'\left(x\right) = \lim_{\delta x \to 0} \Delta_F \frac{f\left(x\right)}{h}$$

Similarly, the 'backward derivative' is:

$$f'\left(x\right) = \lim_{\delta x \to 0} \Delta_B \frac{f\left(x\right)}{h}$$

while the derivative computed from the central difference is defined as:

$$f'\left(x\right) = \lim_{\delta x \to 0} \Delta_C \frac{f\left(x\right)}{h}$$
$$= \lim_{\delta x \to 0} \frac{f\left(x+h\right) - f\left(x-h\right)}{2h}$$

For a smooth (or 'differentiable') function all these operators give the same result.

Using the example in Figure 5.1a, we can compute and plot the derivative values for $h=1$ as illustrated in Figure 5.1b. This is done using the central difference filter to measure the slope of the function. Where we have increasing values in the original function, the derivative values are greater than zero and the slope is positive. Decreasing values in the original function give derivative values less than zero and so the slope is negative. Minimum and maximum values in the original function are where slope changes between positive and negative and so have zero values for the derivative. Thus the derivative function has quantified the relationship between slope and changes (increasing, decreasing, maximum and minimum) in the original function.

The second derivative can be obtained by applying the central difference filter again but to the derivative. The result is the slope of the derivative function, or 'slope of slope', and is closely related to the curvature of the original function. If line sections in the original function have constant slope (linear functions) over the section then the slope of slope is zero so there is no curvature. If sections in the original function have gradually increasing slope, then slope of slope is positive and sections with gradually decreasing slope in the original function have a negative slope of slope. Where the slope changes from gradually increasing to decreasing, or decreasing to increasing, the slope of slope passes through zero and slope (or derivative) is at a local minimum or maximum. Finally, when the slope is zero, the original function is at a local maximum or local minimum. In this case, if the slope of slope is positive it is a minimum and if the slope of slope is negative it is a maximum.

These observations about the value and slope of the derivative function are important for interpreting our original function as they highlight some of the more subtle changes. Since they all relate to the slope of the derivative function, they may be simply quantified by computing the second derivative or $f''\left(x\right)$.

For our example, the second derivative values and plot are as shown in Figure 5.1c. This function clearly identifies the direction and extent of curvature of the original function:

- positive values in $f''\left(x\right)$ show increasing slope in $f'\left(x\right)$ and increasing rate of change in $f\left(x\right)$;
- negative values in $f''\left(x\right)$ show decreasing slope in $f'\left(x\right)$ and decreasing rate of change in $f\left(x\right)$;
- stationary, or local maximum and minimum, points in f(x) have zero f'(x) and (respectively) negative and positive values for f"(x); and
- points of inflection in $f\left(x\right)$, that is where the sign of the curvature changes, are stationary points in $f'\left(x\right)$ and equal to zero in $f''\left(x\right)$. In Figure 5.1c, these occur between pixel values.

Just as the first derivative highlights positive and negative slopes and stationary points in the original function, more subtle changes, such as points of inflection, are significantly easier to detect in the second derivative.

5.1.1 Image-based derivatives

In image data, derivative functions may be simply derived along rows or columns in the image or along defined transects such as diagonals. These derivatives can then be used to indicate the rate of spatial variation in the defined direction. Where a dataset has more than one spatial dimension, the derivatives are computed with respect to one dimension at a time and are termed partial derivatives. For image data with two spatial dimensions, the terminology frequently used is X for across lines and Y down pixel columns in the image, with:

- f_x or $\dfrac{\partial f}{\partial X}$ being the first partial derivative function computed in the X direction for a given Y value (that is, an image line); and

- f_y or $\dfrac{\partial f}{\partial Y}$ being the first partial derivative in the Y direction for a given X value (that is, an image pixel column).

When the image data represent samples from a smooth underlying function (such as geophysical data), the derivative function may be approximated using filters. For example, the difference (or error) between the value for the central difference used as an approximation to the derivative and $f'\left(x_i\right)$ can be shown (Conte and de Boor, 1965) to be:

$$\left| f'\left(x_i\right) - \frac{\left(f\left(x_{i+1}\right) - f\left(x_{i-1}\right)\right)}{2h} \right| = \frac{h^2}{6} \times \left| f'''(y) \right|$$

where $x_{i-1} \leq y \leq x_{i+1}$. That is, when the step h is small, and the underlying function is smooth (hence the third derivative is also small), the central difference can be a very close approximation to the actual derivative.

Figure 5.2 Examples of derivative filters

a. This filter computes the derivative across a line.

-1/2	0	1/2

b. This filter computes the derivative down a line of pixels.

-1/2
0
1/2

The derivative in X and Y can also be used to compute the derivative in a particular direction θ (where θ is measured anti-clockwise from the X axis) as:

$$f_s = f_x \times \cos\theta + f_y \times \sin\theta$$

$$s = x \times \cos\theta + y \times \sin\theta$$

(see Section 6.3). The formation of a directional difference uses the same equation with differences for derivatives.

To calculate the derivative values across a line in an image for $h=1$ pixels, we can compute the difference between pairs of pixels that are separated by one pixel. This spatial difference can be easily calculated using a filter with the weights shown in Figure 5.2a. Similarly, to compute derivative values down columns in the image, we could use the filter shown in Figure 5.2b. The results of applying some simple derivative filters are illustrated in Figure 5.3.

Figure 5.3 Effect of filter direction on differential filters

The original image is shown in Figure 5.3a. The vertical striping patterns visible in the horizontally and diagonally filtered images are due to slight scanner miscalibration which is not discernible in the orginal image. Differential filters accentuate this type of image 'noise' as well as feature boundaries.

a. Original image   b. Horizontal filter   c. Vertical filter   d. Diagonal filter

Source: Harrison and Jupp (1990) Plate 5

A filter with the weights shown in Figure 5.4a also uses the spatial variation of the two adjacent image columns to compute the central value. This filter is formed from the two one-dimensional filters shown in Figure 5.4b, so is equivalent to carrying out the two operations of smoothing and differentiating. Since the differential highlights the differences between pixel values, the use of adjacent pixel values to smooth the image values before differencing usually produces better results.

We can effectively vary the *h* value used for the derivative by varying the filter size. For example, for *h*=2 we could use the filter shown in Figure 5.5a. Similarly, the direction for the filter may be expressed in filter weights. For example, a 'diagonal' filter is shown in Figure 5.5b, where *a* is √2. The '*a*' term accounts for the longer distance to the centres of diagonally adjacent pixels from the centre pixel as compared with those that are in the same row or column. The differential filter weights are usually chosen so that the result is 'exact' for some function, such as a linear function. This implies that the weights sum to zero so that the value is zero when applied to constant data values and the filter produces constant values when applied to a linear function (that is, one with constant slope separately in *X* and *Y*). In practice though, if the '*a*' value equals 1 in this filter, the image scaling would be preserved satisfactorily.

Second derivative values may be computed by re-applying the same filter to a previously filtered channel (or by convolving the two filters as described in Section 3.1). Mixed derivatives may also be computed at this stage by applying the *Y* filter to the first partial derivative in *X* or the *X* filter to the first partial derivative in *Y* to incorporate two-dimensional information.

The second order partial derivatives are:

$$f_{xx} = \frac{\partial}{\partial X}\left(f_x\right) = \frac{\partial}{\partial X}\left(\frac{\partial f}{\partial X}\right) = \frac{\partial^2 f}{\partial X^2}$$

$$f_{xy} = \frac{\partial}{\partial X}\left(f_y\right) = \frac{\partial}{\partial X}\left(\frac{\partial f}{\partial Y}\right) = \frac{\partial^2 f}{\partial X \partial Y}$$

$$f_{yx} = \frac{\partial}{\partial Y}\left(f_x\right) = \frac{\partial}{\partial Y}\left(\frac{\partial f}{\partial X}\right) = \frac{\partial^2 f}{\partial Y \partial X}$$

$$f_{yy} = \frac{\partial}{\partial Y}\left(f_y\right) = \frac{\partial}{\partial Y}\left(\frac{\partial f}{\partial Y}\right) = \frac{\partial^2 f}{\partial Y^2}$$

This set of second partial derivative values can also be treated as a matrix and manipulated to determine the direction and magnitude of the maximum curvature within the filter region. A larger filter size, such as 7×7 should be used to compute surface curvature (see Section 6.4).

Figure 5.4 Combined smoothing and differentiating filter

a. This filter combines smoothing and differentiation.

-1/6	-1/6	-1/6
0	0	0
1/6	1/6	1/6

b. These two one-dimensional filters were used to create the filter above.

-1/2
0
1/2

×

1/3	1/3	1/3

Figure 5.5 Horizontal and diagonal derivative filters

a. Horizontal filter

-1/4	0	0	0	1/4

b. Diagonal filter (a=√2)

-1/2a	0	0
0	0	0
0	0	1/2a

c. First difference

1	-1

The application of these various formulae to edge detection has been widely implemented using expressions for edge magnitude (*M*) and edge angle (*θ*) (Nixon and Aguado, 2002), where in the most usual case, writing $diff\left(x\right)$ and $diff\left(y\right)$ for the derivatives or differences in *X* and *Y*:

$$M = \sqrt{diff\left(x\right)^2 + diff\left(y\right)^2}$$

$$\theta = \tan^{-1}\left(\frac{diff\left(y\right)}{diff\left(x\right)}\right)$$

By thresholding *M,* areas of potential edge are defined and these are usually post-processed by thinning and line joining to define edges and boundaries (Paine and Lodwick, 1989).

The edge magnitude can be viewed as a distance function. Various distance functions (or 'norms', denoted $\|\ \|$), can be defined. Common expressions for norms are (Moik 1980; see Figure 5.6):

$$\|\ \|_2 = \sqrt{diff\left(x\right)^2 + diff\left(y\right)^2}$$

$$\|\ \|_1 = \left|diff\left(x\right)\right| + \left|diff\left(y\right)\right|$$

or

$$\|\ \|_\infty = \max\left[\left|diff\left(x\right)\right|, \left|diff\left(y\right)\right|\right]$$

Norms and distance functions are further discussed in Volume 2E. Some commonly used differential filters are detailed in Excursus 5.2

Figure 5.6 Norm functions

These can be used to estimate the maximum directional difference from two orthogonal distance measurements (see Volume 2E for details).
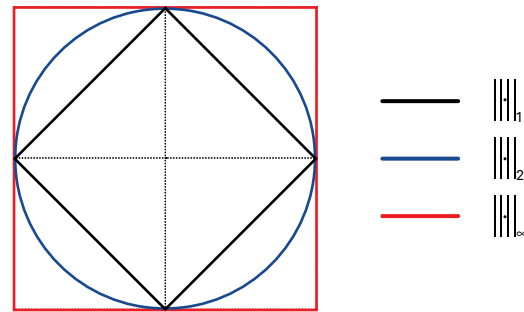
Source: Harrison and Jupp (1990) Figure 78

Imagination reveals the possibilities beyond the edges of our reality.
(L.R. Knost)

Excursus 5.2—Commonly used differential filters

The Roberts cross operator (Roberts, 1965) uses diagonal first differencing in two directions with the filters shown in Figure 5.7, and computes the magnitude and direction of the gradient using the equations given above.

Figure 5.7 Roberts cross operator

0	1
-1	0

and

0	1
-1	0

Figure 5.8 Sobel operator

-1	0	1
-2	0	2
-1	0	1

and

-1	-2	-1
0	0	0
1	2	1
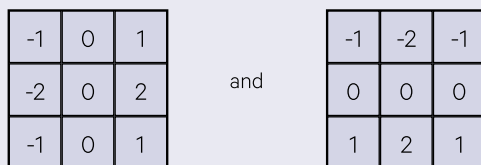
Figure 5.9 Prewitt operator

-1	0	1
-1	0	1
-1	0	1

and

-1	-1	-1
0	0	0
1	1	1
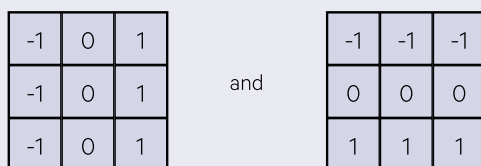
The Sobel operator (Levine, 1985) uses a similar approach to the Roberts operator with the two differential filters shown in Figure 5.8. The Prewitt operator (Prewitt, 1970) varies these weights slightly as shown in Figure 5.9. These filters incorporate local averaging to reduce noise effects so tend to produce superior results to the Roberts operator for 'sharp' edges. The use of zero weights along the 'edge' is reported to reduce jitter (Levine, 1985).

A second difference, which approximates the second derivative, can also be computed as:

$$diff_2\left(x\right) = \left(f\left(x_{i+1}\right) - f\left(x_i\right)\right) - \left(f\left(x_i\right) - f\left(x_{i-1}\right)\right)$$
$$= f\left(x_{i+1}\right) + f\left(x_{i-1}\right) - 2f\left(x_i\right)$$

or with the horizontal filter shown in Figure 5.10. As illustrated in Figure 5.11, the first difference measures the change between two adjacent pixels whereas the second difference indicates the extent of change across a pixel. Images produced by second difference filters highlight edges as adjacent light and dark lines, whereas first difference filters result in single pixel edges.

Figure 5.10 Second difference filter

1	-2	1

Figure 5.11 First and second differences

a. Original image values ($f(x)$)

b. First forward difference ($x_i - x_{i+1}$)

c. First backward difference ($x_i - x_{i-1}$)

d. First difference backward across pixel i ($x_{i+1} - x_{i-1}$)

e. Second difference ($x_{i-1} + x_{i+1} - 2x_i$)
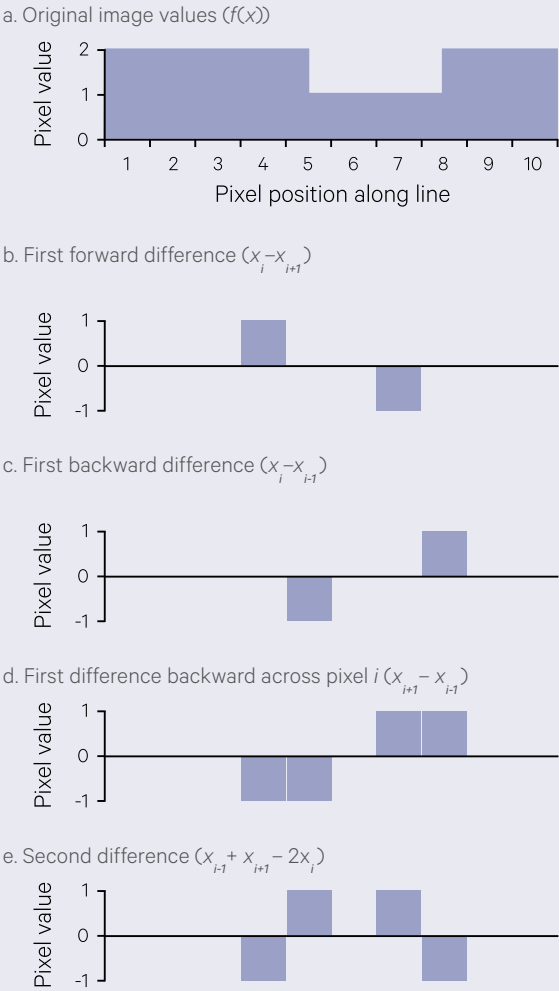
Source: Harrison and Jupp (1990) Figure 79

Figure 5.12 Laplacian edge detector

a. Omni-directional edge detector

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

b. Diagonal filter

a	0	a
0	-4a	0
a	0	a
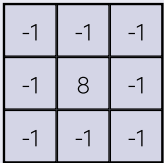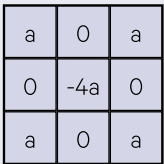
Figure 5.13 Bi-directional filter

0	1	0
1	-4	1
0	1	0

$=$

1	-2	1

$+$

1
-2
1
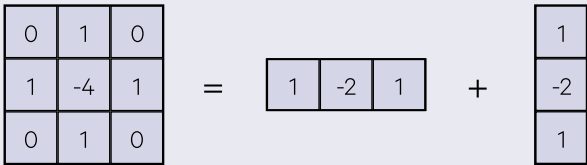
An alternative edge detector is the Laplacian operator (an approximation to the mathematical Laplacian: $f_{xx} + f_{yy}$) which combines second differences in X and Y directions. This omni-directional operator can be implemented with the filter weights shown in Figure 5.12.

This operator is sometimes referred to as the centre-point 2-ring (or nine point) filter in geophysics. Another commonly used bi-directional filter, which is also referred to as a Laplacian operator (or the centre-point 1-ring filter), is shown in Figure 5.13. This basic filter can also be rotated to highlight diagonal edges with the weights shown in Figure 5.12b, where $a = \sqrt{2}$, or convolved to operate on a larger image neighbourhood (see Section 3.1). Some specific derivative filters which were developed for geophysical data and other directional filters are introduced in Volume 2X—Appendix 5.

There are no lines in nature, only areas of colour, one against another.
(Edouard Manet)

5.1.2 Uses of edge detection

Edge detection techniques are useful for a number of EO applications. For example, the effects of using horizontal and vertical filters on landscape surfaces are illustrated in Figure 5.14. Ton *et al.* (1989) outline a method for automatic road identification, which involves the use of edge detection filters to locate roads and categorise their 'status'. Paine and Lodwick (1989) report a multi-stage approach for extracting edges in an image to identify region boundaries that involves smoothing, edge detection, thresholding, thinning and linking. The ultimate result from this type of processing would be a set of polygons that are suitable for use in a Geographic Information System (GIS; see Volume 2D).

Jacobberger (1988) reports an interesting application of directional filtering to map abandoned river channels in Landsat TM imagery. Linear dunes in the study area obscured the morphology and trend of the channel system so directional filtering was used to selectively suppress edges parallel to the dune system. In this case the filter was based on a Laplacian operator (see Excursus 5.2) that was scaled to the width and orientation of the dune edges in the image. The first principal component (PC1—see Section 9) was filtered, then the residual between PC1 and the filtered channel was used for mapping. This residual image enhanced the high frequency detail that is not related to the dune system at the expense of detail that is parallel to it.

Filters that effectively compute higher partial derivatives also exponentially enhance image noise to the power of the order of the derivative. Noisy images should be appropriately pre-processed to reduce noise effects (as discussed in Sections 2.4 and 4.5) before edge detection filtering.

Differential or derivative filters may be applied in most image processing systems, possibly using selected weightings (defined as integer or real values) in an average filter for a selected filter size (see Section 4.1)

As described above, using smoother differences and thresholding an edge magnitude function are normally used to identify edges in noisy images. If the threshold is set too high, low-amplitude edges may not be detected but a low threshold will result in noise being highlighted. Edges in a multi-channel image may be summarised into a single channel by averaging the values across relevant channels (see Section 7) or by using Principal Components Analysis (PCA; see Section 9). Density slicing of edge channels can also be used to separate edge information from other image features (see Volume 2A—Section 9.2.1).

In earlier applications of edge detection, it was popular to use the magnitudes of the Prewitt or Sobel operators and thresholding edge magnitude to define edges. More recently, edge detectors based on more symmetric Gaussian functions and methods that align to the underlying edge more precisely have been used. The Canny detection operator (Canny, 1986) is probably the most popular of these. This operator and variants of it, as well as related applications to template matching and motion detection, can be found in Nixon and Aguado (2002).

Figure 5.14 Horizontal and vertical filters

Landsat TM band 3 image of Fowlers Gap Research Station, NSW, acquired in January 2009.

a. Original image
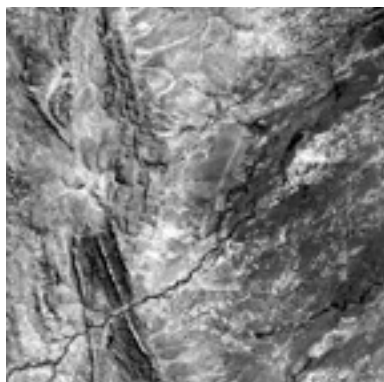
b. After filtering with horizontal filter
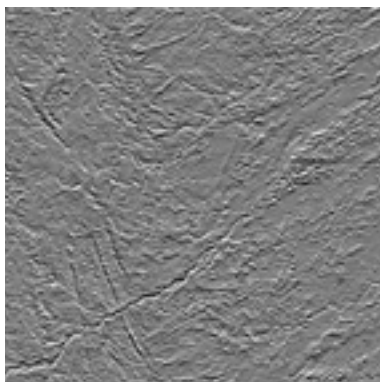
c. After filtering with vertical filter
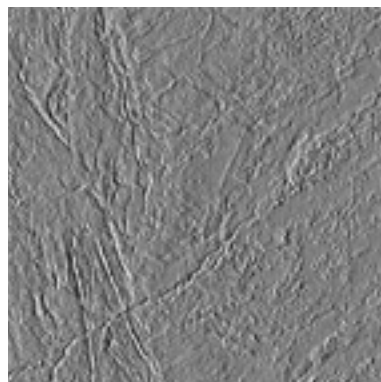
Source: Megan Lewis, University of Adelaide

The interpretation of derivative results depends on the image data being processed. For example, with elevation data the magnitude of the first derivative represents slope (or gradient) and the second derivative indicates surface convexity or curvature (see Excursus 5.1 and Section 6.4). Some image processing packages can compute slope and aspect for an image channel containing elevation data, and typically a 5×5 filter is used to estimate the local derivatives. A linearly scaled slope (or 'gradient') channel may be scaled with the minimum image value indicating a slope of 0° and the maximum value representing a slope of 90°. Aspect (which can also be interpreted as edge direction) may be linearly scaled from north at the minimum value through the mid-range value for an orientation of 180° to the maximum value corresponding to 359°. These results are further discussed in Section 6.4.

5.2 Edge Enhancement (High Pass)

Many image analysis tasks require that the features in an image be clearly defined in terms of spatial patterns and boundaries. For example, accurate positioning tasks require that linear features, such as roads, rivers or various boundaries, be clearly delineated in the image (see Volume 2B). Geological applications, in particular, frequently use edge enhancement methods to highlight lineaments. Image filtering transformations can be used to enhance the edge-related detail and also to identify edges at different scales. Such filters are also referred to as high pass or residual filters since they separate the high frequency variations in imagery from the lower frequency variations.

Section 3.1 and Figure 3.4 introduced the operation of image smoothing and showed that a smoothed image could effectively be defined as:

smoothed image = original image – edges

or

edges = original image – smoothed image

Conversely, an edge enhanced image could be described as:

edge enhanced image = original image + edges

or

edge enhanced image = original image + (original image – smoothed image)
= 2 × (original image) – smoothed image

Using some elementary arithmetic on the original and smoothed image values shown in Figure 5.15, we can use the relationship defined by the last equation to derive the weights for a simple edge enhancement filter.

The effect of edge enhancement using this filter, and varying its size, can be seen in Figure 5.16. This example filter has a typical pattern of weighting values for edge enhancing, with the central value being large and positive and surrounding values being small and negative. The effect of the filter is to exaggerate any difference between the central pixel and those pixels adjacent to it. The operation of this filter may be best understood using example image values as shown in Figure 5.17.

Figure 5.15 Derivation of edge enhancement filter
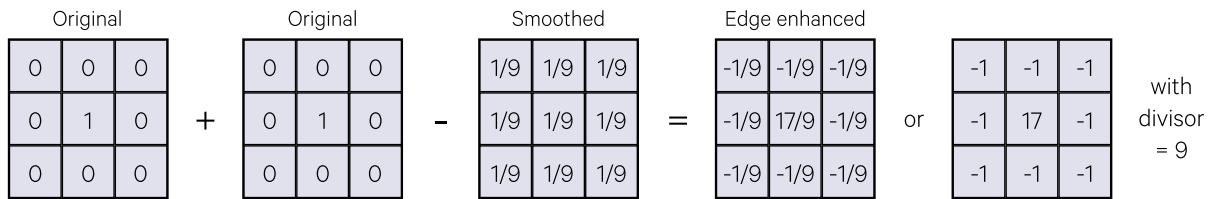
Figure 5.16 Effect of filter size on edge enhancement filters

a. Original Image   b. Filter size: 3×3   c. Filter size: 7×7   d. Filter size: 11×11
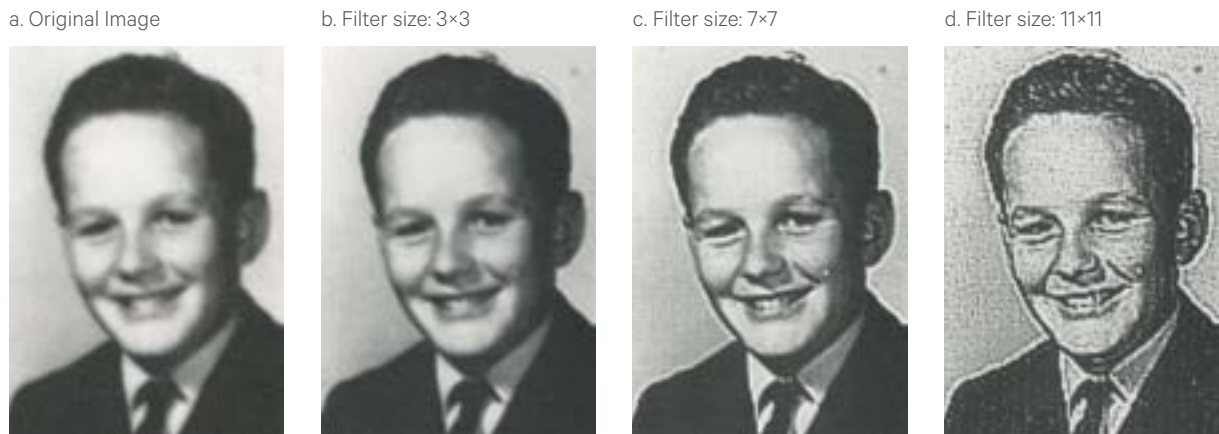
Source: Harrison and Jupp (1990) Plate 4

Various edge detection filters such as the Laplacian operator are described in Section 5.1. As introduced above, an edge enhancement filter can be simply developed by adding an edge detection filter to the original image filter (see Figure 5.18a). An effective edge enhancement filter involves subtraction of the first difference Laplacian from the original (the weightings of this Laplacian required subtraction rather than addition to produce a 'positive' image; see Figure 5.18b).

Niblack (1986) refers to this filter type as being similar to the photographic process of 'unsharp masking'.

Edge enhancement filters can similarly be derived from other edge detection filters, such as directional templates. Chavez and Bauer (1982) developed a method that uses the horizontal first difference to automatically select the best filter size to enhance edges in a particular image.

Edge enhancement transformations may be applied in most image processing systems with user-defined weights in an average filter. As with edge detection filters, thresholding or pre-processing to remove noise is often beneficial to avoid highlighting noise as well as edges.

Figure 5.17 Operation of edge enhancement filter

Differences between the central pixel and its neighbours are exaggerated using a filter that has a large central value with small edge values.

Source: Harrison and Jupp (1990) Figure 80

Figure 5.18 Examples of edge enhancement filters

a. Edge detection filter added to original image creates an edge enhancement filter.

Original				Edges				Edge enhanced		
0	0	0	**+**	-1/9	-1/9	-1/9	**=**	-1/9	-1/9	-1/9
0	1	0		-1/9	8/9	-1/9		-1/9	17/9	-1/9
0	0	0		-1/9	-1/9	-1/9		-1/9	-1/9	-1/9

b. Edge enhancement filter based on Laplacian filter requires subtraction from the original.

Original				Edges				Edge enhanced		
0	0	0	**−**	0	1	0	**=**	0	-1	0
0	1	0		1	-4	1		-1	5	-1
0	0	0		0	1	0		0	-1	0

5.3 Further Information

Castleman (1998) Sections 9, 11 and 12

Jensen (2016) Section 8

Gonzalez and Woods (2018) Section 3

Nixon and Aguado (2002) Section 4

5.4 References

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 8(6), 679–698.

Castleman, K.R. (1998). *Digital Image Processing.* 2nd edn. Prentice-Hall, Inc. 667 pp.

Chavez, P.S., and Bauer, B. (1982). An automatic optimum kernel size selection technique. *Remote Sensing of Environment,* 12, 23–38.

Conte, S.D., and de Boor, C. (1965). *Elementary Numerical Analysis.* McGraw-Hill Kogakusha Ltd, Japan.

Gonzalez, R.C., and Woods, R.E. (2018). *Digital Image Processing.* Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.

Jacobberger, P.A. (1988). Mapping abandoned river channels in Mali through directional filtering of TM data. *Remote Sensing of Environment,* 26, 161–170.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective.* 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Levine, M.D. (1985). *Vision in Man and Machine.* McGraw-Hill Inc. USA.

Moik, J.G. (1980). *Digital Processing of Remotely Sensed Images.* NASA SP–431. Washington, DC. USA.

Niblack, W. (1986). *An Introduction to Digital Image Processing.* Prentice-Hall International. New Jersey.

Nixon, M.S., and Aguado, A.S. (2002). *Feature extraction and image processing for computer vision.* Academic Press, Elsevier. 609 pp. *http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.375.6848&rep=rep1&type=pdf*

Paine, S.H., and Lodwick, G.D. (1989). Edge detection and processing of remotely sensed digital images. *Photogrammetria,* 43, 323–336.

Prewitt, J.M.S. (1970). Object Enhancement and Extraction. In *Picture Processing and Psychopictorics.* (Eds B.S. Lipkin and A. Rosenfeld). Academic Press, New York. pp 75–149.

Roberts, L.G. (1965). Machine perception of three-dimensional solids. In *Optical and Electro-optical Information Processing.* (Ed. J.T. Tippett). MIT Press, Cambridge, Mass. pp 159–197.

Ton, J., Jain, A.K., Enslin, W.R., and Hudson, W.D. (1989). Automatic road identification and labelling in Landsat TM images. *Photogrammetria,* 43, 257–276.

6 Highlighting Surface Variation

Local surface variations are called image 'texture' and can be characterised using a range of statistics (see Section 6.1). A variety of differential filters can also be adapted to accentuate surface variation. Such filters transform imagery to highlight:

- shaded relief (or Sun shadowing)—see Section 6.2;
- exposure to surface winds—see Section 6.3; or
- surface curvature—see Section 6.4.

6.1 Texture (Local Variance)

Image texture has been measured and characterised by a wide variety of methods (Haralick 1979; Levine 1985). Texture transformations use spatial statistics to compute the local standard deviation at each pixel. In this context, the square root of the local variance, is called 'texture' (see Volume 2A—Section 8.2.1). This statistic can be normalised by the mean of the pixel neighbourhood values to produce the 'coefficient of variation'.

A 3×3 filter is used to define the neighbourhood of each pixel. Within this filter region, the texture (T_{ij}) is computed by:

$$T_{ij} = \sqrt{\frac{1}{2n} \times \sum \left(x_{ij} - x_{kj} \right)^2}$$

where

x_{ij} is the central value of a 3×3 window centred at column i and row j;

x_{kl} is the pixel value at column k and row l where $k=i\pm1$ and $l=j\pm1$; and

n is the number of pixels involved in the calculation—usually 8 but sometimes less if there are null pixels within the 3×3 window.

Where $n=8$, this equation becomes:

$$T_{ij} = \sqrt{\frac{1}{16} \times \sum \left(x_{ij} - x_{kj} \right)^2}$$

This texture value can be normalised by $\overline{x}_{ij}$, the mean of the pixels in the filter neighbourhood, to give the Coefficient of Variation (C_{ij}) as:

$$C_{ij} = \frac{T_{ij}}{\overline{x}_{ij}}$$

The maximum texture value is generally difficult to estimate so is usually determined iteratively. (It can be estimated theoretically, using the minimum and maximum values for a given image channel, by computing the texture value of a 3×3 neighbourhood, where the central pixel has the minimum value and the surrounding pixels are equal to the maximum value.)

The texture transformation highlights areas of rapid spatial change, such as edges. In this sense, it is equivalent to the average slope of the image channel 'surface' over the filter region. It has also been used to categorise differences in the spatial structure of image features. For example, at the spatial resolution of a ~50 m pixel size, different forest canopies frequently radiate similar spectral values but can be separated by the difference in their spatial variability or 'texture'. To enhance vegetation texture, this transformation would be best applied to a NIR image channel or to a vegetation index channel (see Section 11 and Volume 3A).

Figure 6.1 Different texture filters

a. Original image   b. Texture   c. Coefficient of variation   d. Edge preserving texture

Source: Harrison and Jupp (1990) Plate 5d to 5f

This spatial analysis is similar to visually considering feature texture in aerial photography interpretation. The coefficient of variation effectively reduces the typically high texture values in bright areas of an image. Pickup and Foran (1987) observed that the coefficient of variation in multi-temporal imagery provides a measure of the effect of plant cover changes on spectral variability within a landscape. As with other spatial statistics, texture results are dependent on both the radiometric and spatial resolutions of an image (see Volume 1B—Section 1). The image texture values can also be used to create an image in which local variance (or standard deviation) is 'constant' (see Excursus 6.1).

As an edge enhancement transformation, texture analysis accentuates boundaries between any spectrally coherent patches (see Figure 6.1). The effect of this enhancement can highlight noise as well as edges. In most cases texture analysis on a smoothed image will reduce this effect.

Edge enhancement effects can be reduced with texture filtering by using adaptive windows (Woodcock and Ryherd, 1989—see Section 3.5). As illustrated in Figure 3.5, the adaptive window technique selects the 'best' sub-window to use within a filter region to derive the filtered value. In texture filtering, if the sub-window is selected on the basis of

minimum variance, the resulting texture value is more likely to relate to a feature patch in the image than to an edge between two features.

Adaptive windows can be used to compute either smoothing or texture measures for an image. The texture results are generally selected from either the:

- minimum variance (based on the mean);
- deviation (associated with the median); or
- entropy (related to the mode)

of values within a sub-window of each filter region. The minimum variance result is illustrated in Figure 6.1, labelled as the 'Edge preserving texture'.

In conjunction with image classification, texture results are best used in a post-classification stage to split up specific, spectrally-defined classes (see Volume 2E). Unlike most classification algorithms, which are pixel specific, texture analysis considers the neighbouring pixel values so can add new information to the classification process. However, if texture channels are used to define the classification, a large number of classes is likely to be generated, as the boundaries themselves will tend to form separate classes. Gordon and Philipson (1986) suggest the use of ratioing texture-enhanced channels to reduce this characteristic edge brightness.

*Each man is always in the middle of the surface of the earth
and under the zenith of his own hemisphere,
and over the centre of the earth.
(Leonardo da Vinci)*

60

Earth Observation: Data, Processing and Applications.   Volume 2: Processing

Excursus 6.1—High Pass Filtering

Source: David Jupp, CSIRO

Image texture values can be used to create an image in which the local variance (or standard deviation) is 'constant'. This type of enhancement has been variously referred to as 'constant variance' (Harris 1977), 'adaptive contrast' (Levine 1985) or 'statistical differencing' (Niblack 1986) and results in even contrast over the whole image. The general transformation is defined as:

$$\frac{x_{ij} - \overline{x}_{ij}}{\sigma_{ij}}$$

where

x_{ij} is the image value at pixel i and line j;
$\overline{x}_{ij}$ is the mean value of the neighbourhood of x_{ij}; and
σ_{ij} is the standard deviation of the neighbourhood of x_{ij}.

Figure 6.2a illustrates the results of applying this transformation to the image shown in Figure 6.1a using a range of filter sizes. To return this result to the original image data range, it can be multiplied by the mean of the local standard deviations, then added back to the local neighbourhood mean ($\overline{x}_{ij}$; see Figure 6.2b). These images illustrate the play off between high pass filtering and local variance equalising. Such differences will be data dependent, and in this example they are small when the high pass filtered and equalised image is added back to the original image. The high pass images lend themselves to further blur and stretching enhancements.

This enhancement is particularly effective for grey-scale imagery with a high dynamic range but little local variation. Medical imagery such as X-rays can be clarified using this technique.

Figure 6.2 Constant variance images

The scanned photograph shown in Figure 6.1a has been transformed using high pass filtering.

a. High pass filtering using equation above with four different filter sizes that are equivalent to 1%, 2.5%, 5% and 10% of the image width (from left to right)

b. High pass filtering using the same filter sizes with rescaling back to the original image data range.

Source: David Jupp, CSIRO

6.2 Relief Shading

As introduced in Volume 1, the term 'Sun shadowing' describes the relief shading produced on the Earth's surface by a particular Sun position. Relief shading is a useful transformation for elevation data as it highlights geomorphological patterns. When elevation data are registered with EO imagery, the relief shading corresponding to the time of image acquisition can be computed to help reduce topographic shading effects in the image.

Relief shading can be computed from relief data such as a Digital Elevation Model (DEM—see Volume 2D) if:

- a reflectance model is assumed for the terrain; and
- the Sun position can be defined in terms of:
  - ◆ altitude (or elevation)—angle above the horizon; and
  - ◆ azimuth (angle clockwise from north).

A range of surface reflectance models have been devised for various purposes (see Volume 1X—Appendix 5). For example, the simple cosine reflectance model defines the radiance (R) of a Lambertian surface (see Volume 1A—Section 5 or Volume 1B—Section 3) as:

$$R = R_n \left(\cos i \right)$$

where

> R is the radiance of a target surface;
> R_n is radiance viewed in a direction perpendicular to the target surface; and
> i is the angle between upwelling radiance and the normal to the surface (see Figure 6.3).

Figure 6.3 Cosine reflectance model
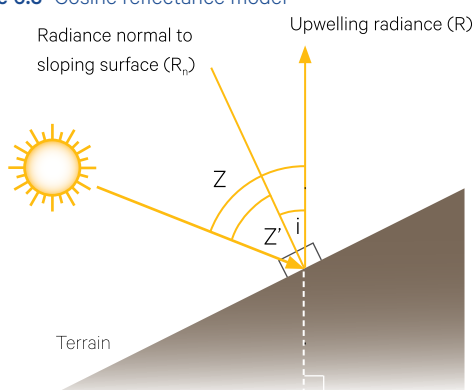
Source: Harrison and Jupp (1990) Figure 83

This effectively models surface radiance for:

- a Sun position that is normal to the surface; and
- a sensor viewing the surface from a position directly overhead.

The modelled radiance can be modified for other Sun positions using the topographic modulation factor:

$$\frac{\max\left(\cos Z', 0 \right)}{\cos\left(Z \right)}$$

where

> Z is the zenith angle of the Sun (the angle from the directly vertical position), and
> Z' is the angle between the Sun vector and the normal to the slope (Ahmad *et al.*, 1989).

As discussed in Section 5.1, the first partial derivatives of an elevation surface can be used to define the direction and magnitude of the slope. The 'downhill' direction of a slope vector is defined as the two-dimensional vector:

$$\begin{bmatrix} -f_x \\ -f_y \end{bmatrix}$$

In a three-dimensional space, a vector can be described in terms of the angles it makes with each of the X, Y and Z axes. This is commonly represented as the direction cosines (that is, the cosines of each of these angles). The vector which is normal (vertically orthogonal) to the slope is described by the direction cosines l, m and n where:

$$l = \frac{p}{\sqrt{1 + p^2 + q^2}}$$

$$m = \frac{q}{\sqrt{1 + p^2 + q^2}}$$

$$n = \frac{1}{\sqrt{1 + p^2 + q^2}}$$

and

$$p = f_x$$

$$q = f_y$$

The Sun vector can similarly be defined as:

$$l_s = \frac{p_s}{\sqrt{1+p_s^2+q_s^2}}$$

$$m_s = \frac{q_s}{\sqrt{1+p_s^2+q_s^2}}$$

$$n_s = \frac{1}{\sqrt{1+p_s^2+q_s^2}}$$

where

$$p_s = \tan Z \times \sin\theta$$
$$q_s = \tan Z \times \cos\theta$$

and

Z is the Sun zenith angle; and
θ is the Sun azimuth angle.

The cosine of the angle between two vectors can be computed as the scalar product of their direction cosines, that is:

$$\cos Z' = \begin{bmatrix} l & m & n \end{bmatrix} \begin{bmatrix} l_s \\ m_s \\ n_s \end{bmatrix}$$

$$= l \times l_s + m \times m_s + n \times n_s$$

The cosine of the angle between the Sun and the 'normal to the slope' vectors (that is Z') can thus be derived as:

$$\frac{p \times p_s + q \times q_s + 1}{\sqrt{1+p^2+q^2} \times \sqrt{1+p_s^2+q_s^2}}$$

Relief shading (or Sun shadowing) is then computed as:

$$\frac{R \times \cos Z'}{\cos Z}$$

This gives a relative value between 0 and 1 for 0°≤ Z≤ 90° and Z'> 0, or 0 for any other angles.

Some image processing systems compute relief shading, slope and aspect for an elevation channel using the cosine reflectance model for the terrain (see Figure 6.4). A 5×5 filter (detailed in Volume 2X—Appendix 5.2.2) may also be used to estimate local derivatives (see Section 5.1). The first partial derivative results can be used to define the principal gradient for the filter region at each pixel, that is, the vector that is perpendicular to the dominant direction of contour lines within the region. The slope of the gradient corresponds to the length of this vector and may be computed as:

$$\sqrt{f_x \times f_x + f_y \times f_y}$$

The direction of the gradient vector, or aspect, can also be computed from the partial derivatives as:

$$\tan^{-1}\left(\frac{-f_x}{-f_y}\right)$$

Relative relief shading can be computed as detailed above with f_x and f_y of the specified image channel defining the slope vector and user-selected Sun azimuth and zenith angles[4] defining the Sun vector. To compute relief shading parameters, the geometry of the input image needs to be defined in terms of scaling in the:

- X dimension (pixel width);
- Y dimension (pixel depth); and
- Z dimension (height scaling—the relationship between actual height and the values coded in the elevation channel); as well as the
- orientation relative to north.

Figure 6.4 Parameters derived from elevation

Based on a digitial elevation model, the relief shading, slope and aspect can be computed as image channels.

a. Elevation | b. Relief shading | c. Slope | d. Aspect

Source: Harrison and Jupp (1990) Plate 13

4   The zenith angle is 90° minus the Sun altitude angle.

An atmospheric correction factor may also be specified to account for absorption and backscatter, which would reduce the measurement of surface radiance in conditions of haze or cloud.

Insolation, 'that amount of solar radiation received on a given surface in a given time period' (Landsberg and Sands, 2011), is a significant parameter in environmental physics and ecology (see Figure 6.5). Relief shading can serve as a surrogate measure for insolation to indicate the radiance that would be measured from a surface with uniform land cover for given parameters of Sun position and atmospheric transmittance conditions.

In conjunction with EO-based analyses, topographic data allow investigations of the relationship between spectral data and surface morphology (see Volume 2D). For example, registration between an image-based land cover classification and elevation data allows quantification of the occurrence and co-occurrence of classes at different topographic locations. Tabular summaries of such comparisons may be produced automatically in most image processing systems (see Volume 2E). Alternatively, land cover categories may be stratified by topography or morphology. The derivation of slope information also allows elevation categories to be stratified by slope classes for various management purposes.

Figure 6.5 Total insolation

Maximum insolation occurs when the surface is perpendicular to incoming radiation. For a flat surface, insolation depends on Sun position and atmospheric transmittance.

Maximum insolation

Variable insolation

No insolation
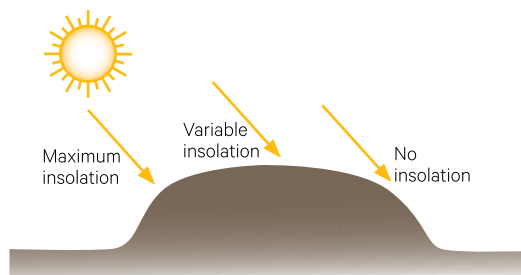
Source: Harrison and Jupp (1990) Figure 84

The best way to study Mars is with two hands, eyes and ears of a geologist, first at a moon orbiting Mars … and then on the surface.
(Buzz Aldrin)

6.3 Exposure (Directional)

A special implementation of the differential filter is the exposure transformation. This transformation indicates the relative exposure for each pixel in an image channel 'surface' to a user-defined direction as illustrated in Figure 6.6. The 'exposure' result is similar to relief shading with the 'Sun position' effectively on the horizon.

This process is relevant to shallow water mapping applications where parameters for water depth and prevailing wind direction can be used to derive a value indicating the degree of exposure of the sea floor at each pixel to that wind direction. In shallow water mapping exercises visible blue or green channels from EO imagery can be used as a surrogate for water depth (Jupp *et al.*, 1985). Directional parameters are typically expressed either in degrees, as a compass bearing clockwise from north, or in terms of X and Y components using polar coordinates, as illustrated in Figure 6.7 (see also Volume 2B).

Figure 6.6 Exposure transformation

a. Perspective view

Wind direction

x

Surface gradient at x
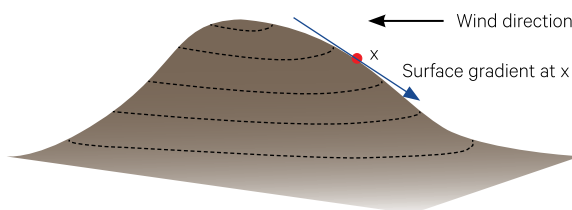
b. Plan view (E=exposure value)

Wind direction

Surface gradient at x

E

E=0

x

E=max

E

E
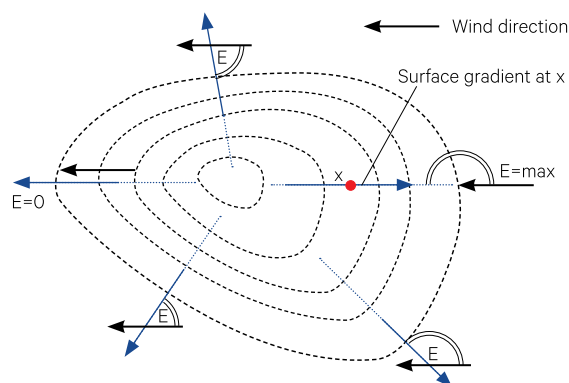
Source: Harrison and Jupp (1990) Figure 81

An exposure transformation uses two smoothing differential filters over the neighbourhood of each pixel. Typical weights for these filters are shown in Figure 6.8, with both filters having a divisor of 12. The exposure at each pixel is then derived from the equation:

$$E = -\left(\sin\theta \times f_x + \cos\theta \times f_y\right)$$

where

> $\sin\theta$ represents the X component of the compass direction;
>
> f_x is the first partial derivative in the X direction;
>
> $\cos\theta$ represents the Y component of the compass direction;
>
> f_y is the first partial derivative in the Y direction; and
>
> f is the image data, such as water depth (assumed to be oriented to north).

(Note: the form of this equation differs from the gradient direction equation given in Section 6.2. This difference is due to the way the compass direction is defined. Here the direction is given as clockwise from north whereas in the context of relief shading the direction is effectively measured as anti-clockwise from east.)

Figure 6.7 Representing compass direction as X and Y components

Since $\sin\theta$ = a/c and $\cos\theta$ = b/c, for c=1, $\sin\theta$ = a (X component) and $\cos\theta$ =b (Y component), where θ measures degrees clockwise from north.
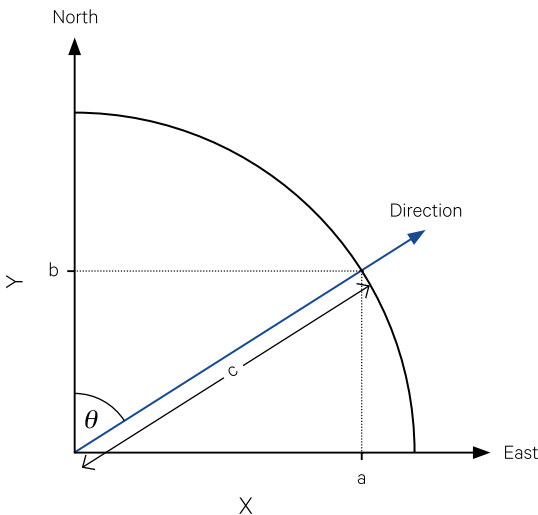
A non-linear scaling of the resulting exposure value E may be used to increase contrast in low image values, in which case, the final exposure image values are derived from the product of the scaled exposure component and the input image value at each pixel. For best results, images should be destriped and despiked before computing the exposure transformation as differential filters will enhance image noise such as striping patterns or spikes (see Sections 2.4 and 4.5). Some algorithms incorporate a dark value to indicate the data level corresponding to atmospheric noise (see Section 10.2.1 for more details about dark value selection). An expected maximum value may also be requested for rescaling purposes. This is best determined iteratively, but the 99% maximum is usually satisfactory. The resulting exposure channel may be displayed as a grey-scale and contrast-enhanced, or used to modify the brightness values of a colour image to indicate surface relief (see Section 8.5).

As illustrated in Figure 6.9, the exposure transformation has proved to be valuable for mapping sub-surface morphology in coral reef studies. In this example, comparable exposure images have been derived from both Landsat MSS imagery (green band) and a sparse grid of water depth samples. In the exposure channel, bright areas indicate a higher local slope and aspect to the selected wind direction, and in the prevailing wind direction these areas indicate greater exposure. Darker areas infer sheltered positions relative to that wind direction. Thus, by using the relevant range of wind directions, this transformation allows a set of 'impact images' to be iteratively compiled for any suitably imaged location.

Figure 6.8 Filter weights for exposure transformation

These filters both have divisors of 12 to retain image scaling.

f_x

-1	0	1
-4	0	4
-1	0	1

f_y

1	4	1
0	0	0
-1	-4	-1

Figure 6.9 Exposure (directional) filtering

Comparable images have been derived for John Brewer Reef, Queensland, by directional filtering of Landsat MSS imagery and depth data. The resulting images highlight reef morphology and show exposure to the prevailing wind direction (southeast).

a. Transect samples of actual water depth formed a sparse grid of points when registered with Landsat MSS imagery. (Note: null pixels are shown as black.) This sparse dataset was interpolated (see Section 4.6) to create a continuous depth image before exposure (directional) filtering.

Real depth data   Interpolated depth data   Exposure on interpolated depth

b. A Landsat MSS4 (green band) image for John Brewer Reef was despiked and smoothed before directional filtering to show exposure to the prevailing wind direction. An exposure image derived from EO data offers an economical approximation to sourcing real depth data.

Landsat MSS4 image   Preprocessed MSS4 image   Exposure on preprocessed MSS4

Source: Harrison and Jupp (1990) Plate 6

The ability to predict the exposure of sub-surface relief to prevailing weather conditions provides valuable information to coastal zone managers in formulating management priorities for different sites. Spatial variations in the distribution of plant and animal species are usually strongly influenced by the wind and wave energy regimes that are operative in the littoral and sub-littoral zones (see Volume 3B).

As a general directional filter, the exposure transformation can be applied to topographic surfaces, magnetics and gravity data, or any data where a directional spatial pattern can be observed. The resulting channel is similar to relief shading (see Section 6.2) but does not produce any 'black' shadow areas on slopes away from the 'Sun'. The exposure transformation has also been used to enhance nighttime thermal imagery for structural interpretation.

Sometimes I enjoy just photographing the surface
because I think it can be as revealing as going to the heart of the matter.
(Annie Leibovitz)

6.4 Surface Curvature (Surface Shape)

The use of differential calculus to quantify the changes in curvature of a two-dimensional surface, such as a DEM channel, was introduced in Section 5.1, with:

- the first derivative representing the rate of change, or slope, of the surface; and

- the second derivative representing the change in slope, or curvature.

In two-dimensional data, these changes are expressed as partial derivatives, that is, two derivative values are computed separately indicating changes in the two dimensions. These second order statistics are invariant to rotation and slope, and describe intrinsic curvature properties of the data surface (Smith, 1983). Such mathematical analyses may be applied to image data, using filtering techniques with appropriately defined weights, to determine the curvature in both across line (*X*) and down column (*Y*) directions for each pixel in the image.

As detailed in Section 5.1, four second-order partial derivatives can be computed for two-dimensional data to quantify the changes in surface curvature:

$$f_{xx} = \frac{\partial^2 f}{\partial X^2}$$

$$f_{xy} = \frac{\partial^2 f}{\partial X \partial Y}$$

$$f_{yx} = \frac{\partial^2 f}{\partial Y \partial X}$$

$$f_{yy} = \frac{\partial^2 f}{\partial Y^2}$$

These derivatives can be treated as a matrix to determine the direction and magnitude of maximum curvature within the region. When such results are derived from elevation data, they indicate the relative size and orientation of the major landform features within an image filter region.

Volume 2X—Appendix 6 describes matrix algebra techniques which may be used to convert the 2×2 matrix of second partial derivatives of:

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

to the form:

$$\begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$$

where

> k_1 is called the first eigenvalue and indicates the magnitude of a special vector in the direction of maximum data variation, called the first eigenvector; and

> k_2 is the second eigenvalue, which indicates the magnitude of the second eigenvector, this being defined as orthogonal to the first eigenvector.

In terms of elevation data, the first eigenvector of the second partial derivative matrix is aligned with the direction of curvature of the major landform feature (within a filter region at an image pixel), while the corresponding eigenvalue indicates the relative magnitude of curvature of the feature. The second eigenvalue then indicates the extent of curvature in the direction at 90° to the first eigenvalue.

Such results can then be used to categorise different landform features in elevation data, or in fact the surface curvature of any two-dimensional dataset such as an image channel. In terms of image processing then, the curvature values of each pixel describe the surface changes within the image sub-area defined by the filter size used to derive them.

It is appropriate to compute the surface curvature of an image channel using a 7×7 filter region (see Volume 2X—Appendix 5.2.3). As with relief shading (see Section 6.2), the input image geometry needs to be specified in terms of pixel width (*X*), pixel depth (*Y*) and the height scaling used to convert actual height to image values (*Z*). Using appropriate filters weights (see Volume 2X—Appendix 5.2.3 for examples), the following statistics could be computed as new image channels (see Figure 6.10):

- the first eigenvalue of the second derivative matrix (also referred to as λ_1) indicates the extent of surface shape in the direction of maximum shape change;

- the second eigenvalue of the second derivative matrix (also referred to as λ_2) indicates the extent of the surface shape change in the direction that is orthogonal to the maximum change direction in the 7×7 filter region;

- the Laplacian: $\frac{\lambda_1 + \lambda_2}{2}$ indicates the average extent or magnitude of the surface curvature within the filter region; and

- the Discriminant: $\frac{\lambda_1 - \lambda_2}{2}$ summarises the directionality of surface curvature within the filter region (such as a ridge or gully).

Figure 6.10 Parameters related to surface curvature

Using the digitial elevation model shown in Figure 6.4, surface curvature parameters can be computed as image channels. The Laplacian is the average of the first and second eigenvalues, while the discriminant is their difference. To compute this example, a 7×7 filter region was used with the filter weights specified in Volume 2X—Appendix 5.2.3.

a. First eigenvalue

b. Second eigenvalue

c. Laplacian

d. Discriminant
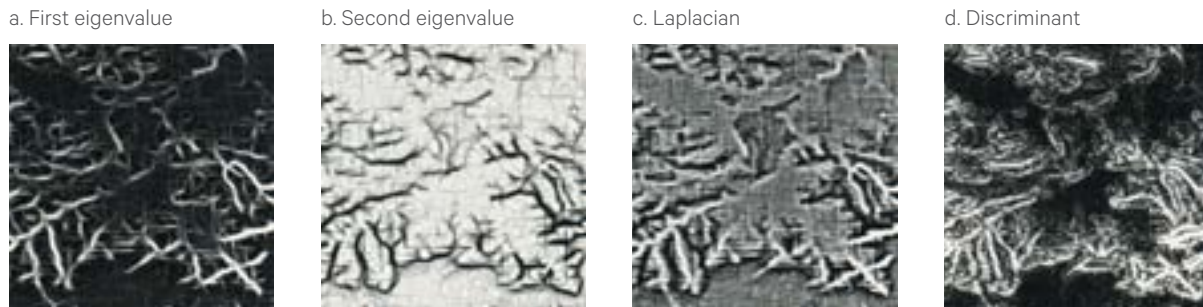
Source: Harrison and Jupp (1990) Plate 15

The scaling of these statistics depends on the ranges specified for the expected minimum and maximum values. As with many transformations, the actual range of transformed values that will be relevant to an image can not be predicted so these range values are best determined iteratively.

When applied to elevation data these four parameters clearly differentiate the major landform features of ridges, gullies, hills, basins, saddles and flat ground within each 7×7 pixel region. The relationship between these landform features and the output image values are given in Table 6.1.

These surface curvature parameter channels could be used to form a colour composite. If the first eigenvalue is displayed as blue, the Laplacian as green and the Discriminant as red, ridges would show as well-connected red lines, gullies as pink lines, undulating ground as aqua and flat ground as green (see Figure 6.11). It should be noted that the category 'flat' only applies to surface curvature within the filter region, not surface orientation, so that a uniformly sloping area with no change in surface shape within the filter would be described as 'flat'. A composite display of slope as blue, the Laplacian as green, and Discriminant as red provides enhanced contrast in the undulating ground. In geophysics, the Laplacian statistic is also called the second vertical derivative.

Table 6.1 Curvature image categories

λ_1 is first eigenvalue and λ_2 is second eigenvalue.

Surface Curvature Category	Curvature Image Values			
	λ_1	λ_2	Laplacian	Discriminant
Flat ground	~0	~0	~0	small
Basin (bowl shape)	>>0	>>0	>>0	small
Hill (cap shape)	<<0	<<0	<<0	small
Saddle	>>0	<<0	~0	large
Gully line	>>0	~0	>>0	large
Ridge line	~0	<<0	<<0	large

These statistics may be derived from either elevation data or appropriate EO imagery (Li *et al.*, 1986). Various methods have been proposed to isolate topographic shading (that is, effectively Sun shadowing) information from EO imagery (Ahmad *et al.*, 1989). Where land cover variation is relatively small, topographic shading is the major factor causing changes in image brightness and can be reasonably estimated as the first principal component (PC1; see Section 9). As with other analyses, Sun shadowing imagery that relates to a high Sun angle provides better discrimination of the landform features than is possible with low Sun angle imagery.

While the results of curvature analysis on Sun shadowing imagery will necessarily differ from those of elevation data, some interesting landform features are highlighted. At the very least, these results can be used to stratify the image into 'flat' and non-flat areas. Such stratifications then allow the analysis of land use and land cover features in the flat areas without influencing class definition by topographic brightness variations (see Volume 2E).

Figure 6.11 Curvature composite

This curvature composite image is derived from the digital elevation model shown in Figure 6.4 and used in Figure 6.10, here shown in pseudo-colour with blue indicating low elevation and red showing high elevation. The curvature composite image displays the first eigenvalue as blue, the Laplacian as green and the discriminant as red to highlight major landform categories.

a. Pseudo-coloured elevation

b. Curvature composite
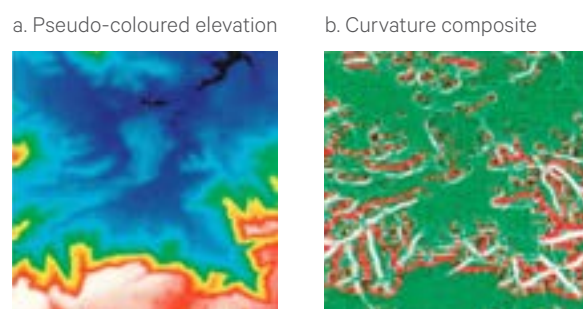
Source: Harrison and Jupp (1990) Plate 15

6.5 Further Information

Texture

Jensen (2016) Section 8

6.6 References

Ahmad, W., Jupp, D.L.B., and Nunez, M. (1989). Land cover mapping in a rugged terrain area using Landsat MSS data. *Int. J. Remote Sensing,* 13(4), 673–683. *https://doi.org/10.1080/01431169208904145*

Gordon, D.K., and Philipson, W.R. (1986). A texture-enhancement procedure for separating orchard from forest in Thematic Mapper data. *International Journal of Remote Sensing,* 7, 301–304.

Haralick, R.M. (1979). Statistical and Structural Approaches to Texture. *Proceedings of IEEE,* 67, 786–804.

Harris, J.L. Sr. (1977). Constant Variance Enhancement: A Digital Processing Technique. *Applied Optics,* 16, 1268–1271. *https://doi.org/10.1364/AO.16.001268*

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective.* 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Jupp, D.L.B., Mayo, K.K., Kuchler, D.A., Van R. Classen, D., Kenchington, R.A., and Guerin, P.R. (1985). Remote sensing for planning and managing the Great Barrier Reef of Australia. *Photogrammetria,* 40, 21–42.

Landsberg, J.J., and Sands, P.J. (2011). Weather and Energy Balance. Ch 2 in *Terrestrial Ecology Volume 4: Physiological Ecology of Forest Production.* Elsevier. doi: *10.1016/B978-0-12-374460-9.00002-0*

Levine, M.D. (1985). *Vision in Man and Machine.* McGraw-Hill Inc. USA.

Li, R., Jupp, D.L.B., Harrison,B.A., and Ahmad, W. (1986). Extraction of landform information from remotely sensed data: an initial study. *Proceedings of Beijing International Symposium on Remote Sensing,* Beijing, China, 18–22 November, 67–73.

Niblack, W. (1986). *An Introduction to Digital Image Processing.* Prentice-Hall International. New Jersey.

Pickup, G., and Foran, B.D. (1987). The use of spectral and spatial variability to monitor cover changes on inert landscapes. *Remote Sensing of Environment,* 23, 351–363.

Smith, G.B. (1983). Shape from shading: an assessment. *Proc. NASA Symposium on Mathematical Pattern Recognition and Image Analysis,* June 1–3, 1983. Johnson Space Centre, Houston, Texas, 543–576.

Woodcock, C.E., and Ryherd, S.L. (1989). Generation of texture images using adaptive windows. *Tech. Papers, 55th Annual Meeting ASPRS,* 2, 11–22.

Linear Operations
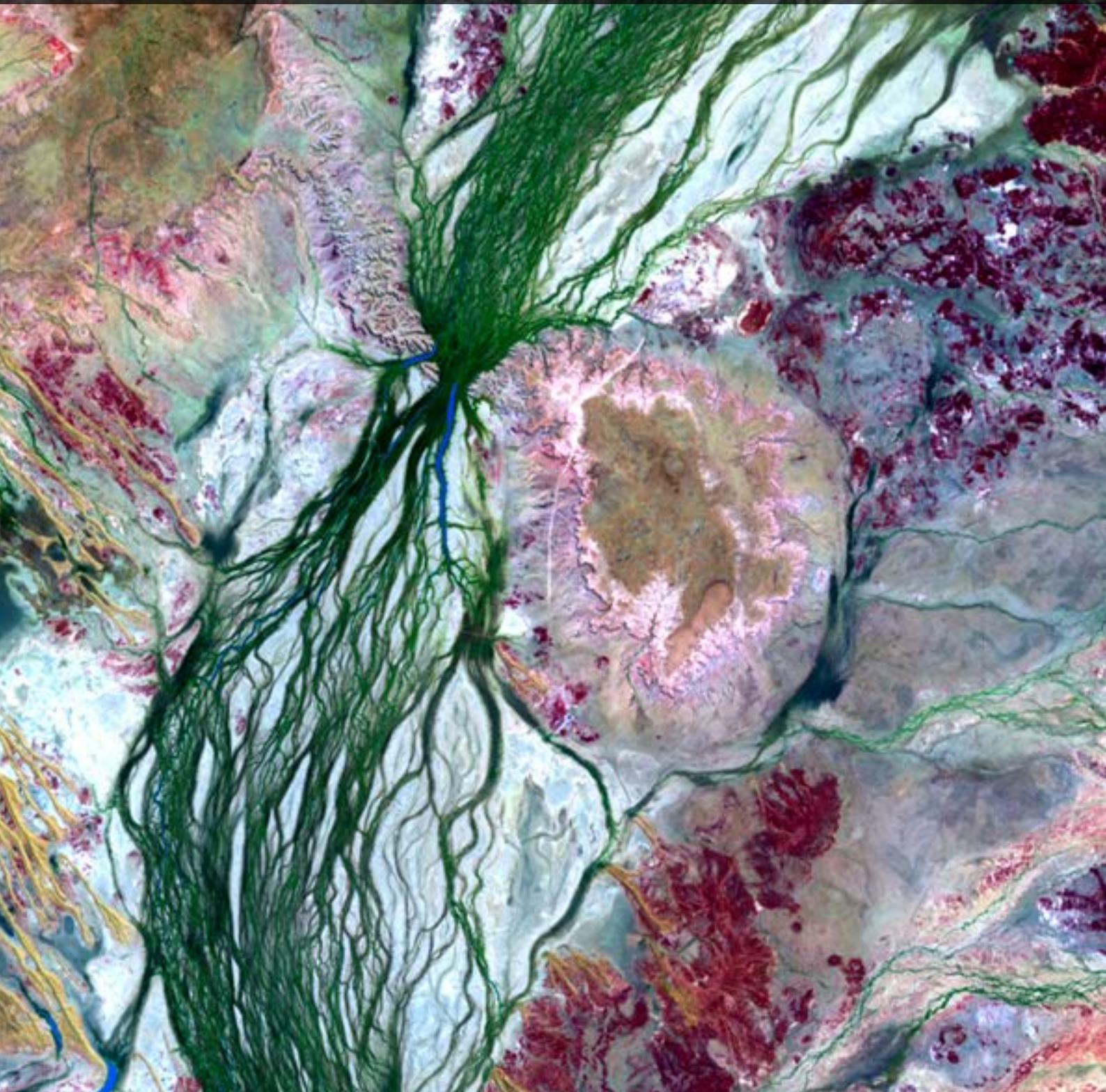
The following three sections are concerned with techniques that apply linear operations to one or more image channels. Linear operations rely on matrix algebra and can implement simple operations (such as simple rescaling or inversion of individual channels) or powerful multidimensional manipulation of image data. Non-linear operations involve multiplication and division of channel values (see Section 1.4).

Most image processing systems offer a wide range of multi-channel transformations. Some of the most commonly available linear transformations involving multiple channels include:

- modifying image brightness (see Section 7);
- converting image colour space (see Section 8); and
- Principal Components Analysis (PCA; see Section 9);

Some image processing systems provide flexible 'band maths' options, whereby a sequence of linear and non-linear operations can be applied to selected sets of image channels. Such options offer significant flexibility, but need to be used intelligently with EO imagery to ensure that the transformed results are meaningful.

Contents

7 Affine Transformations

An affine transformation is one that preserves points, straight lines and planes between the input and output datasets (see Volume 2B—Section 3.2.1). This linear transformation is defined by a multidimensional matrix plus an offset, or shift, and can be used to implement many image processing operations. Such techniques rely on matrix algebra, which is described in Volume 2X—Appendix 6.

In this section, we initially review matrix notation in Excursus 7.1, then Section 7.1 introduces the general affine transformation. Some simple uses of the affine transformation with EO imagery are described in Excursus 7.2, including:

- adding and subtracting image channels;
- averaging channels;
- inverting and rescaling channels;
- processing log-transformed channels; and
- implementing regression models.

These linear transformations can also be used for removing spatial 'noise' (see Section 7.2).

Affine transformations are useful for a range of image processing operations (see Section 7.3), including converting between colour spaces (see Section 8). Another commonly available linear transformation is Principal Components Analysis (PCA; see Section 9), possibly with the option to define selected transformation parameters and/or to rotate the transformed channels back to their original data space. A specific channel rotation operation, which is used to enhance vegetation greenness in EO imagery, is described in Section 11.1.3.

Life is a linear equation in which you can't cross multiply!
If you think you can do it, you can do it.
If you think you can't do it, you can't do it.
It's a simple formula!
(Isrealmore Ayivor)

Background image: Landsat-8 OLI image of Cubbie Station in southern Queensland, which is the largest irrigation property in the southern hemisphere.
Source: Craig Shephard, DSITI

Excursus 7.1—Review of Matrix Notation

Matrix operations allow new image channels to be defined as some linear combination of the original channels. For example, to compute the sum and difference of two channels, X and Y, we would use the equations:

$$X' = X + Y \text{ (or } X' = 1 \times X + 1 \times Y)$$

$$Y' = X - Y \text{ (or } Y' = 1 \times X + -1 \times Y)$$

In matrix notation, we assume that each new X' or Y' value can be computed as some weighted sum of the original X and Y values. This allows us to define a table of weighting values. For example, the general transformation equations:

$$X' = a \times X + b \times Y$$
$$Y' = c \times X + d \times Y$$

could be represented by their coefficients as summarised in Table 7.1. This tabulation is more generally represented as the matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Similarly, the matrix required for the initial sum and difference equations given above would be:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

The general matrix operation in Table 7.1 is usually described as:

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} X \\ Y \end{bmatrix}$$

where X, Y, X', Y', a, b, c, and d are defined above.

Table 7.1 Matrix coefficients

Output Value	Coefficients for Input Values	
	X	Y
X'	a	b
Y'	c	d

Such matrices can be used to transform imagery by computing new channels X' and/or Y' for each image pixel using its values in input channels X and Y. However, image data spaces usually have more than the two dimensions (channels) used in this example. For example, for a four channel image the general matrix operation is defined using a square four-dimensional matrix as shown in Figure 7.1.

These linear algebraic operations allow very flexible manipulation of the image data space. In algebraic terms, the operation is simply a linear combination of the image channel values for each pixel. However in geometric terms, a wide range of sophisticated manipulations of the data space may be performed. Guidelines for defining matrix coefficients for general reflection, rescaling, skewing and rotation operators on the image data space are detailed in Volume 2B—Section 3.2.1 and Volume 2X—Appendix 6.

For example, a simple geometric operation may be to reflect the data space about the line $Y = X$ as shown in Figure 7.2. In terms of coordinates referenced by the X' and Y' axes, the geometric operation can be summarised algebraically by the two simple equations:

$$X' = Y \text{ (or } X' = 0 \times X + 1 \times Y)$$

$$Y' = X \text{ (or } Y' = 1 \times X + 0 \times Y)$$

These equation can be represented by the matrix:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Figure 7.1 General matrix operation

$$\begin{bmatrix} output\ channel\ 1 \\ output\ channel\ 2 \\ output\ channel\ 3 \\ output\ channel\ 4 \end{bmatrix} = \begin{bmatrix} A & B & C & D \\ E & F & G & H \\ I & J & K & L \\ M & N & O & P \end{bmatrix} \times \begin{bmatrix} input\ channel\ 1 \\ input\ channel\ 2 \\ input\ channel\ 3 \\ input\ channel\ 4 \end{bmatrix}$$

or

output channel 1 = A × input channel 1 + B × input channel 2 + C × input channel 3 + D × input channel 4

output channel 2 = E × input channel 1 + F × input channel 2 + G × input channel 3 + H × input channel 4

output channel 3 = I × input channel 1 + J × input channel 2 + K × input channel 3 + L × input channel 4

output channel 4 = M × input channel 1 + N × input channel 2 + O × input channel 3 + P × input channel 4

Figure 7.2 Reflecting image axes about the line *Y* = *X*

Simple matrix operations can be used to modify the geometry of an image data value space.

a. Original axes
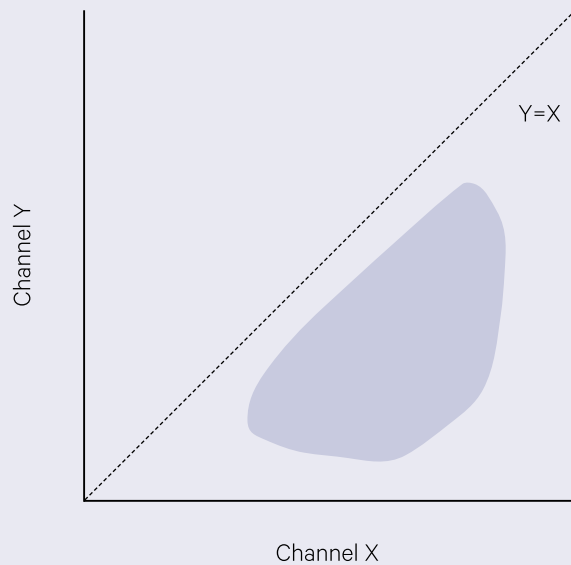
b. Reflected axes
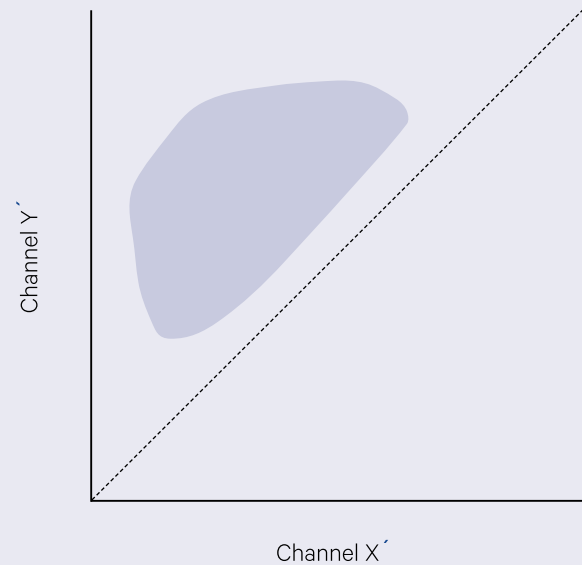
Source: Harrison and Jupp (1990) Figure 85

7.1 General Affine Transformation

Derivation of the linear transformation is introduced in Excursus 7.1 above. This operation allows a new channel to be created as a linear combination (or weighted sum) of the original image channels. Affine transformations are commonly used to define vegetation or soil brightness indices in EO datasets (see Section 11).

An affine transformation adds an offset value to the result of a linear transformation to form the matrix equation:

$$\begin{bmatrix} output \\ image \end{bmatrix} = \begin{bmatrix} transformation \\ matrix \end{bmatrix} \times \begin{bmatrix} input \\ image \end{bmatrix} + \begin{bmatrix} offset \end{bmatrix}$$

For a single channel image, this transformation computes the following for each pixel:

$$output\ value = matrix\ coefficient \times input\ value$$
$$+ offset\ value$$

In image processing, the offset allows the transformed values to be shifted back into the image data range. If a transformation is defined where one channel has a coefficient of one and all other coefficients are zero, the offset can be used to add or subtract a constant value to or from each pixel value in an image channel.

Weighted averaging of multiple channels in an image can be easily implemented using an affine transformation (see also Volume 2X—Appendix 6). For example, the equation:

$$output\ channel = a \times input\ channel\ 1$$
$$+ b \times input\ channel\ 2$$
$$+ c \times input\ channel\ 3$$

can be represented by the matrix equation:

$$\begin{bmatrix} output \\ channel \end{bmatrix} = \begin{bmatrix} a & b & c \end{bmatrix} \times \begin{bmatrix} input\ channel\ 1 \\ input\ channel\ 2 \\ input\ channel\ 3 \end{bmatrix}$$

The combined use of affine transformation, logarithmic or exponential transformation (see Section 2.2) and channel ratios (see Section 10) can produce most of the indices commonly encountered in EO literature (see Excursus 7.2).

Excursus 7.2—Using Affine Transformations

Adding and subtracting channels

An affine transformation can be used to perform simple channel arithmetic, such as addition or subtraction. (Addition and differencing of all channels in pairs of images are discussed in Volume 2D). These arithmetic operations are required to produce indices of particular image features. For example, a simple vegetation greenness index could be based on the difference between pixel values in NIR and red channels of an image (see Section 11):

$$Index\ value = NIR\ value - red\ value$$

This index could be computed using the affine transformation:

$$Index = \left(1 \times NIR\right) + \left(-1 \times red\right)$$

or

$$\begin{bmatrix} index \end{bmatrix} = \begin{bmatrix} 1 & -1 \end{bmatrix} \times \begin{bmatrix} NIR \\ red \end{bmatrix}$$

with an offset equal to half the channel data range[5].

The popular Normalised Difference Vegetation Index (NDVI) involves ratioing this difference by the sum of the NIR and red channels:

$$NDVI = \frac{NIR - red}{NIR + red}$$

or

$$NDVI = \frac{\left(1 \times NIR\right) + \left(-1 \times red\right)}{\left(1 \times NIR\right) + \left(1 \times red\right)}$$

This can be implemented by computing the sum and difference of the two channels by affine transformation then ratioing the results (see Section 11.1.2). The affine transformation matrix required to compute this sum and difference would be:

$$\begin{bmatrix} difference \\ sum \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} NIR \\ red \end{bmatrix}$$

with offsets equal to half the data range for *difference* and 0 for *sum*[6]. The NDVI could then be computed as a ratio of the output difference channel (as the numerator) and the sum channel (as the denominator; see Section 11.1.2). NDVI has demonstrated effective and usable correlations with plant greenness metrics, such as biomass or leaf area index (see Volume 3A). It can saturate, however, in vegetation with high greenness, such as irrigated sports fields (Tony Sparks, *pers. comm.*).

Various channel combinations have been suggested to enhance and summarise image information relating to vegetation condition or soil brightness, such as the 'Tasselled Cap' transformations (see Section 11.1.3). Other specific transformations relating to vegetative cover and condition are considered in Section 11 and Volume 3A.

Averaging channels

The average of multiple channels may be required to summarise specific pattern information in an image, such as edges (see Section 5.1.2). For a three channel image, the average would be computed as:

$$average = \frac{channel\ a + channel\ b + channel\ c}{3}$$

and the required affine transformation matrix would be:

$$\begin{bmatrix} average \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \times \begin{bmatrix} channel\ a \\ channel\ b \\ channel\ c \end{bmatrix}$$

Inverting and rescaling channels

Channel data ranges can also be inverted and linearly rescaled using appropriate coefficients in affine transformations. For example, to invert the range of an image channel, for each pixel we can compute:

$$inverted\ channel = \left(channel\ value \times -1\right)$$
$$+ channel\ data\ range$$

This can be implemented as an affine transformation using -1 as the coefficient for the channel to be inverted plus a channel offset. For example, to invert channel *a* in a three channel image we would use the matrix equation shown in Figure 7.3. This transformation would only change values in channel a, leaving the values in channel b and channel c the same.

---

5   In many image processing systems, one image value is reserved to indicate 'null' pixels (see Volume 2A—Section 1.2), which effectively reduces the data range by one. With byte data, for example, the null value is often value 0 or 255, and the data range reduces to 255 instead of 256 values. In this case the offset would equal 127 rather than 128.

6   In many image processing systems, value 0 or 255 may be reserved in byte resolution imagery to indicate 'null' pixels, in which case the value 127 would be used as the offset.

Figure 7.3 Matrix equation for channel inversion

$$
\begin{bmatrix} \text{inverted channel a} \\ \text{channel b} \\ \text{channel c} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \text{channel a} \\ \text{channel b} \\ \text{channel c} \end{bmatrix} + \begin{bmatrix} \text{offset} \\ 0 \\ 0 \end{bmatrix}
$$

Linear transformations on log-transformed channels

When an affine transformation is applied to log-transformed imagery (see Sections 1.2.3 and 2.2), channel addition and subtraction respectively represent multiplication and division of the original channels. For example,

$$
\log\left(\text{channel 1} \times \text{channel 2}\right) = \log\left(\text{channel 1}\right) + \log\left(\text{channel 2}\right)
$$

$$
\log\left(\frac{\text{channel 2}}{\text{channel 3}}\right) = \log\left(\text{channel 2}\right) - \log\left(\text{channel 3}\right)
$$

$$
\log\left(\left(\text{channel 4}\right)^3\right) = 3 \times \log\left(\text{channel 4}\right)
$$

(In the case of channel ratios, dark values are implemented in this sequence during the logarithmic transformation; see Sections 2.2 and 10.2.1).

A complex non-linear combination such as:

$$
\frac{\left(\text{channel 1}\right) \times \left(\text{channel 2}\right)}{\left(\text{channel 3}\right) \times \left(\text{channel 4}\right)^3}
$$

would then be a linear concatenation of the previous operations. In terms of matrix coefficients, these transformations (with offsets equal to 0) would be represented as shown in Figure 7.4.

The exponentiation option, which should be available in most image processing systems (see Section 2.2), could then be used to convert the resulting channels back to a linear scaling.

Implementing regression models

Where a known linear relationship exists between image values and some ground parameter, such as crop yield, and can be expressed as a regression of the form:

$$
\text{parameter} = a_i \times x_i + b_i
$$

where

x_i is the image value in channel i;

a_i is the model coefficient for channel i; and

b_i is the model constant for channel i.

then a new channel, which represents a scaled value for the ground parameter, can be computed using the affine transformation:

$$
\begin{bmatrix} \text{parameter} \end{bmatrix} = \begin{bmatrix} a_i \end{bmatrix} \times \begin{bmatrix} \text{channel } i \end{bmatrix} + \begin{bmatrix} b_i \end{bmatrix}
$$

The regression technique described in Sections 1.3.4 and 7.2.1 can be used to determine the relationship between two image channels. One channel can then be rescaled to match the other using the regression coefficients and offsets as exampled in Section 7.2.1 (see also Volume 2D).

Figure 7.4 Example matrix equation

$$
\begin{bmatrix} \log\left(ch1 \times ch2\right) \\ \log\left(ch2 / ch3\right) \\ \log\left(ch4^3\right) \\ \log\left(\dfrac{ch1 \times ch2}{ch3 \times ch4^3}\right) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 3 \\ 1 & 1 & -1 & -3 \end{bmatrix} \times \begin{bmatrix} \log ch1 \\ \log ch2 \\ \log ch3 \\ \log ch4 \end{bmatrix}
$$

7.2 Removing Spatial 'Noise'

Spatial 'noise' can occur in EO imagery for a number of reasons, such as atmospheric haze, specular reflection or uneven illumination (due to topography or uneven lighting). These spatial variations are not constant over the image, so we cannot correct for them by simply subtracting a constant 'noise' level from each pixel.

Methods that may be used to adjust for these spatial variations include weighted channel subtraction (see Excursus 7.2) or channel ratioing (see Section 10). The choice of correction technique depends on the nature of the model that describes the interaction between surface reflectance and the image noise. Where the spatial noise is related to ground reflectance by a multiplicative model (that is the detected noise level is affected by target reflectance, such as with topographic shading) image ratios should be used. The models for atmospheric noise and sunglint (specular reflection from water) are linearly related to surface reflectance (that is the noise level is independent of target reflectance) so an affine transformation can be used to correct for these components in the image radiance values. That is:

$$x_{ij} = b_{ij} + d_{ij}$$

where

x_{ij} is image radiance value at pixel i in channel j ;
b_{ij} is the reflectance component of pixel i in channel j ; and
d_{ij} is the atmospheric noise and/or sensor calibration offset component.

7.2.1 Using a reference channel

This approach uses one image channel that contains the pattern as a 'reference' channel (in the case of scan-digitised imagery, this should always be specifically produced using a blank target) and requires that the reference channel be 'sacrificed' in any subsequent processing. Thus the reference channel is best selected from channels that would not normally be used. Assuming some constants α_j and β_j:

$$d_{ij} = \alpha_j \times d_{ir} + \beta_j$$
$$x_{ij} - \left(\alpha_j \times d_{ir} + \beta_j \right) = b_{ij} - \alpha_j \times b_{ir}$$

where

when $j=r$, $x_{ir} = d_{ir}$ is the image radiance value at pixel i in the reference channel r ; and
b_{ir} is the reflectance component of pixel i in the reference channel r. This reference channel is usually chosen so that this component is not significant for further processing.

In this equation, the result is free of d_{ij} and retrieves b_{ij} if $b_{ir} = 0$.

In an image to be used for water studies, for example, a haze pattern could be reduced by using an infrared (especially SWIR) channel as the reference channel, provided it contains the same haze pattern. Since the NIR and SWIR wavelengths are virtually totally absorbed by water, it is unlikely that they would be required for subsequent processing. Where several channels are available in this wavelength region, the first principal component of those channels could be used (see Section 8).

The relationship between each image channel and the reference channel can be determined by image crossplotting (see Section 1.3.4 and Volume 2A—Section 8.1.3). As illustrated in Figure 7.5, the regression line for each crossplot can then be computed as:
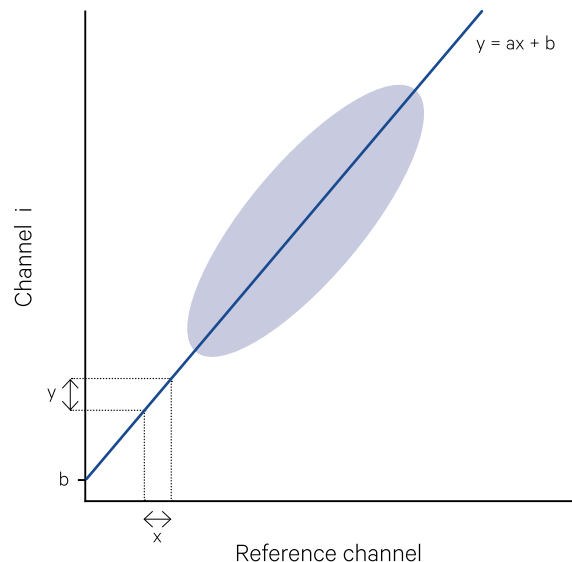
channel i = a$_i$ × reference channel + b$_i$

where

a$_i$ is the regression coefficient
b$_i$ is the regression offset.

Figure 7.5 Using channel crossplot to derive regression line with reference data

The regression equation summarises the relationship between two image channels as y = ax+b, where a = y/x.

Source: Harrison and Jupp (1990) Figure 86

Figure 7.6 Filter to remove image noise

$$
\begin{bmatrix} \text{corrected ch 1} \\ \text{corrected ch 2} \\ \text{corrected ch 3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -a_1 \\ 0 & 1 & 0 & -a_2 \\ 0 & 0 & 1 & -a_3 \end{bmatrix} \times \begin{bmatrix} \text{input ch 1} \\ \text{input ch 2} \\ \text{input ch 3} \\ \text{reference ch} \end{bmatrix} + \begin{bmatrix} a_1 \times \overline{R} \\ a_2 \times \overline{R} \\ a_3 \times \overline{R} \end{bmatrix}
$$

or

corrected ch 1 = $1 \times$ input ch 1 $- a_1 \times$ reference ch $+ a_1 \times \overline{R}$

corrected ch 2 = $1 \times$ input ch 2 $- a_2 \times$ reference ch $+ a_2 \times \overline{R}$

corrected ch 3 = $1 \times$ input ch 3 $- a_3 \times$ reference ch $+ a_3 \times \overline{R}$

The proportion of our image signal in channel i which is modelled as the spatial 'noise' can then be subtracted by the equation:

channel i' = channel $i - a_i \times$ reference channel

To ensure that the scaling of channel i' is consistent with channel i (and that they have the same mean values), we usually add an offset of:

$a_i \times \overline{R}$

where

$\overline{R}$ is the mean of the reference channel.

In an affine transformation of a four channel input image, this model would be specified as shown in Figure 7.6 above.

The reference channel may need pre-processing for some image sources. For example, in scan-digitised imagery the blank target scan can have considerable local variation, so may need to be heavily smoothed (with a 5×5 or 7×7 average filter with all weights equal to 1—see Section 4 above) before being used for illumination correction. Other sources may require calibration to radiance units. To assess the impact of this correction, the transformed and original image pairs can be differenced (see Volume 2D).

7.2.2 Removing 'Limb Brightening'

'Limb brightening', a characteristic increase in brightness towards the edges of the image, commonly occurred in early airborne scanner imagery. This resulted from the wide scan angle, which caused edge pixels to be imaged through a significantly longer atmospheric path distance. This effect can be corrected using an affine transformation.

The brightness trend across the image can be estimated for each channel from the average brightness down each pixel column. The average brightness line may be represented in an image by duplicating the line to match the number of lines in the original image (see Volume 2A—Section 7.2.1.4). This is then subtracted from the original image (with coefficients of 1 and an offset equal to the mean of the original image) to produce a brightness corrected image :

corrected channel = $1 \times$ original channel

$- 1 \times$ average brightness line

$+$ original channel mean

To correct all channels in an image, this operation would be performed more efficiently using image differencing (see Volume 2D). The statistics associated with along-line brightness variation within an image can be accumulated in some image processing systems. For example, the mean intensity for each channel could be extracted to enable the correction described above.

During a transformation, it's understandable to be hesitant about moving too fast, especially at the outset.
(Irene Rosenfeld)

7.3 Other Affine Transformations

A range of affine transformations have been proposed for multivariate EO imagery to variously highlight data or noise components for particular purposes. These methods analyse the spectral and/or spatial correlation of one or more images and are particularly relevant to change detection studies (see Volume 2D). Examples of such transformations include:

- Principal Components Analysis (PCA; see Section 9)—rotates image channels such that the data are arranged along axes of decreasing variance;

- Canonical Variates Analysis (CVA) transformation matrix (see Volume 2E)—maximises the separation between class sample image values;

- Maximum Autocorrelation Factor (MAF)—derives new, orthogonal channels in which neighbouring pixels have maximum autocorrelation to highlight contiguous regions (Switzer and Green, 1984; Conradsen *et al.*, 1985; Nielsen *et al.*, 1998; Nielson, 2011). Autocorrelation statistics are derived from analysis of the dispersion matrix of the difference between each image channel and its spatially shifted self;

- Minimum/Maximum Noise Fraction (MNF)—orders transformed image components in terms of image quality (Green *et al.*, 1988; Nielsen, 2011); and

- Multivariate Alternation Detection (MAD)—analyses canonical correlations between pairs of image channels to highlight differences (Nielsen *et al.*, 1988; see Volume 2D). Nielsen (2007) describes an extension to this transformation, termed the Iteratively Reweighted MAD method (IR-MAD).

Factor analysis methods are further discussed in Volume 2E.

7.4 Further Information

Sawyer (1955)

7.5 References

Conradsen K., Ersboll B. K., Thyrsted T. (1985). A comparison of min/max autocorrelation factor analysis and ordinary factor analysis. *Nordic Symposium in Applied Statistics*, Lyngby. pp. 47-56.

Green, A.A., Berman, M., Switzer, P., and Craig, M.D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing,* 26(1), 65–74.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Nielsen, A.A., Conradsen, K. and Simpson, J.J. (1998): Multivariate alteration detection (MAD) and MAF post-processing in multispectral, bi-temporal image data: new approaches to change detection studies. *Remote Sensing of Environment,* 64, pp. 1-19.

Nielson, A.A. (2007). The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data. *IEEE Transactions on Image Processing*, 16(2), 463–478.

Nielson, A.A. (2011). Kernel Maximum Autocorrelation Factor and Minimum Noise Fraction Transformations. *IEEE Transactions on Image Processing*, 20(3), 612–624.

Sawyer, W.W. (1955). *Prelude to Mathematics*. Penguin Books Ltd. London.

Switzer, P., Kowalik, W.S., and Lyon, R.J.P. (1981). Estimation of atmospheric path-radiance by the covariance matrix method. *Photogrammetric Engineering and Remote Sensing,* 47, 1469–1476.

Switzer, P., and Green, A.A. (1984). Min/max autocorrelation factors for multivariate spatial imagery. *Technical Report 6,* April 1984, Dept. Statistics, Stanford.

8 RGB-HSI Conversions

The RGB (Red, Green, Blue) and HSI (Hue, Saturation, Intensity) models for colour specification were introduced in Volume 2A—Section 5.1.2. These models are related by their intensity axes and can be interconverted using affine transformations.

In this section, after reviewing the relationship between the RGB and HSI colour systems (see Sections 8.1 and 8.2), we consider some transformations that modify image colour by manipulating these colour models:

- enhancing chromaticity (see Section 8.3);
- merging higher spatial resolution panchromatic data with multispectral data (see Section 8.4); and
- combining topographic shading with colour imagery for relief effects (see Section 8.5).

By isolating the intensity information in the image to a single channel, various transformations can be used to modify image intensity before converting the HSI colour values back to the RGB colour space. Two examples of transformations that use the HSI colour system to enhance image appearance are given in Excursus 8.1.

8.1 RGB Colour Model

In the Red-Green Blue (RGB) colour cube, the intensity axis is defined by the line:

$$R=G=B$$

as illustrated in Figure 8.1. Colours in the RGB colour space can also be defined as triangular coordinates on equal intensity planes across the cube (see Figure 8.2a). These coordinates can be normalised within each plane such that the values along each axis from the triangle centre to an apex vary from 0–1 (see Figure 8.2b).

Cartesian coordinates in the RGB cube can be converted to normalised triangular coordinates by the equations:

$$r = \frac{R}{R+G+B}$$
$$g = \frac{G}{R+G+B}$$
$$b = \frac{B}{R+G+B}$$

By definition, these normalised coordinates sum to 1. Since, for any colour r+g+b=1, only two components need to be known to deduce the third (see Figure 8.2c). For example at the white point, r=1/3 and g=1/3, so:

$$b=1-r-g$$
$$=1/3$$

We are like chameleons, we take our hue and the color of our moral character
*from those who are around us.*
(John Locke)

Figure 8.1 RGB colour cube

The colours formed by differing proportions of the additive primary colours, red green and blue, can be represented as a three-dimensional data space. The diagonal axis in this space, referred to as the intensity axis, represents the shades of grey, from black to white, which have equal proportions of each primary colour. On planes orthogonal to the intensity axis, all colours have equal intensity, with saturation increasing away from the axis and hue varying around it.

a. Orientation of RGB colour cube
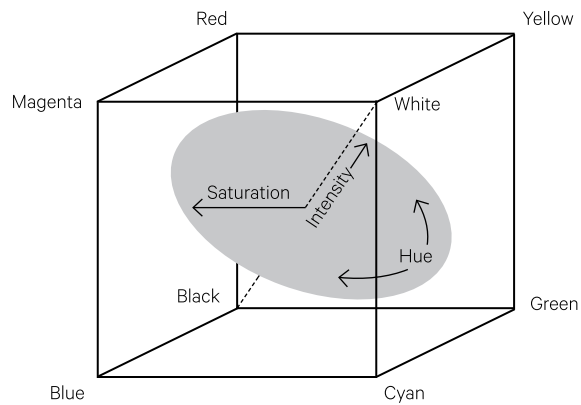
b. Colour planes on surface of RGB colour cube
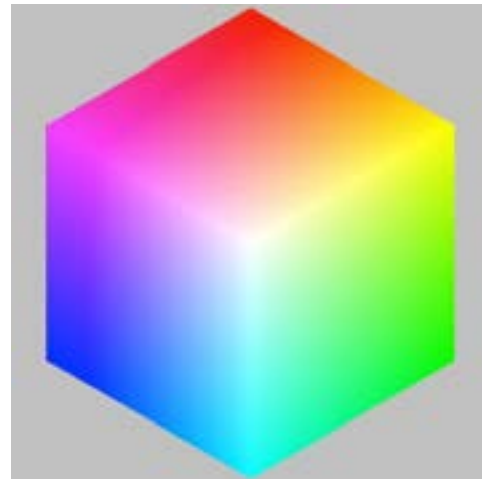
Source: a. Harrison and Jupp (1990) Figure 11; b. Frank Clark. Retrieved from: *http://www.frank-t-clark.com/Professional/Papers/ColorHCW/ColorHCW.html*)

Figure 8.2 RGB triangular coordinates

a. Definition of triangular planes in RGB cube
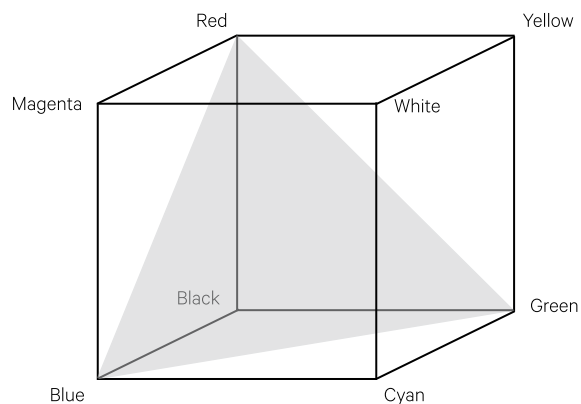
b. Normalised triangular coordinates
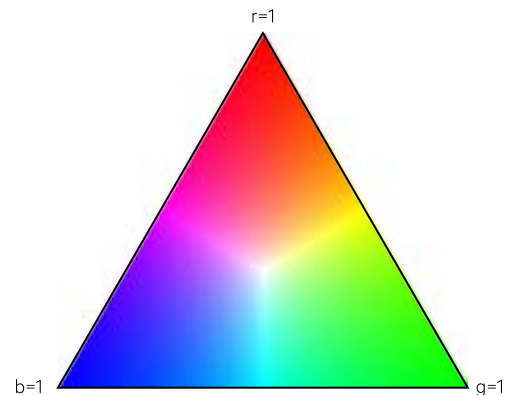
c. Normalised chromaticity coordinates. Since, for any colour r+g+b=1, only two components need to be known to deduce the third. For example at the white point, r=1/3, g=1/3, so b=1-r-g=1/3.

Source: Harrison and Jupp (1990) Figure 102

82

Earth Observation: Data, Processing and Applications. Volume 2: Processing

8.2 HSI Colour Model

The Hue-Saturation-Intensity (HSI or HSV: Hue-Saturation-Value or HSB: Hue-Saturation-Brightness[7]) model defines colours on the equal intensity planes as coordinates of hue and saturation, with hue b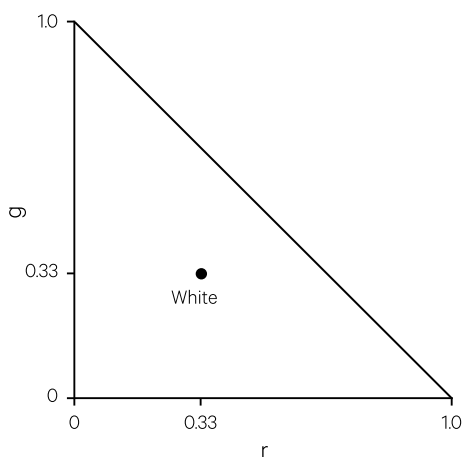eing measured as an angle around the plane and saturation as the radial distance from the centre of the plane (see Figure 8.3). This colour model can be useful to vary the relationships between individual colours, and is a common feature in image processing systems for enhancing photographic images. In terms of EO imagery, it is also a valuable approach for substituting intensity information before visual interpretation.

To measure hue, we define a hue angle of 0° radiating from the centre to the blue apex. The hue angle can then be defined in degrees in an anti-clockwise direction from the blue radial as:

$$H = \cos^{-1}\left(\frac{2b - g - r}{\sqrt{6\left(b - \frac{1}{3}\right)^2 + \left(g - \frac{1}{3}\right)^2 + \left(r - \frac{1}{3}\right)^2}} \right)$$

or

$$H = \cos^{-1}\left(\frac{0.5 \times \left((b-g) + (b-r)\right)}{\sqrt{(b-g)^2 + (b-r)(g-r)}} \right)$$

On a triangular plane, colours of equal saturation are represented as concentric triangles and are measured as:

$$S = 1 - 3 \times \min\left(r, g, b\right)$$

HSI cylindrical coordinates can also be computed from the RGB Cartesian coordinates (see Volume 2B) as:

$$\text{Intensity} = R + G + B$$

$$\text{Hue} = \tan^{-1}\left(\frac{v_2}{v_1} \right)$$

$$\text{Saturation} = \sqrt{v_1^2 + v_2^2}$$

where

$$v_1 = \frac{1}{\sqrt{6}}R + \frac{1}{\sqrt{6}}G - \frac{2}{\sqrt{6}}B$$

$$v_2 = \frac{1}{\sqrt{2}}R - \frac{1}{\sqrt{2}}G$$

or, in matrix notation:

$$\begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The intensity value is better divided by √3 to give a unit vector, that is the length of the vector from the origin to the point (√3, √3, √3) is 1. The HSI dimensions of an example image are shown in Figure 8.4.

Figure 8.3 HSI coordinates

a. Intensity is measured along the diagonal axis between black and white.
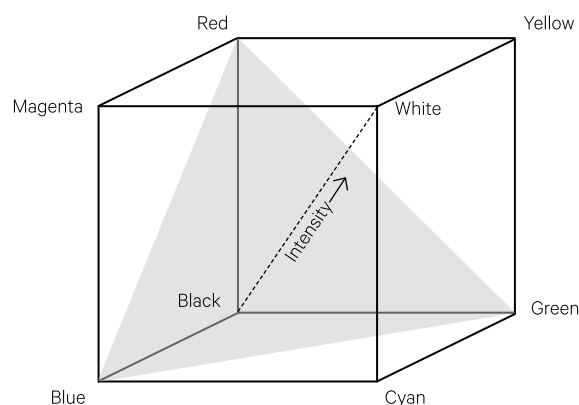
b. Saturation is the linear distance from the intensity axis to the selected colour on an equal intensity plane. Hue is the angular distance from a primary colour (usually blue) to the selected colour on an intensity plane.

Source: Harrison and Jupp (1990) Figure 104

7   Note that the HSL (Hue-Saturation-Luminosity) colour model is different to HSI.
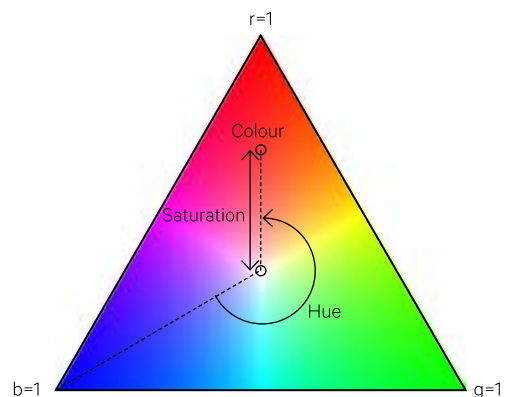
Figure 8.4 Hue, Saturation and Intensity dimensions

This example image was acquired by Sentinel-2a on 10 September 2018 in the Kimberley region, northwest WA.

a. Colour (bands 4, 3 and 2 as RGB)

b. Hue

c. Saturation

d. Intensity

Source Norman Mueller, Geoscience Australia

Colours are the smiles of nature.
(Leigh Hunt)

HSI coordinates can similarly be converted to RGB coordinates using the inverse transformation of:

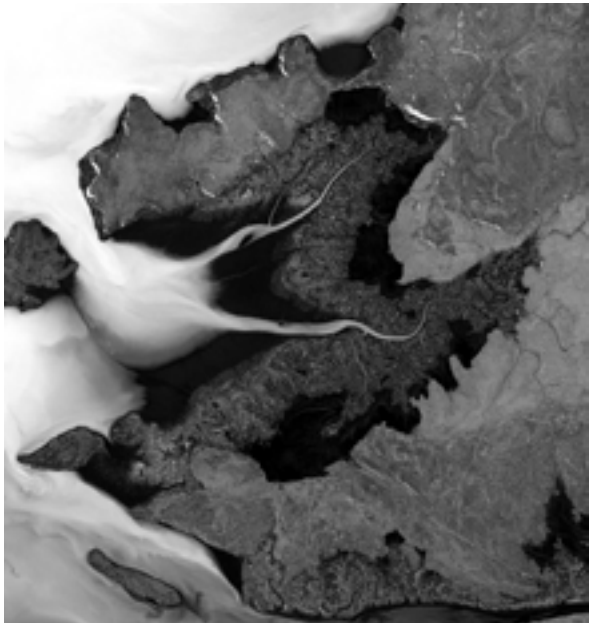$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix} \times \begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix}$$

where

v_1 = saturation × cos (hue); and

v_2 = saturation × sin (hue).

Since v_1 and v_2 are defined above, this transformation can also be expressed as:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{\sqrt{3}} & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{\sqrt{3}} & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \times \begin{bmatrix} I \\ R \\ G \\ B \end{bmatrix}$$

This single operation effectively replaces the intensity component of the RGB colour composite with another channel defined as the image intensity (I). The rotation from RGB to HSI can be varied to weight hues by their perceived brightness. For example, Niblack (1986) suggests a variation (derived from the luminosities of a NTSC (National Television System Committee: US) standard monitor), which effectively tilts the planes of constant intensity with respect to the intensity axis.

The HSI model has been developed to produce various colour spaces shaped as single and double cones, hexcones and the Munsell irregular shape (Munsell 1929). Spherical coordinates for HSI are also defined by Gillespie *et al.* (1987; see Section 8.3). Other systems for specifying colour are introduced in Volume 2X—Appendix 3. Examples of processes using the RGB-HSI transformation with EO imagery are described in Excursus 8.1.

8.3 Chromaticity Transformation

Gillespie *et al.* (1987) demonstrate the use of the rgb chromaticity coordinates to enhance colour composite imagery. Their method involves converting the channels displayed as red, green and blue to normalised rgb coordinate channels and computing an intensity channel as the sum of their values (see Section 8.1).

This transformation can be implemented using:

- an affine transformation to produce the intensity channel (I; see Section 7);

    I=R+G+B

then

- channel ratioing to divide each composite channel by the intensity channel (see Section 8.2):

$$r = \frac{R}{I}$$

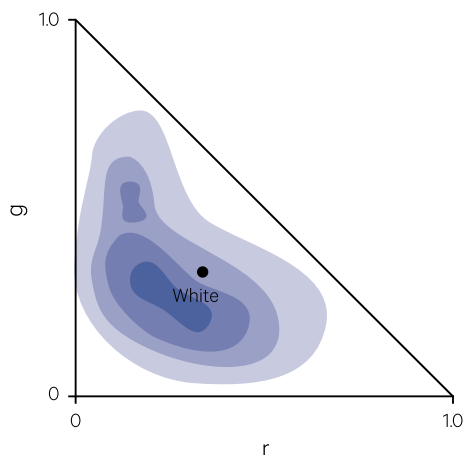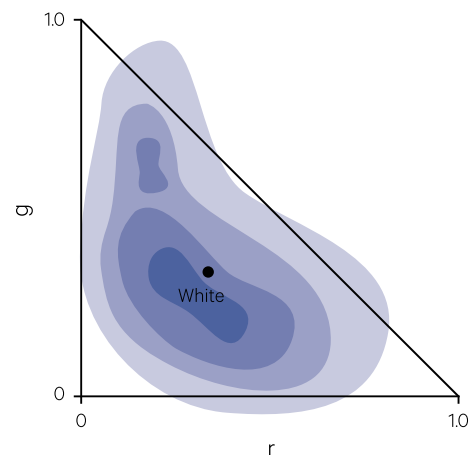$$g = \frac{G}{I}$$

$$b = \frac{B}{I}$$

The spectral geometry of this procedure is illustrated in Figure 8.5.

Figure 8.5 Chromaticity transformation

a. Original image data values are shown on a crossplot of r versus g.

b. Each colour component is expanded individually to fill the colour space.

Source: Harrison and Jupp (1990) Figure 103 [Adapted from Gillespie *et al.* (1987)]

The resulting channels can then be contrast-enhanced individually and transformed back to the original channels by multiplying the rgb channels by the intensity channel, I:

R=r×I

G=g×I

B=b×I

Gillespie *et al.* (1987) suggest that colour pictures prepared this way are similar to those enhanced by the HSI and decorrelation methods (see Sections 8.2 and 9.6 respectively).

As further discussed in Section 9.7, alternative intensity data could also be used in the back transformation if required. For example, pan-sharpening describes the operation of merging finer scale panchromatic channel with registered, coarser scale multispectral imagery (see Section 8.4). Image fusion based on the chromaticity transformation is commonly called the Brovey transform (Hallada and Cox, 1983; Jensen, 2016). However, as with many other pan-sharpening techniques, colour distortions can occur in the merged image, especially when the spectral range of the intensity channel differs from the spectral ranges of the colour channels (Liu and Moore, 1998).

8.4 Pan-Sharpening

Pan-sharpening (or panchromatic sharpening) describes the process of merging finer scale panchromatic imagery with a coarser scale colour composite to improve the apparent spatial resolution of the colour image. As further detailed in Volume 2D a wide range of methods have been proposed for pan-sharpening EO imagery (Thomas *et al.*, 2008; Amro *et al.*, 2011; Jalan and Sokhi, 2012).

One approach to pan-sharpening uses an affine transformation to effectively convert between the RGB and HSI colour spaces (see Section 8.2). In this case, the colour composite image would need to be resampled to the same pixel size as the panchromatic

imagery and both datasets would need to precisely overlay a common spatial grid (see Volume 2B). The affine transformation matrix shown in Figure 8.6 could then be used to replace the multispectral intensity information with the panchromatic channel.

An example of this transformation applied to Landsat-8 OLI and Panchromatic imagery is shown in Figure 8.7. Spectral distortions can occur in the merged image if the spectral range of the intensity channel differs from the spectral ranges of the colour channels (Liu and Moore, 1998; Thomas *et al.*, 2008; see Volume 2D). Other methods for pan-sharpening are described in Section 9.7 and Volume 2D.

Figure 8.6 Matrix equation to merge high resolution data

$$
\begin{bmatrix} \text{high res ch 1} \\ \text{high res ch 2} \\ \text{high res ch 3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & \frac{1}{\sqrt{3}} \end{bmatrix} \times \begin{bmatrix} \text{rescaled ch 1} \\ \text{rescaled ch 2} \\ \text{rescaled ch 3} \\ \text{panchromatic ch} \end{bmatrix}
$$

or

$$\text{high res ch 1} = \frac{2}{3} \times \text{rescaled ch 1} - \frac{1}{3} \times \text{rescaled ch 2} - \frac{1}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{panchromatic ch}$$

$$\text{high res ch 2} = -\frac{1}{3} \times \text{rescaled ch 1} + \frac{2}{3} \times \text{rescaled ch 2} - \frac{1}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{panchromatic ch}$$

$$\text{high res ch 3} = -\frac{1}{3} \times \text{rescaled ch 1} - \frac{1}{3} \times \text{rescaled ch 2} + \frac{2}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{panchromatic ch}$$

Figure 8.7 Integrating imagery with differing spatial resolution

Landsat-8 panchromatic and OLI images were acquired over Canberra on 25 October 2018. These images have been merged to create a colour image with the spatial resolution of the panchromatic image.

a. Landsat-8 panchromatic image with 15 m pixels

b. Landsat-8 OLI image with 30 m pixels, displayed using bands 6, 5, 3 as RGB

c. Panchromatic and OLI images are merged to create a colour image with 15 m pixels.

Source: Norman Mueller, Geoscience Australia

8.5 Adding Relief Shading

A filter-based process for generating shaded relief imagery from topographic data was introduced in Section 6.2. Such relief imagery can be integrated with the brightness of the composite colours in a colour composite image to represent changes in elevation. For example, in Figure 8.8 a shaded relief channel for southeast Australia has been merged with a thematic image of the Murray-Darling basin to emphasis the role of topography in regional drainage.

The merging process requires that the data range of the relief channel matches those channels forming the composite image. If the data range of an elevation channel is much larger than the channels forming the colour composite, the image channels must first be rescaled to the same range (usually with the scaling parameters used to display the composite image) so that the relief shading will affect all channels equally.

The relief shading can then be merged into the colour image using the same matrix coefficients given in Section 8.4 for pan-sharpening (see Figure 8.9). The resulting shaded composite will have modified brightness values but the same hues as illustrated in Figure 8.8. Weighting values can also be derived for this operation if required. The resulting shaded relief image is sometimes non-linearly scaled to enhance contrast in dark areas.
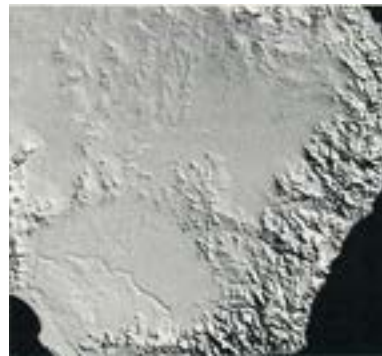
Figure 8.8 Modifying colour brightness using relief shading

An affine transformation can be used to replace the intensity axis of a colour composite image with some other intensity scale. In this example, a Sun shadowing channel (derived from elevation data for southeast Australia) has been used to highlight surface relief within colour-coded drainage divisions of the Murray-Darling basin. A segmentation mask has been used to restrict the effect of the transformation to within the basin.

a. Original colour image   b. Shaded relief image   c. Colour image with shaded relief

Source: Harrison and Jupp (1990) Plate 17

Figure 8.9 Matrix equation to add relief shading

$$\begin{bmatrix} \text{shaded ch 1} \\ \text{shaded ch 2} \\ \text{shaded ch 3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & \frac{1}{\sqrt{3}} \end{bmatrix} \times \begin{bmatrix} \text{rescaled ch 1} \\ \text{rescaled ch 2} \\ \text{rescaled ch 3} \\ \text{shaded relief ch} \end{bmatrix}$$

or

$$\text{shaded ch 1} = \frac{2}{3} \times \text{rescaled ch 1} - \frac{1}{3} \times \text{rescaled ch 2} - \frac{1}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{shaded relief ch}$$

$$\text{shaded ch 2} = -\frac{1}{3} \times \text{rescaled ch 1} + \frac{2}{3} \times \text{rescaled ch 2} - \frac{1}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{shaded relief ch}$$

$$\text{shaded ch 3} = -\frac{1}{3} \times \text{rescaled ch 1} - \frac{1}{3} \times \text{rescaled ch 2} + \frac{2}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{shaded relief ch}$$

Excursus 8.1—Examples of HSI-based transformations

Various transformations of EO image channels can be informative for particular applications. To maximise the visual content of an image, selected transformations can be combined into a colour composite, either directly or via an HSI transformation (see Section 8.2). For example:

- information in multi-channel images may be enhanced by substituting appropriate transformations for the hue, saturation and/or intensity components; or
- multiple transformations applied to a single channel can be combined into a colour composite using the HSI transformation.

Two examples of these options are detailed below.

Substituting transformed channels for HSI components

Jaskolla and Henkel (1988) reported a transformation sequence that extends the colour variation in EO imagery compared with a standard false colour composite. Using three input channels containing green, red and NIR reflectance, this sequence computes three intermediate channels to represent the intensity, hue and saturation components:

- intensity=high pass filter on the mean value of green, red and NIR bands;
- hue=NIR–(green+red)/2; and
- saturation=red–green.

These three channels can then be combined as an HSI composite and transformed back to RGB space (see Section 8.2). This transformation was developed for SPOT-MSS imagery and is shown for an example Landsat-5 image in Figure 8.10. This enhancement clearly improves colour contrast in the example image and highlights linear features. As such, it has been reported to be useful for geological interpretation (Jaskolla and Henkel, 1988)
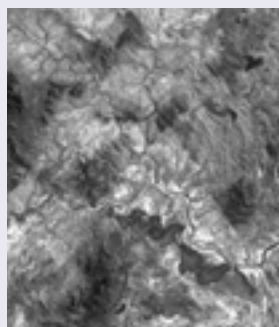
Figure 8.10 Channel substitution using HSI transform

Bands 2, 3 and 4 of a Landsat-5 image, acquired over Charters Towers region, Queensland, on 7 October 1987, are processed to create multiple channels that can be presented as a colour composite using the HSI transformation.
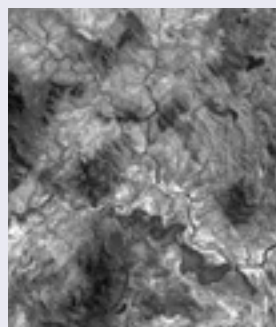
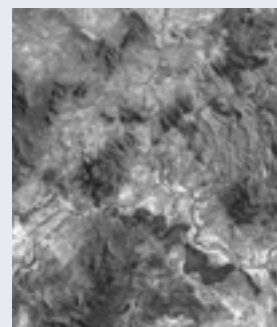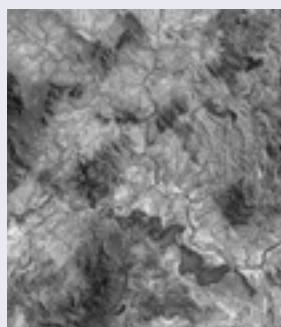a. Original colour composite (bands 4, 3, 2 as RGB)

b. Band 2 (green)
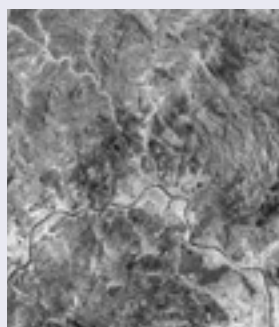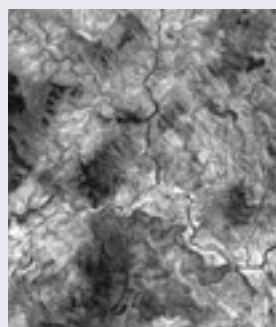
c. Band 3 (red)

d. Band 4 (NIR)

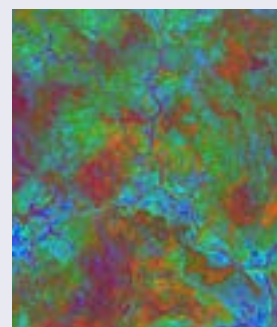e. High pass filter on average of original channels represents intensity

f. (NIR-(green+red) / 2) assumed to represent hue

g. (red-green) assumed to represent saturation

h. Colour composite of HSI transformed back to RGB space

Source: Tony Sparks, Icon Water

89

Presenting single channel data as a colour composite

Any relevant transformations of an image channel could be substituted for the hue, saturation and/or intensity components in a pseudo-HSI composite. For example, the HSI transformation may be used in conjunction with differential filtering of an elevation channel k to display an HSI pseudo-colour composite, which comprises:

intensity=elevation (channel k)

hue=aspect ($v_1 = \delta k\,/\,\delta x$)

saturation=slope ($v_2 = \delta k\,/\,\delta y$)

(Please refer to Sections 5.1.1 and 6 for derivative equations.)

These intensity, hue and saturation components can then be converted to RGB coordinates using the inverse HSI transformation (see Section 8.2):

$$
\begin{bmatrix} R \\ G \\ B \end{bmatrix} =
\begin{bmatrix}
\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0
\end{bmatrix} \times
\begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix}
$$

or

$$R = \frac{1}{\sqrt{3}} \times I + \frac{1}{\sqrt{6}} \times v_1 + \frac{1}{\sqrt{2}} \times v_2$$
$$G = \frac{1}{\sqrt{3}} \times I + \frac{1}{\sqrt{6}} \times v_1 - \frac{1}{\sqrt{2}} \times v_2$$
$$B = \frac{1}{\sqrt{3}} \times I - \frac{2}{\sqrt{6}} \times v_1$$

This representation is applied to an elevation channel in Figure 8.11. It is also reported as being a useful enhancement sequence for visual interpretation of single channel geophysical data (Harrison and Jupp, 1990).

Figure 8.11 Single channel presented as colour composite

Slope and aspect channels have been computed from an elevation channel. When combined as an HSI composite image (in which elevation is represented as intensity, aspect is represented as hue, and slope represented as saturation), these channels can be back-transformed to form an RGB colour composite.

a. Elevation

b. Slope

c. Aspect

d. Colour composite

Source: Harrison and Jupp (1990) Plates 13 and 14

8.6 Further Information

Gonzalez and Woods (2018)

Jensen (2016) Section 8

Sawyer (1955)

8.7 References

Amro, I., Mateos, J., Vega, M., Molina, R., and Katsaggelos, A.K. (2011). A survey of classical methods and new trends in pansharpening of multispectral images. *EURASIP Journal of Advances in Signal Processing 2011,* 2011: 79. *https://doi.org/10.1186/1687-6180-2011-79*

Fung, T., and LeDrew, E. (1987). Application of Principal Components Analysis to change detection. *Photogrammetric Engineering and Remote Sensing,* 53, 1649–1658.

Gillespie, A.R., Kahle, A.B., and Walker, R.E. (1987). Colour enhancement of highly correlated images. II. Channel ratio and 'chromaticity' transformation techniques. *Remote Sensing of Environment,* 22(3), 343–365.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing.* Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.

Hallada, W.A., and Cox, S. (1983). Image Sharpening for Mixed Spatial and Spectral Resolution Satellite Systems. *Proc 17th International Symposium on Remote Sensing of Environment.* 9–13 May 1983, Ann Arbor. pp 1023–1032.

Jalan, S., and Sokhi, B.S. (2012). Comparison of different pan-sharpening methods for spectral characteristic preparation: multi-temporal CARTOSAT-1 and IRS-P6 LISS-IV imagery. *International Journal of Remote Sensing,* 33(18), 5629–5643.

Jaskolla, F., and Henkel, J. (1988). Evaluation and digital processing of multispectral SPOT data. *International Journal of Remote Sensing,* 9, 1629–1637.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective.* 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Liu, J.G., and Moore, J.Mc.M. (1998). Pixel block intensity modulation: Adding spatial detail to TM band 6 thermal imagery. *International Journal of Remote Sensing,* 19(13), 2477–2491. doi: *10.1080/014311698214578*

Munsell, A. H. (1929). *Munsell Book of Color.* Munsell Colour Co., Inc.

Niblack, W. (1986). *An Introduction to Digital Image Processing.* Prentice-Hall International. New Jersey.

Sawyer, W.W. (1955). *Prelude to Mathematics.* Penguin Books Ltd. London.

Singh, A., and Harrison, A. (1985). Standardised Principal Components. *International Journal of Remote Sensing,* 6, 883–896.

Thomas, C., Ranchin, T., Wald, L., and Chanussot, J. (2008). Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geoscience and Remote Sensing,* 46, 1301–1312.

9 Principal Components Analysis

The Principal Components transformation (PCA, also referred to as eigenvalue, Hotelling or discrete Karhunen-Loeve (KL) transforms) uses image spectral statistics to define a rotation of the original image channels such that the data are arranged along axes of decreasing variance. The coordinates for the new axes are computed by an affine transformation (that is, a linear combination) of the original data coordinates (see Section 7).

The transformation parameters for PCA are derived from the covariance or correlation matrix (see Section 1.3.3 above) of an image (or sub-image). This transformation is particularly useful for imagery with many similar channels, such as hyperspectral imagery. The resulting Principal Components (PCs) are totally uncorrelated and can be computed by adding proportions of the original image channel values. Many image processing systems offer options for adjusting channel variance both when defining and applying the PC transformation matrix, and also to back transform from the PCA space to the original data channels.

In this section we introduce the PCA transformation (see Section 9.1), interpretation of principal components (PCs; see Section 9.2), and the transformation parameters (see Section 9.3), then present common applications of PCA in the context of EO imagery:

- decorrelation stretching (see Section 9.4);
- data reduction (see Section 9.5);
- highlighting specific features (see Section 9.6);
- multi-scale intensity enhancement (see Section 9.7); and
- change detection (see Section 9.8).

9.1 PCA Transformation

As discussed in Section 7.1, a multi-channel image can be considered as a multidimensional matrix, or multi-vector space. For a two channel image, the two-dimensional data space can be represented in a crossplot (see Section 1.3.4 above). It is possible to redefine pixel values in an image by geometrically changing the positions of, or relationship between, the axes of this data space. For example, in Figure 9.1, by rotating and shifting the channel axes, the point A becomes defined as position p' and q'. Data space axes can similarly be rescaled, reflected or rotated as detailed in Volume 2X—Appendix 6. Such changes can allow spectral patterns and relationships in the data to be enhanced and analysed more directly.

Mathematical matrices allow changes to the 'spectral' geometry of two and higher-dimensional data to be represented algebraically. For example, in Figure 9.1, channels X and Y show a high degree of correlation in the original data space. As detailed in Volume 2X—Appendix 6, the original axes can be rotated and each pixel assigned a new value relative to the new coordinate system. In Figure 9.1 the rotated channel X' is aligned with the direction of maximum variance so that the data channels are uncorrelated in the rotated space. This type of rotation is used in Principal Components Analysis for data reduction and enhancement.

The first rule of intelligent tinkering is to save all the parts. (Paul R. Ehrlich)

93

The principle of a PCA transformation is illustrated in Figure 9.2 for a two-channel image. The first principal component, PC1, is defined as the axis that passes through the mean data value and is aligned with maximum variance in the image. The second principal component, PC2, is defined as being orthogonal (perpendicular) to PC1.

The same concept can be applied to images with more than two channels, with the rotation transformation being computed from the covariance matrix (see Section 1.3.3). This matrix would typically be derived from the whole image but, if PC1 was required to be aligned with the variation in a particular cover type to enhance a specific feature, it may also be based on a training set (see Volume 2A— Section 9.1.2) or spectral theme (see Volume 2A— Section 9.1.3).

The covariance or correlation matrices indicate the degree to which image channels are correlated (see Section 1.3.3 and Volume 2A—Section 8.1.4). The PCA transformation produces an image with totally uncorrelated channels. This is done by mathematically defining a rotation that creates an image with all zero covariance values. In terms of the covariance matrix, this would mean that all off-diagonal elements were equal to zero. Since only the diagonal elements are non-zero, such a matrix is referred to as a diagonal matrix.

As detailed in Section 7 above, a general linear transformation can be represented as a matrix operator. This transformation redefines each image pixel relative to a new coordinate system. Special vectors exist between the two coordinate systems such that points on these vectors are simply redefined by the transformation as a constant multiple of their original coordinates. These vectors are called eigenvectors of the matrix and the constant values are called eigenvalues. Mathematically this is expressed as shown in Figure 9.3. A simple image example of the computation of these statistics is presented in Excursus 9.1. The eigenvectors and eigenvalues of a matrix may be found using matrix algebra as detailed in Volume 2X—Appendix 6 (which gives the mathematical derivation of the PCA transformation; see also Volume 2E).

Figure 9.1 Redefining image axes

Image data values can be redefined relative to modified position and scaling of axes.
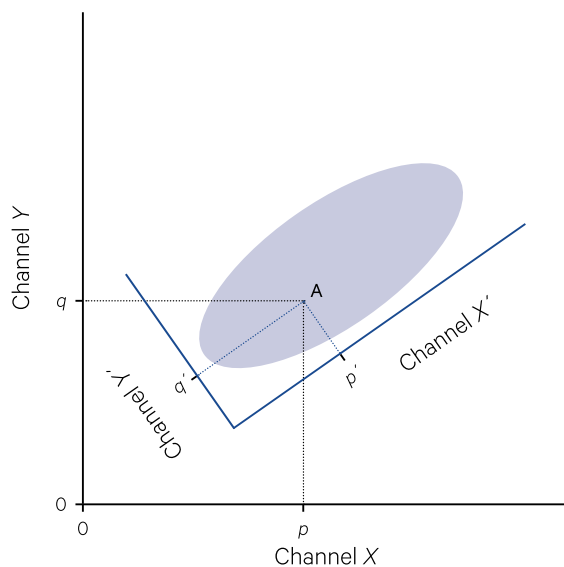
Source: Harrison and Jupp (1990) Figure 87

Figure 9.2 Principle of PCA transformation

A new primary data axis (PC1) is defined in the direction of maximum image variance (over all image channels) then other axes are defined to be orthogonal to this direction.
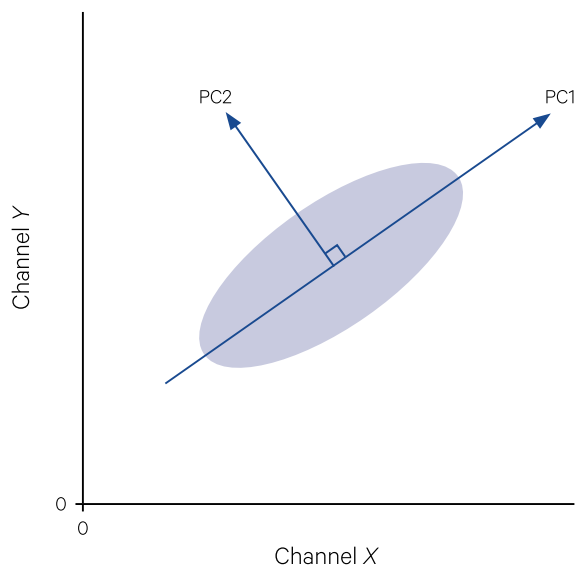
Source: Harrison and Jupp (1990) Figure 88

Nothing can be divided into more parts than it can be constituted of.
But matter (i.e. finite) cannot be constituted of infinite parts.
(Isaac Newton)

Figure 9.3 Eigenvectors of a matrix

Special vectors called eigenvectors exist between the original and transformed image coordinates, which can represent the transformation as a constant multiple (called the eigenvalue) of the original coordinates.

$$
\begin{array}{c} \text{Channel 1} \\ \\ \text{Channel 2} \end{array}
\underset{\text{Transformation matrix}}{\begin{bmatrix} a & b \\ \\ c & d \end{bmatrix}}
\times
\underset{\text{Eigenvectors}}{\begin{bmatrix} x \\ \\ y \end{bmatrix}}
=
\underset{\text{Eigenvalue}}{k}
\times
\underset{\text{Eigenvectors}}{\begin{bmatrix} x \\ \\ y \end{bmatrix}}
$$

or

Channel $1 = a \times x + b \times y = k \times x$

Channel $2 = c \times x + d \times y = k \times y$

Source: Harrison and Jupp (1990) Figure 89

The major direction of data variation in multi-channel EO imagery rarely lies along one original channel axis. Similarly the variance gradient for specific image features is usually defined in at least two data dimensions. These statistical characteristics of EO data make it difficult to neatly analyse the component features or divide up the image brightness variations as channel themes (see Volume 2A—Section 9.1). By aligning the data axes with the variance of the whole image dataset, or just with a specific feature in it, we can access the changes within the major variance using a single channel.

Thus, the PCA transformation allows a greater proportion of the image variance to be represented in fewer uncorrelated channels. These PCs can be combined into a colour composite, which typically displays a larger range of colours than any composite images formed from the original channels. Colour composite images formed from PCs are frequently used to highlight linear features for geological analyses (see also Excursus 8.1).

An example of the PCA transformation is shown in Figure 9.4. In this example, the contrast both between and within features in Figure 9.4a is visually enhanced in Figure 9.4b. In this case, PC1 is dominated by cloud (showing as red or pink in the PCA image) since these features have large contrasting values and create high multidimensional variance in the original data.

Figure 9.4 Principal components transformation

PCA was applied to a 10 band Sentinel-2a image over central Tasmania, acquired on 5 October 2018. This analysis is based on bands 1, 2, 3, 4, 5, 6, 7, 8, 11, and 12 (all visible, red-edge, and NIR bands except 8a; see Figure 1.2).

a., Original cloudy image is displayed as a natural colour composite using bands 4, 3 and 2 as RGB.

b. PCA image formed by displaying PC1 as red, PC2 as green and PC3 as blue.
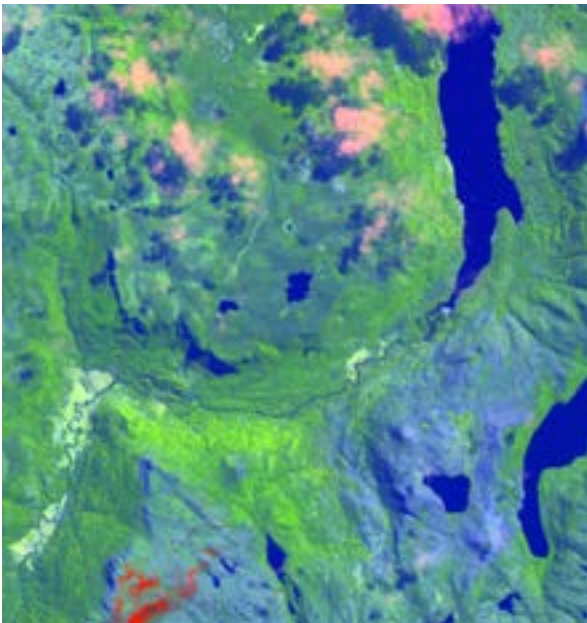
Source: Norman Mueller, Geoscience Australia

Excursus 9.1—PCA statistics

To show the relationship between image statistics and PC equations, we will use the simple example image from Section 1.3.4 (see Figure 1.8). As detailed in Volume 2A—Section 8.1 and reviewed in Section 1.3.3, the covariance matrix for this three channel image can be computed as:

$$C_x = \begin{bmatrix} 2.9380 & -0.3840 & 2.3000 \\ -0.3840 & 4.7200 & -0.0144 \\ 2.3000 & -0.0144 & 2.2656 \end{bmatrix}$$

The eigenvalues of the matrix are derived as:

$$\begin{vmatrix} 2.938 - k & -0.384 & 2.3 \\ -0.384 & 4.72 - k & -0.0144 \\ 2.3 & -0.0144 & 2.2656 - k \end{vmatrix} = 0$$

which gives eigenvalues of 5.1, 4.5 and 0.3 for the three channels (see Volume 2X—Appendix 6.3). The covariance matrix of the PCA image (which by definition has no correlation between channels) can then be written as:

$$\begin{bmatrix} 5.1 & 0 & 0 \\ 0 & 4.5 & 0 \\ 0 & 0 & 0.3 \end{bmatrix}$$

The eigenvectors of the example image covariance matrix can also be computed to form the rows of the PCA transformation matrix (see Volume 2X—Appendix 6.3 for details of derivation). For our example, the final transformation matrix is computed as:

$$\begin{bmatrix} 0.6306 & -0.5879 & 0.5068 \\ 0.4153 & 0.8072 & 0.4195 \\ 0.6556 & 0.0541 & -0.7531 \end{bmatrix}$$

This matrix can now be used to rotate the data space of original image to the uncorrelated PCA image space via the equations:

$$PC1 = 0.6306 \times ch\,A - 0.5879 \times ch\,B + 0.5068 \times ch\,C$$

$$PC2 = 0.4153 \times ch\,A + 0.8072 \times ch\,B + 0.4195 \times ch\,C$$

$$PC3 = 0.6556 \times ch\,A + 0.0541 \times ch\,B - 0.7531 \times ch\,C$$

Other statistics relevant to the PCA transformation are introduced in Volume 2X—Appendix 6.3. For example, the degree of correlation between each original image channel and each principal component can also be determined. For our example image, this is summarised by the normalised matrix:

$$\begin{bmatrix} 0.8307 & -0.6109 & 0.7604 \\ 0.5136 & 0.7875 & 0.5909 \\ 0.2096 & 0.0197 & -0.2742 \end{bmatrix}$$

The elements of this matrix indicate the degree of correlation between the original channels and each PC. For example, channels A and C are strongly correlated with PC1 (0.8307 and 0.7604 respectively) while channel B is negatively correlated with PC1 (-0.6109) and more strongly correlated with PC2 (0.7875).

The percentage variation accounted for by each PC can also be computed for our example image (see Table 9.1), indicating that PC1 and PC2 together represent 97.3% (51.8+45.5) of the variation contained in the three channels of the original image.

Table 9.1 Principal component percentage variation

PC	Variance	Percentage
1	5.1	51.8
2	4.5	45.5
3	0.3	2.7
Total	9.9	100

9.2 Interpretation of Principal Components

Interpretation of principal component channels depends on the features of the imagery being transformed. The matrix coefficients specify the contribution each original image channel is making to the transformed channels, so can indicate the 'meaning' of each PC. In Landsat TM imagery, for example, PC1 usually represents (at least) 90% of the data variation and is typically computed using all positive coefficients. Since the PC1 channel is a useful summary of image brightness, including topographic shading, it may be used in transformations that require a relief shading channel (see Section 8.4) or to split spectral classes derived from log-ratio channels into categories which indicate surface shading. In Landsat

TM imagery containing extensive areas of vegetation, PC2 is usually related to vegetation greenness, while PC3 is generally indicative of vegetation 'yellowness' or soil colour (see also Section 11.1.3). In this case, the coefficients for PC2 would be positive in the infrared channels and negative in the visible channels (or vice versa), however these definitions vary widely depending on characteristics of the scene and the image channels being processed. Low order PCs are generally accepted to represent the noise level in the image and typically contain (much) less than 5% of the data variation.

An example of PCA results for a Landsat TM image is given in Excursus 9.2.

Excursus 9.2—PCA Example

Source: Megan Lewis, University of Adelaide

This excursus presents an example of PCA transformation applied to the optical bands in Landsat TM imagery. The original image channels are shown in Figure 9.5 (see Section 1.1 for band definitions) and a colour composite of three original bands is shown in Figure 9.6.

Figure 9.5 Original image channels

Landsat TM image channels over Fowlers Gap Research Station, NSW, acquired in June 1993.

a. Band 1 (blue)
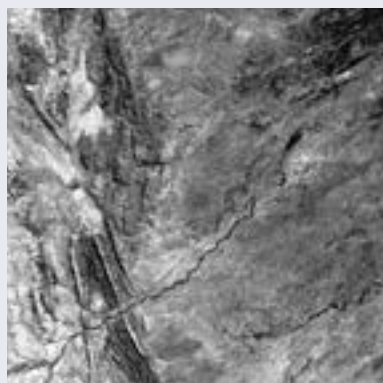
b. Band 2 (green)

c. Band 3 (red)

d. Band 4 (NIR)

e. Band 5 (SWIR1)

f. Band 7 (SWIR2)

Figure 9.6 Colour composites
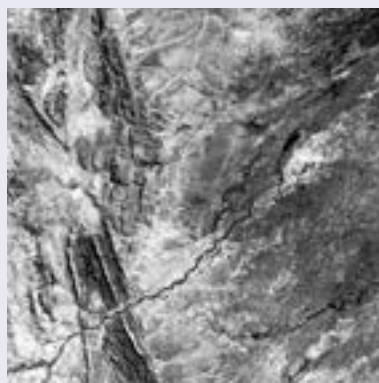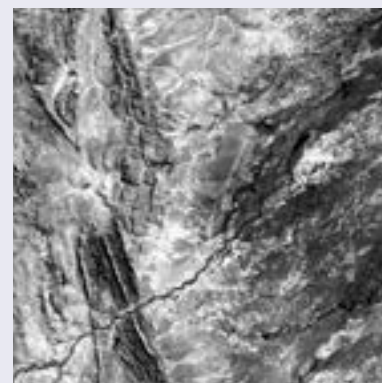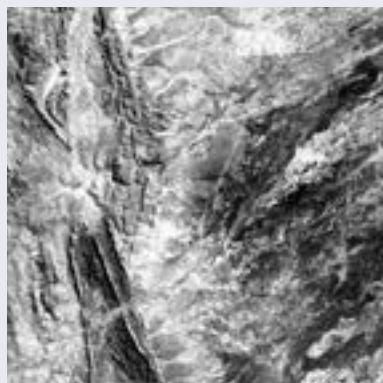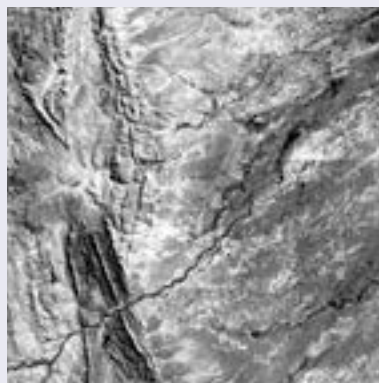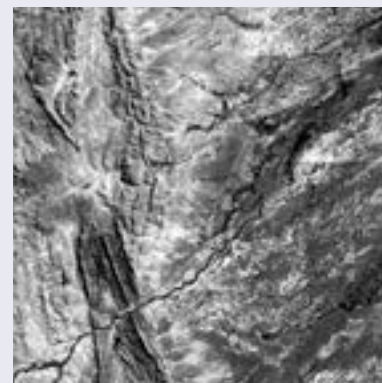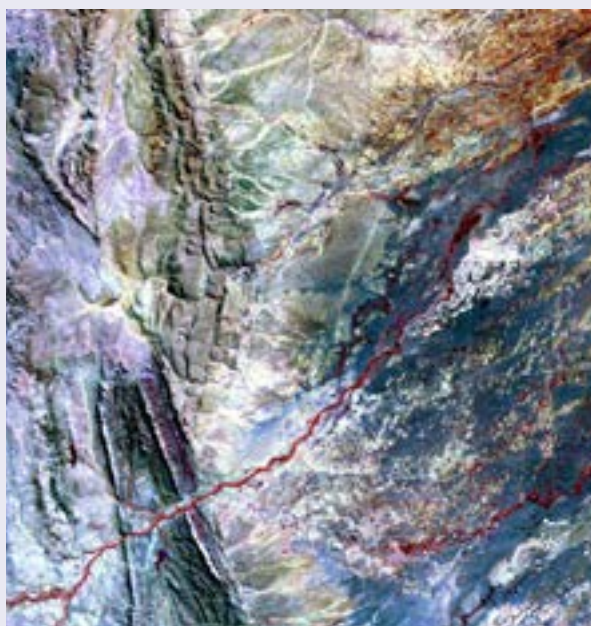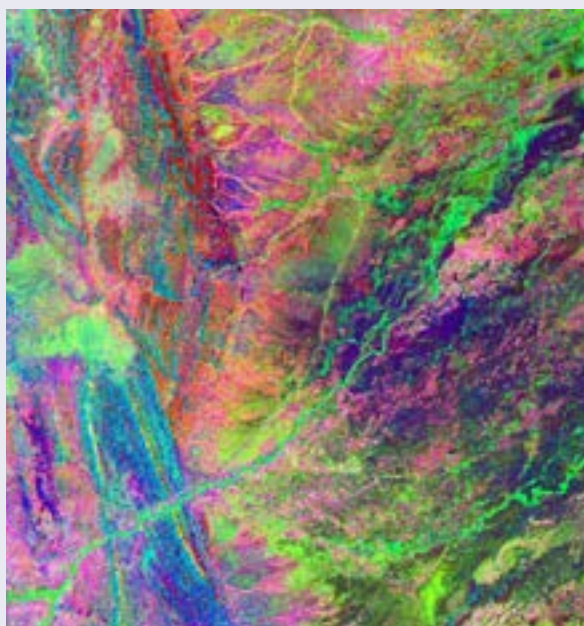
a. Landsat TM bands 4, 3, 2 as RGB

b. PC1, PC2 and PC3 as RGB

The correlation matrix for this image is shown in Table 9.2. As reviewed in Section 1.3.3 (and detailed in Volume 2A—Section 8.1), this table summarises the correlation between image channels. This example matrix indicates that the original TM bands are strongly correlated, with little spectral contrast between land surfaces. For example, the correlation between bands 3 (red) and 7 (SWIR) is 0.879 (~88%), meaning that 88% of the information in band 3 is also available from band 7.

The PC transformation is derived from the correlation matrix to give the transformation matrix shown in Table 9.3. This table summarises the PC equations to convert from the original channels to PCs (see Section 9.1). For example, the first row contains the coefficients required to compute PC1 from the original bands:

$$PC1 = 0.058 \times band1 + 0.059 \times band2 + 0.045 \times band3$$
$$+ 0.375 \times band4 + 0.593 \times band5 + 0.706 \times band7$$

This transformation enables a new set of channels to be computed which isolate sensor noise to one component (PC6 in this example) and maximises the spectral information in the higher components (see Figure 9.7). The range of colours displayed in the false colour composite based on the original channels (see Figure 9.6a) clearly contrasts with the colour composite formed by PCs 1, 2 and 3 (Figure 9.6b), which demonstrates that the three highest PCs contain more 'unique' information than the three original bands displayed in the colour composite.

Table 9.2 Correlation matrix for original image

Band	Band					
	1	2	3	4	5	7
1	1.0	0.788	0.65	0.546	0.554	0.531
2	0.788	1.0	0.902	0.803	0.775	0.773
3	0.65	0.902	1.0	0.902	0.873	0.879
4	0.546	0.803	0.902	1.0	0.836	0.821
5	0.554	0.775	0.873	0.836	1.0	0.928
7	0.531	0.773	0.879	0.821	0.928	1.0

Table 9.3 PC transformation matrix

PC	Band					
	1	2	3	4	5	7
1	0.058	0.059	0.045	0.375	0.593	0.706
2	0.153	0.154	0.137	0.534	0.432	-0.680
3	0.336	0.271	0.243	0.540	-0.652	0.196
4	0.453	0.758	-0.131	-0.426	0.150	-.008
5	0.603	-0.408	-0.675	0.115	-0.021	-0.016
6	0.540	-0.398	0.669	-0.298	0.117	0.006

Figure 9.7 Principal component channels

a. PC1
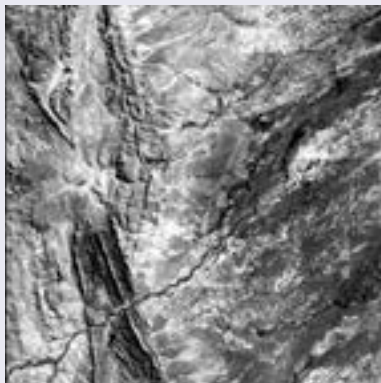
b. PC2
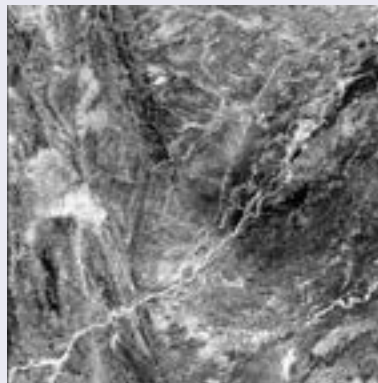
c. PC3

d. PC4

e. PC5

f. PC6

9.3 Transformation parameters

The PCA transformation matrix is generally computed using the specified image statistics (channel means and covariance or correlation matrix), possibly with adjustment for selected channel weights and a balance factor. This matrix can then be applied to selected images as a linear transformation. The resulting PC values are usually rescaled to the full image data range using user-defined expected minimum and maximum values, which are often best determined iteratively.

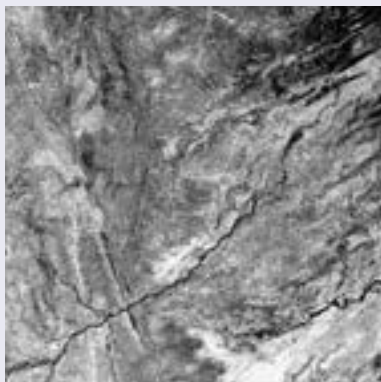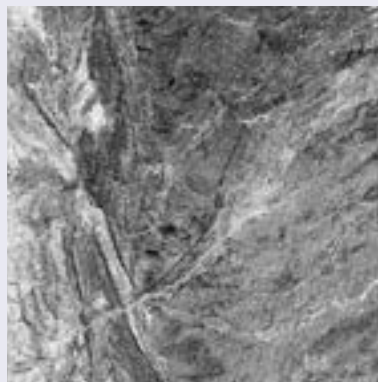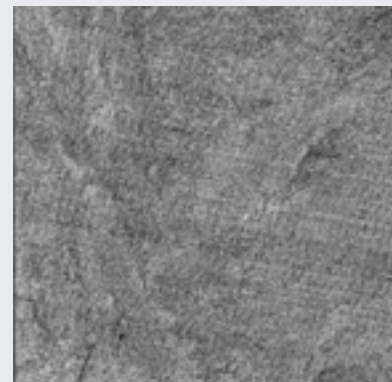A principal component may also be removed in the 'back transformation' from the PCA axes to the original image space. This is particularly relevant to noisy imagery where the 'noise feature' can be isolated as one component. Back transformation without that component then effectively removes the noise from the original data space. In four channel Landsat MSS imagery, for example, the PC4 usually contains (much) less than 5% of the total image variance and visually appears to only summarise image noise.

The image statistics used to define the PC transformation may be derived from one or more images, or may be defined by previous analyses. The latter, for example, may apply to a specific image

feature (see Volume 2A—Section 8) or an image classification (see Volume 2E). Once defined, to compute the PC transformation the channels of the input matrix may be weighted in different ways such as:

- equal—unweighted; this option would be applicable when channel scaling is correct and equal in all channels of the image;

- inversely proportional to channel standard deviation—this converts the covariance matrix to a correlation matrix so each channel is considered to have the same variance; this option may be useful when the image channels are from different sources (in terms of sensor or date) and/or different transformations, and have totally different scaling;

- inversely proportional to channel data range—this can be used for nominal data or when variance is not related to image distribution, for example in an image with a bi-modal histogram; or

- defined by the user—you may already know the relationship between channels and want to specify which one(s) make the greatest contribution to the PCA; this option is also useful to calibrate the data channels to radiance or some appropriate units before processing.

In general terms, unweighted PCA assumes the error component in the image data is uncorrelated and has the same variance in all channels whereas weighted PCA assumes that the error in all channels is uncorrelated and has variance proportional to the total image variance. Local variance statistics (see Volume 2A—Section 8.2.1) can be considered as an estimate of relative error levels between channels so, for each channel, weights for the PCA could be determined as:

$$\frac{1}{\text{local standard deviation}}$$

Singh and Harrison (1985) claim that a significant improvement in the signal-to-noise ratio of PC images can be achieved using the correlation rather than covariance matrix for PCA of EO data. They refer to the results of a transformation based on the correlation matrix as 'standardised PCs'. This standardisation is suggested to be especially useful in multi-temporal analyses to minimise differences due to atmospheric conditions or Sun angle between images (Fung and LeDrew, 1987).

User-defined channel weighting may be required for nominal data where the scaling in the original channels has been specifically adjusted for category definition, and the range and variance statistics are not comparable between channels. In most EO applications, however, channel weighting is not required since the scaling between channels of a single image is relatively consistent in current image data sources. It cannot be assumed, however, that the scaling will be consistent between different images (from multiple dates or sources; see Volume 2D).

The scaling of the PC channels can also be modified by a 'balance factor', which could range from a minimum value (often 0), where all PC scales are equal, to a maximum value (often 1), which scales each PC in proportion to variance. This parameter is most relevant to back rotation from the PCA space to the original image data space as it effectively allows contrast enhancement of the transformed data before conversion to the original data space. This produces similar results to a decorrelation stretch as described in Section 9.4.

9.4 Decorrelation Stretching

Imagery produced by PCA typically has much more saturated colours than standard colour composites as illustrated in Figure 9.8. This is due to the decorrelation between PC channels. In terms of the RGB colour cube described in Section 8.1, EO imagery generally displays little variation about the intensity axis, with no potential to form the subtractive primary colours at the corners of the colour cube (see Figure 9.9). A colour composite comprising the first three PC channels however, effectively fills the volume of this cube so that both additive and subtractive primaries are visible in the displayed imagery (see Excursus 9.2).

Figure 9.8 Principal Component transformation

Landsat TM image over Flinders Ranges, South Australia

a. Original false colour composite is displayed using bands 4, 3 and 2 as RGB.

b. Principal component image based on PCs 1, 2 and 3 highlights lithology, soils and vegetation.

Source: Megan Lewis, University of Adelaide

Decorrelation stretching provides a useful visual enhancement for increasing colour saturation in composite imagery (Gillespie *et al.*, 1986). As illustrated in Figure 9.10 this process involves:

- transforming image channels to the PC space;
- stretching the image data along the PC axes to take on a spherical distribution; then
- transforming the data back onto the original channel axes.

This enhancement is particularly useful for visual interpretation of EO imagery in a geological application (Rothery, 1987). The back transformation can also be applied with one component removed. This is useful for removing certain spatial noise patterns from imagery where the noise can be isolated as one component.

Figure 9.9 RGB colour cube showing original image data range

The correlation between channels in EO imagery generally means that the subtractive primary colours (yellow, magenta and cyan) do not occur in colour composite images.
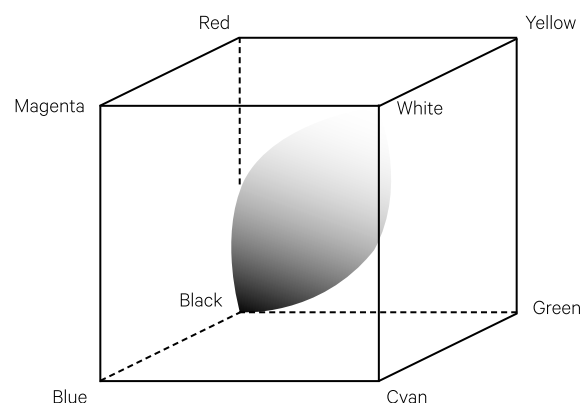
Source: Harrison and Jupp (1990) Figure 90

Figure 9.10 Decorrelation stretching

Images can be transformed to the principal component space for contrast stretching then back transformed to the original image channels. By rescaling the data in an uncorrelated data space, the resulting composite image has enhanced colour contrast.
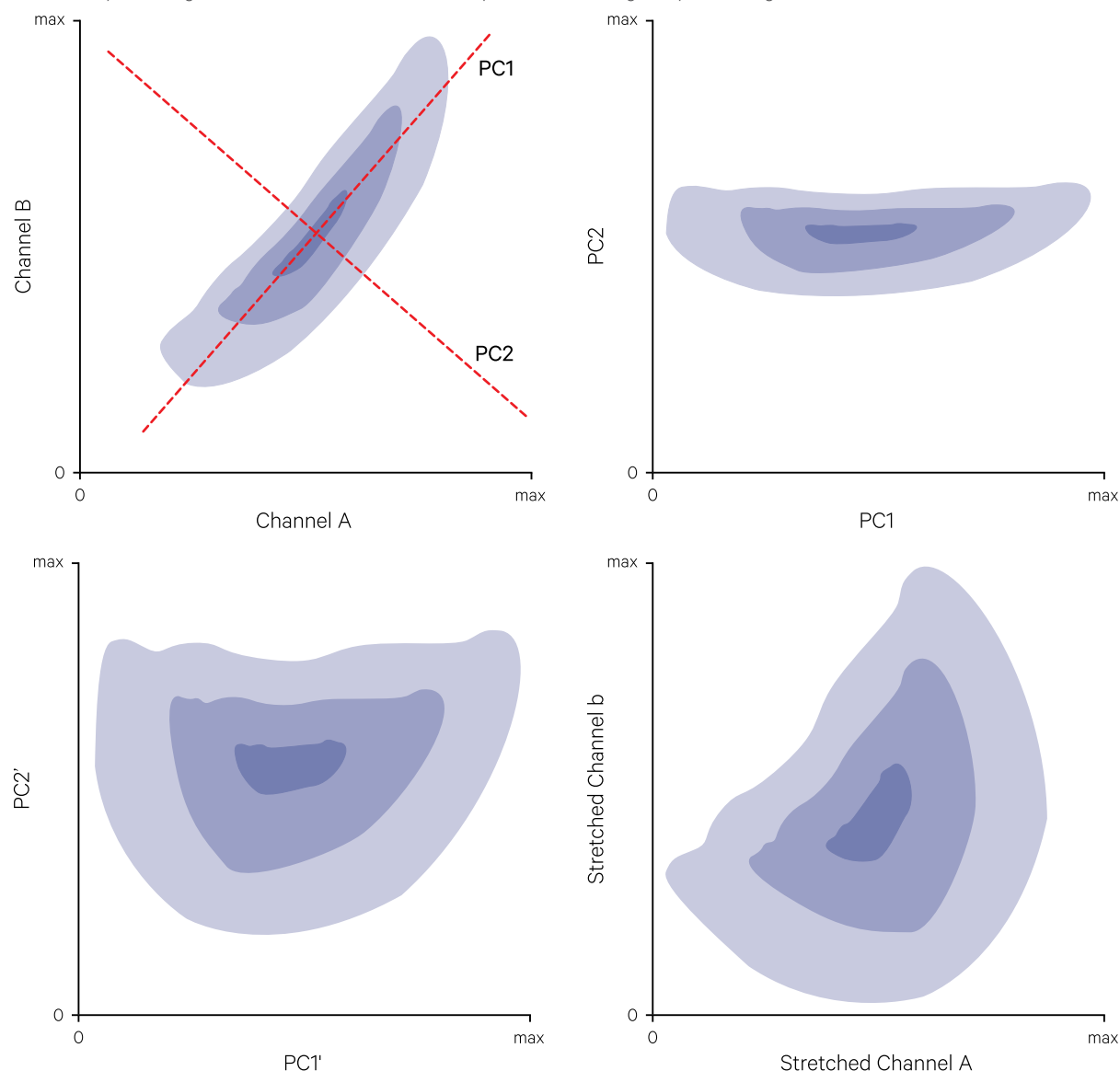
Source: Harrison and Jupp (1990) Figure 91 [Adapted from: Gillespie *et al.* (1986)]

9.5 Data Reduction

Most EO images contain some redundant information. For example, Landsat TM2 and TM3 bands contain similar data values since most ground cover types give similar responses in these two channels (that is, they are highly correlated). When redundancies do occur in an image, it is possible to 'remove' the duplicate information and represent the data in fewer channels. For many EO datasets, the vast majority of data variation contained within multiple image channels can be represented by their first few PCs (Chuvieco and Huete, 2010; Jensen, 2016).

PCA is principally used for data reduction, although frequently this is an intermediate step in other processes such as image classification, enhancement or integration. By representing the image data in fewer channels (that is, reducing the data dimensionality), various operations may be applied to the transformed channels before transforming back to the original data space. Such data reduction techniques are particularly relevant to hyperspectral imagery. Analysis of 128 channel HyMap airborne imagery, for example, showed that:

- PCs 1–4 contained 99% of the image variation;
- PCs 5–50 mostly contained local variation information and anomalies; and
- PCs 50–128 mostly contained instrument noise, although may also contain useful information about anomalies or small features (Megan Lewis, *pers. comm*; see Volume 2D)

This data reduction operation is particularly relevant to geological applications (where feature identification relies on visual interpretation of enhanced imagery) and hyperspectral imagery (with numerous, often correlated bands). In image classification, PC1 and PC2 provide a useful basis for autoseeding methods based on defining channel themes (see Volume 2A— Section 9.1). Volume 2E describes this procedure in more detail. Once the image has been classified using a set of basic classes defined this way, they may be transformed back to the original image data channels using the 'spectral transfer' operation described in Volumes 2D and 2E.

A sentence should contain no unnecessary words,
a paragraph no unnecessary sentences,
for the same reason that a drawing should have no unnecessary lines
and a machine no unnecessary parts.
(William Strunk, Jr.)

9.6 Highlighting Specific Features

PCA can be advantageous for isolating specific image features as individual PCs. Some applications of PCA to EO imagery include mapping soils (Mitternicht and Zinck, 2003; Viscara-Rosel and Chen, 2011) and detecting cloud (Amato *et al.*, 2008).

Fraser and Green (1987) describe an interesting technique called 'directed PCA' (DPCA), which is intended to map specific image information to a given PC. This technique involves calculating PCs on two input ratio channels. For geological analyses, these ratio channels are selected so that one is a geological

discriminant that is often confused with vegetation (such as 1.65 μm/2.22 μm; see Figure 9.11a) and the other is primarily a vegetation index (such as NIR/red; see Figure 9.11b). PCA of these two channels produces two uncorrelated data axes: the first aligned with vegetation and second with non-vegetation (see Figure 9.11c). For geological analyses this technique is reported to substantially reduce the effects of vegetation in woodland environments. In this example, vegetation is visibly suppressed along drainage lines.

Figure 9.11 Software defoliant

A Landsat-5 image (path 95, row 73) over Charters Towers region, Queensland, acquired on 7 October 1987, was processed to produce two ratio channels, one highlighting clay soils and the other highlighting vegetation (see also Figure 8.10). When PCA is applied to these two ratio channels, PC2 effectively maps non-vegetation, so can be considered as a 'software defoliant'. In this case, suppression of vegetation is most evident in drainage channels.
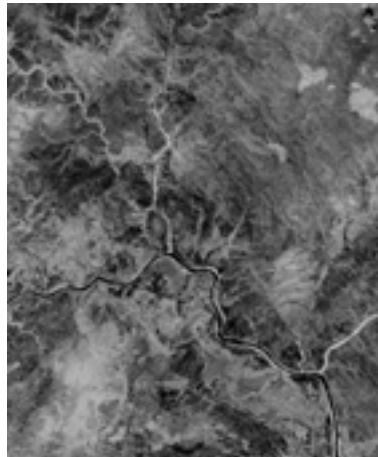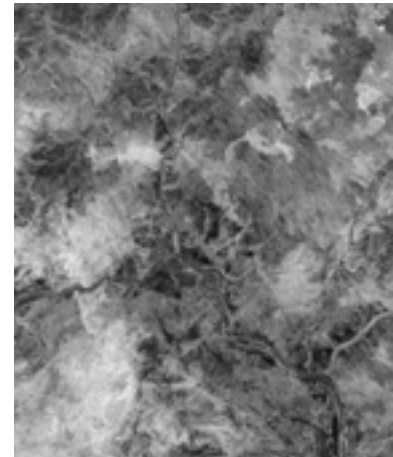
a. Ratio channel highlighting clay soils
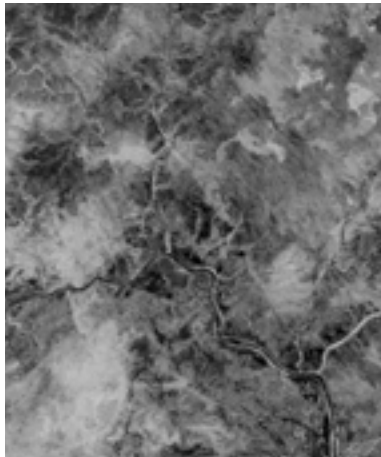
b. Ratio channel highlighting vegetation

c. PC2 highlighting non-vegetation

Source: Tony Sparks, Icon Water

9.7 Multi-scale Intensity Enhancement

Panchromatic imagery with finer spatial resolution than a multispectral image can be integrated with the multichannel data to visually enhance its spatial detail. This pan-sharpening process is introduced in Section 8.4 in terms of the RGB-HSI transformation (see Figure 8.7).

Various methods have been suggested for integrating image pairs with differing spatial resolutions (such as Cliche *et al*., 1985 and Price, 1987a). PCA can also be used for pan-sharpening a spatially-registered panchromatic/multispectral image pair as part of the following processing sequence:

- resample multispectral image to the pixel size of the panchromatic image (see Volume 2A—Section 6, and Volume 2B);
- transform selected channels in the resampled multispectral image to PCs;
- replace PC1 with the registered panchromatic channel (since PC1 represents the brightness or intensity variations in EO data). The panchromatic channel should first be rescaled to match PC1 using the regression technique described in Section 7.2.1 (see also Volume 2D); and
- back-transform the resulting image to the original multispectral image channels to produce an image in which the spatial detail of the panchromatic channel is supplemented by the colours of the coarser composite image.

The back transformation matrix will need to be adjusted to account for the scaling between the actual PC ranges and their representation in the PC image. This matrix is implemented as an affine transformation (see Section 7.1).

This approach offers a very flexible method of merging new information with any spatial dataset, since the new 'vector' (or channel) can be used to replace any existing vector in the original data. However, for replacing the intensity component of a colour image, the direct method discussed in Section 8.4 is most appropriate.

Most pan-sharpening applications involve panchromatic and multispectral images from the same sensor (such as SPOT MSS and SPOT Pan). Procedures involving PCA and multivariate regression have also been suggested for substituting signals from one sensor for signals in another (such as Suits *et al*., 1988). If required, such methods can also be used to transfer spectral interpretation techniques from one sensor to another. Data fusion approaches are further discussed in Volume 2D in the context of image integration.

103

9.8 Change Detection

PCA is frequently used with multi-temporal imagery as a change detection method, by either:

- transforming two or more images separately to PCs, then selecting the most appropriate PCs for comparison; or
- registering multiple images and transforming them to PCs as a merged dataset.

The use of PCA for change detection studies is further discussed in Volume 2D.

9.9 Further Information

Chuvieco and Huete (2010) Section 7

Gonzalez and Woods (2018) Section 12

Jensen (2016) Section 8

9.10 References

Amato, U., Antoniadis, A., Cuomo, V., Cutillo, L., Franzese, M., Murino, L., and Serio, C. (2008). Statistical Cloud Detection from SEVIRI Multispectral Images. *Remote Sensing of Environment*, 112, 750–766.

Chuvieco, E., and Huete, A., (2010). *Fundamentals of Satellite Remote Sensing*. CRC Press, Boca Raton.

Cliché, G., Bonn, F., and Teillet, P. (1985). Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement. *Photogrammetric Engineering and Remote Sensing*, 51, 311–316.

Fraser, S.J., and Green, A.A. (1987). A software defoliant for geological analysis of band ratios. *International Journal of Remote Sensing*, 8, 525–532.

Gillespie, A.R., Kahle, A.B., and Walker, R.E. (1986). Colour enhancement of highly correlated images. I. Decorrelation and HSI contrast stretches. *Remote Sensing of Environment*, 20(3), 209–235.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Metternicht, G.I., and Zinck, J.A. (2003). Remote Sensing of Soil Salinity: Potentials and Constraints. *Remote Sensing of Environment*, 85, 1–20.

Price, J.C. (1987a). Combining Panchromatic and Multispectral Imagery from Dual Resolution Satellite Instruments. *Remote Sensing of Environment*, 21, 119–128.

Price, J.C. (1987b). Calibration of satellite radiometers and the comparison of vegetation indices. *Remote Sensing of Environment*, 21, 15–27.

Rothery, D.A. (1987). Decorrelation stretching as an aid to image interpretation. *International Journal of Remote Sensing*, 8, 1253–1254.

Suits, G., Malila, W., and Weller, T. (1988). Procedures for using signals from one sensor as substitutes for signals of another. *Remote Sensing of Environment*, 25, 395–408.

Viscarra-Rossel, R.A., and Chen, C. (2011). Digitally Mapping the Information Content of Visible-Near Infrared Spectra of Surficial Australian Soils. *Remote Sensing of Environment*, 152, 1–14.

104

Earth Observation: Data, Processing and Applications. Volume 2: Processing

Non-Linear Operations
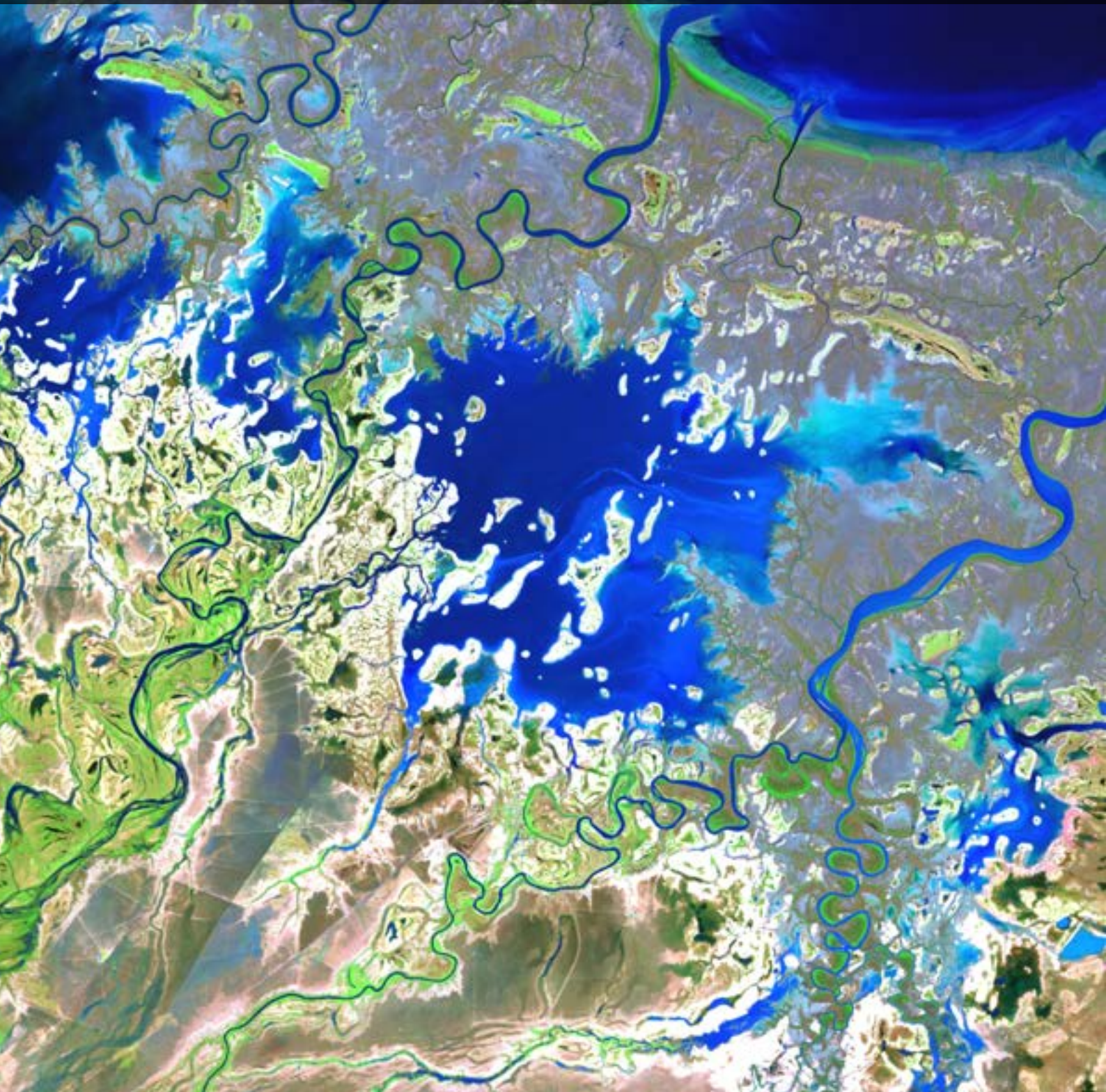
Sections 7, 8 and 9 discuss linear operations that are commonly applied to one to more EO image channels. In this context, addition and subtraction of channels are examples of simple linear operations. The following two sections are concerned with non-linear operations applied to pairs of channels in an image (see Section 1.4).

Non-linear operations include multiplication and division, with the most frequently used non-linear operation for image processing of EO imagery being ratioing of image bands (see Section 10). Multiplying image bands is not a commonly used operation, but may be appropriate for masking out regions of an image (Castleman, 1998) or for image modulation (Harrison and Jupp, 1990).

Many of the band indices that are popularly used with EO imagery rely on a combination of linear and non-linear operations. Some of the indices that have been developed to highlight and condense information in EO imagery relating to vegetation condition are introduced in Section 11. These indices are discussed in more detail, and in the context of specific application areas, in Volume 3A.

Contents

10 Channel Ratios

A simple but effective transformation that is commonly used with EO imagery involves computing the ratio values between two image channels. The simplest implementation of this transformation involves dividing the value of a pixel in one channel by its value in a second channel and rescaling the result to the available image data range for all pixels in the image. The ratio value resulting from this calculation is generally rescaled according to user-defined parameters to fill the full image data range (see Section 2).

In this section ratioing of image channels will be discussed in terms of:

- computational methods—simple ratios, log ratios, smoothed ratios and directed ratios (see Section 10.1); and
- processing parameters—dark values and calibration impact (see Section 10.2).

Firstly, some ratios that are frequently used with EO imagery are introduced in Excursus 10.1.

Transformation literally means going beyond your form.
(Wayne Dyer)

Background image: Landsat TM colour ratio image over Flinders Ranges, SA, where the ferric iron (iron oxide) ratio (TM3/TM1) is shown as red, the ferrous minerals ratio (TM5/TM4) is shown as green, and the clay minerals ratio (TM5/TM7) is shown as blue. **Source:** Megan Lewis, University of Adelaide

Excursus 10.1—Commonly Used Ratios

While the ratioing of image channels is a very simple operation computationally, it is a very effective way of highlighting the differences between two channels. Ratios of various channels are used to highlight differences in particular features of the Earth's surface. As introduced in Volume 1, different land covers reflect and absorb EM radiation in different and often characteristic ways (see Figure 10.1 and Section 1.1). When image channels are ratioed, these differences are emphasised.

For example, healthy vegetation typically reflects NIR wavelengths strongly, but absorbs most wavelengths in the visible red region of the EM spectrum, whereas while NIR reflectance is relatively strong for most soils, the difference between their NIR and red reflectances is less dramatic. These differences are highlighted in a single channel when the ratio of NIR and red bands are computed. In this case a NIR/red ratio channel would allow soil and vegetation features to be clearly distinguished in the image.
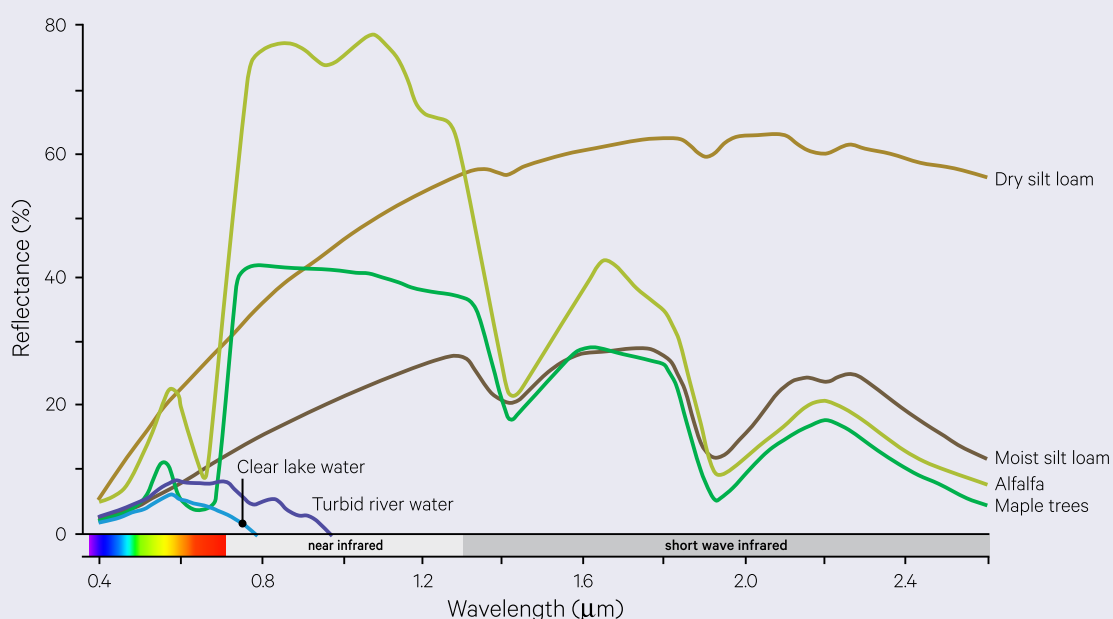
Some of the channel ratios that are commonly computed for EO imagery are summarised in Table 10.1. These ratios have been computed for a Landsat-8 OLI image of Lake Carnegie, WA. The original Landsat-8 OLI channels are shown in Figure 10.2 and the ratio channels are shown in Figure 10.3. Whereas several original pairs of channels (such as blue and green, and SWIR1 and SWIR2) are similar, the ratio channels highlight the differences between channels.

Table 10.1 Popular EO image channel ratios

Ratio channels		Highlights
Numerator	Denominator	
Green	Red	Soil colour / Water colour
Green	Blue	Water colour
NIR	Red	Vegetation with green foliage
Red	Blue	Ferric iron
SWIR1 (~1.65 μm)	NIR	Ferrous minerals
SWIR1 (~1.65 μm)	SWIR2 (~2.22 μm)	Geological features

Figure 10.1 Idealised spectral reflectance curves

Reflectance from different Earth surface features varies for different wavelengths. Characteristic spectral reflectance curves can be created for different features to indicate a 'typical' shape over a particular range of wavelengths. This information can be used to differentiate between surface features on the basis of their observed reflectance characteristics.

Adapted from: Harrison and Jupp (1989) Figure 6

108

Earth Observation: Data, Processing and Applications. Volume 2: Processing

Figure 10.2 Landsat-8 OLI image channels

Six original image channels are shown for a Landsat-8 image over Lake Carnegie, WA, acquired on 8 July, 2017.

a. Blue channel (band 2: 0.452–0.512 μm)
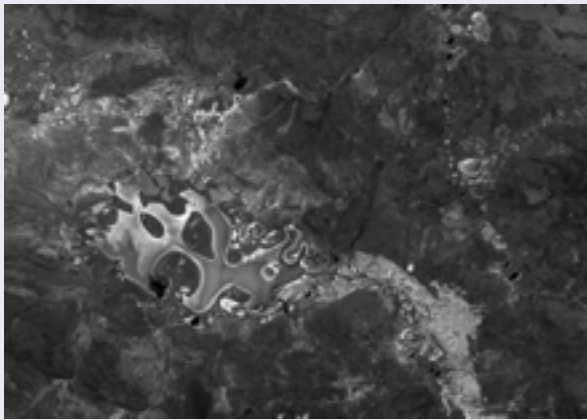
b. Green channel (band 3: 0.533–0.590 μm)
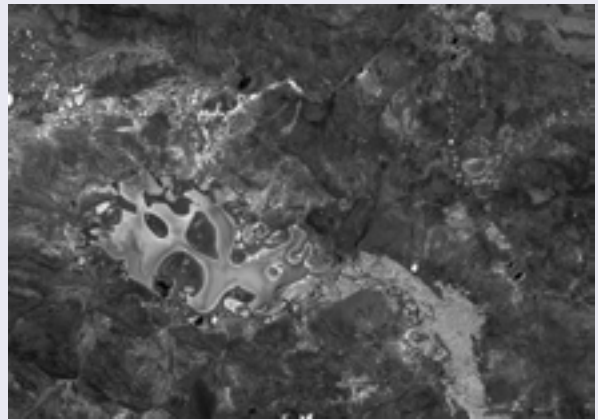
c. Red channel (band 4: 0.636–0.673 μm)
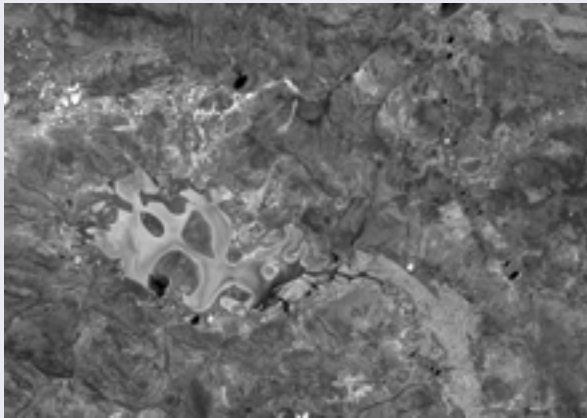
d. NIR channel (band 5: 0.851–0.879 μm)
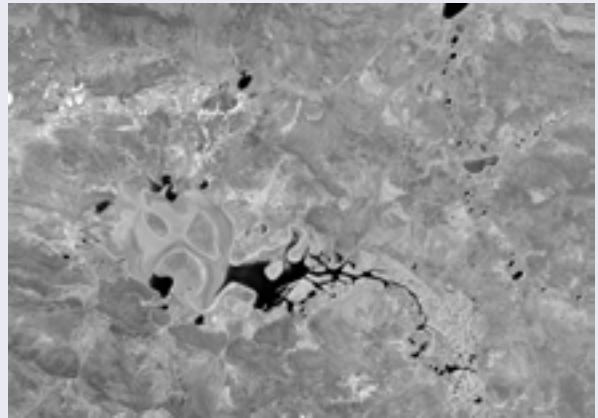
e. SWIR1 channel (band 6: 1.566–1.651 μm)
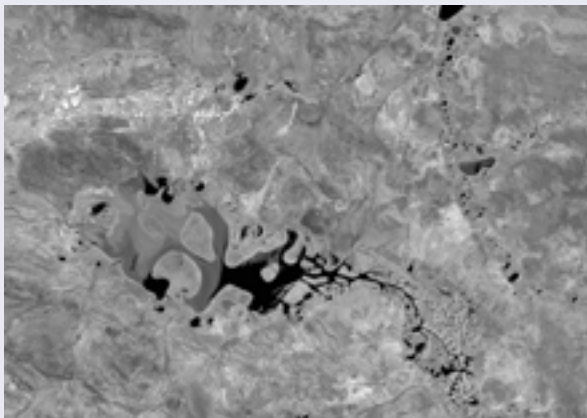
f. SWIR2 channel (band 7: 2.107–2.294 μm)
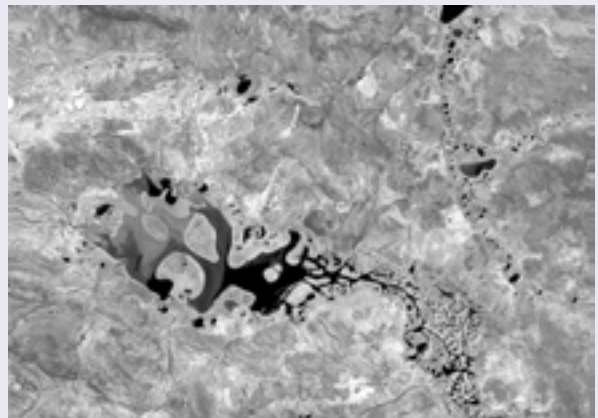
Source Norman Mueller, Geoscience Australia

Figure 10.3 Ratio channels

Several ratio channels have been computed for the image channels shown in Figure 10.2.

a. Ratio of green/red
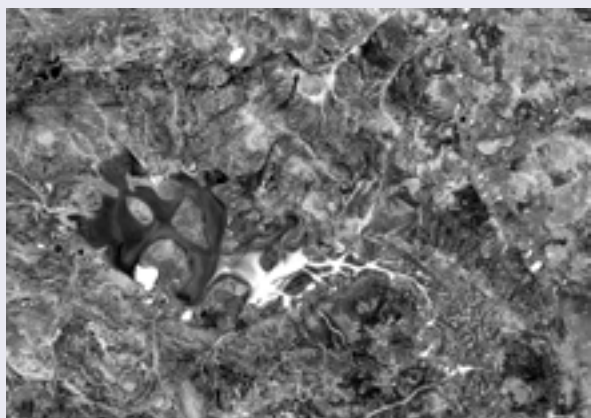
b. Ratio of green/ blue
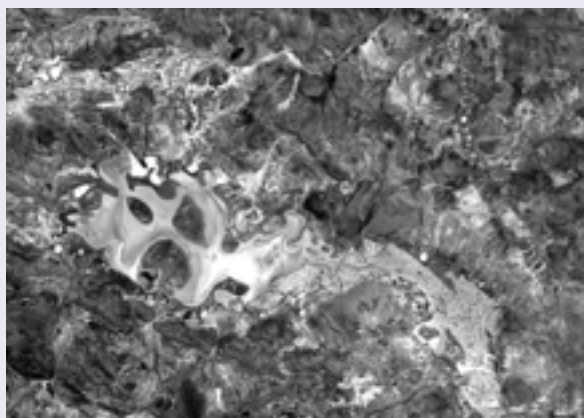
c. Ratio of NIR/red
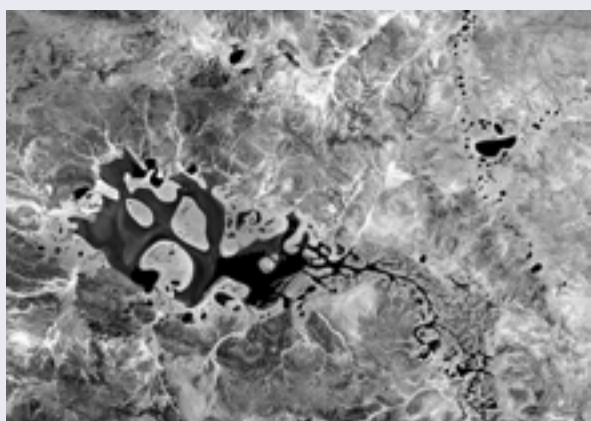
d. Ratio of red/ blue
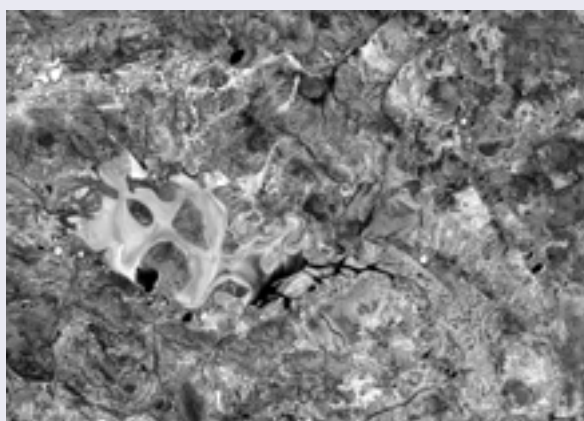
e. Ratio of SWIR1/NIR
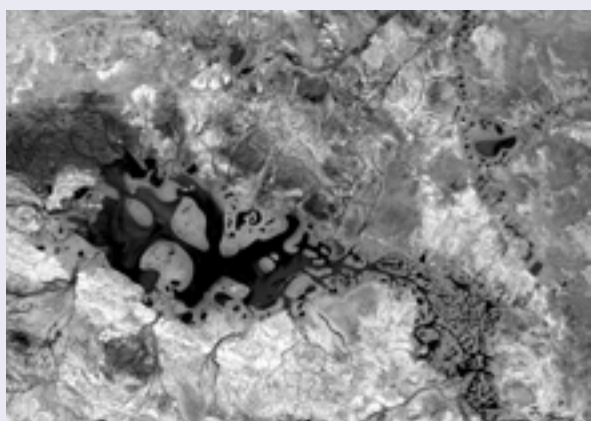
f. Ratio of SWIR1/SWIR2
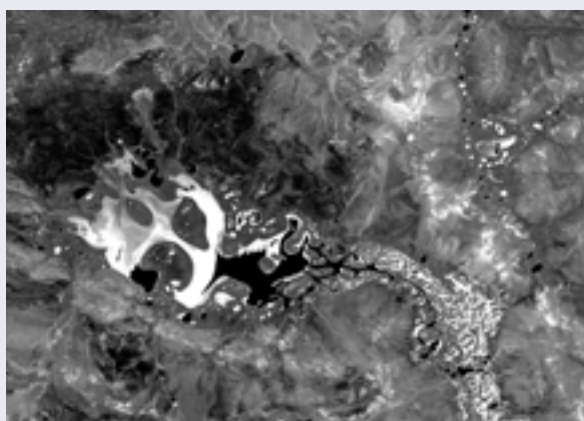
Source Norman Mueller, Geoscience Australia

Colour composites of various ratio channels may be used to highlight particular surface features. For example, to maximise lithological spectral information in well-exposed terrain Crippen (1988) suggests a colour composite of selected Landsat TM ratio channels (see Figure 1.2 for band definitions):

- TM5/TM7 (SWIR1/SWIR2), which highlights clay (see Figure 10.4a), as red;

- TM5/TM4 (SWIR1/NIR), which highlights ferrous minerals, as green; and

- TM3/TM1 (red/ blue), which highlights ferric iron (iron oxide; see Figure 10.4b) as blue.

A variation on this combination is illustrated in Figure 10.5.

110

Earth Observation: Data, Processing and Applications. Volume 2: Processing

Figure 10.4 Ratios to highlight lithology

Two ratio channels are computed for a Landsat TM image of a portion of Flinders Ranges, SA.

a. Clay minerals (SWIR1/SWIR2)

b. Iron oxide (visible red/visible blue)
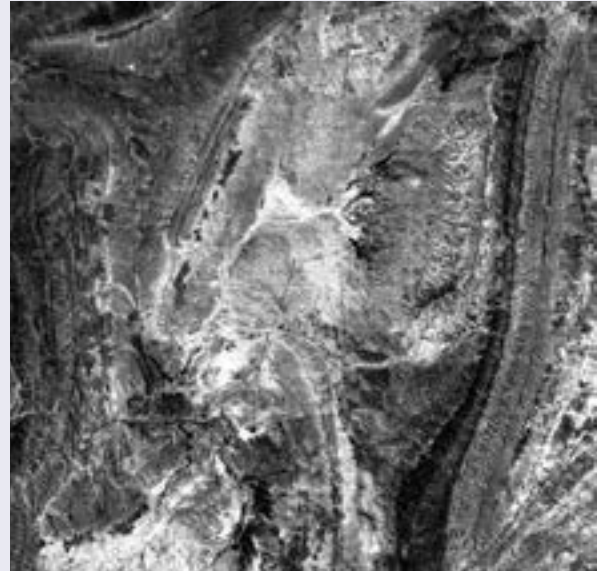
Source: Megan Lewis, University of Adelaide

Satterwhite and Henley (1987) evaluated a number of band ratioing techniques with Landsat TM data for discriminating spectral characteristics of arid region soils and vegetation conditions. Vogelmann and Rock (1988) found Landsat TM channel ratios of TM5/TM4 (SWIR1/NIR) and TM7/TM4 (SWIR2/NIR) correlated well with ground measurements of forest damage (% foliage loss in spruce-fir forests). A composite of the Landsat MSS bands MSS4/MSS5 (green/red) as blue, MSS5/MSS6 (red/NIR1) as green and MSS6/MSS7 (NIR1/NIR2) as red has been reported to enhance discrimination between altered rock types by many authors (such as Moik 1980; Drury 1987).

Chavez *et al.* (1982) discuss the selection of ratio channels to form colour composites for geological applications. These authors suggest an index of:

$$\frac{\sum \sigma}{\sum \text{absolute correlation values}}$$

where

σ is the correlation coefficient (normalized covariance; see Section 1.3.3 and Volume 2A—Section 8.1.4)

be computed for each possible combination of channels, then the combination with the largest index value be selected. The resulting composite image would have the most variation within, and least correlation between, channels so should be the most visually informative (see Volume 2A—Section 5.2.2). This method confirmed the popular ratio composite suggested by Crippen (1988).

Figure 10.5 Colour composite of band ratios

Landsat TM colour ratio image over Flinders Ranges, SA, where the ferric iron (iron oxide) ratio (TM3/TM1) is shown as red, the ferrous minerals ratio (TM5/TM4) is shown as green, and the clay minerals ratio (TM5/TM7) is shown as blue.
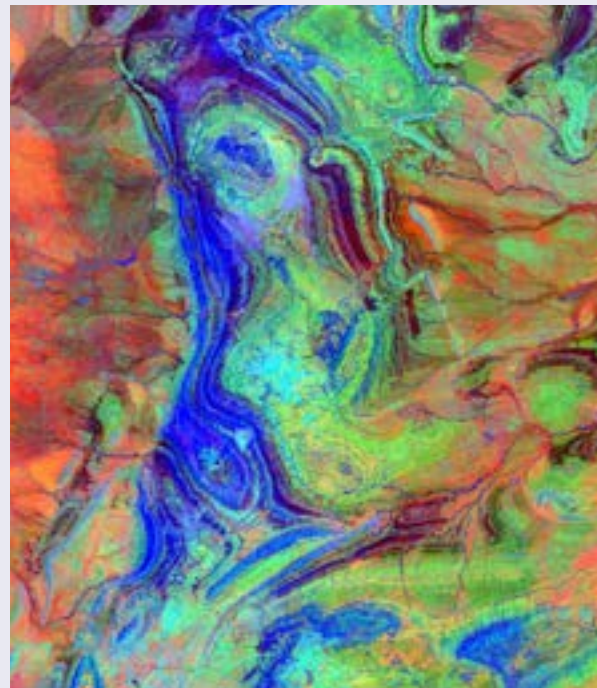
Source: Megan Lewis, University of Adelaide

10.1 Computation Methods

Channel ratios may be computed in a variety of ways. For example, a simple ratio may be computed between pairs of channels as shown in Figure 10.6 (see Section 10.1.1). Ratios of a pair of image channels may also be computed by differencing the logarithms of the channels (see Section 10.1.2). Alternatively a 'smoothed ratio' may be computed as detailed in Section 10.1.3. This produces a more 'stable' result by removing small variations in image values that would be accentuated by the ratioing process. Finally, a modified approach to computing channel ratios, called 'Directed Band Ratioing' (Crippen *et al.*, 1988), which enables the ratioed values to retain intensity information, is described in Section 10.1.4.

Figure 10.6 Operation of simple channel ratio

a. Process—each pixel value in channel a is divided by its corresponding pixel value in channel b then the resulting ratio value is rescaled in the output channel.

b. In this numeric example, the minimum ratio value ($\frac{1}{7}$) is rescaled to output channel value 0 and the maximum ratio value ($\frac{9}{3}$) is rescaled to value 100. In the output channel, values are computed as $\dfrac{\left(\text{ratio}-\frac{1}{7}\right)\times 100}{\left(\frac{9}{3}-\frac{1}{7}\right)}$

Source: Harrison and Jupp (1990) Figure 93

10.1.1 Simple ratios

Most image processing systems can construct multiple ratio channels from pairs of input channels. For each pair of corresponding pixels in the input images, this ratio is computed as:

$$\frac{x-a}{y-b}$$

where

- x and y are the paired pixel values in the input channels; and
- a and b are the respective 'dark' values for these channels.

The 'dark' value for an image channel represents the minimum 'data' response in that channel and corresponds to the darkest feature in the image. As detailed in Section 10.2.1, it provides an origin for computing the ratio of two channels such that values below the dark value are assumed to be due to atmospheric effects and/or sensor 'noise' (see Volume 1A—Section 13).

The ratio values resulting from this calculation are generally rescaled so that they fill the full image data range in the output image (see Figure 10.6b). The optimum rescaling parameters may need to be determined iteratively, and should be recorded if the ratio channel is to be subsequently used with models that relate ground measurements to image radiance (see Section 2). A wide distribution, but low frequency, of image values close to the minimum and maximum values in a channel can sometimes make this rescaling

operation difficult. This data-dependent problem can be resolved by reference to the ratio image histogram (see Section 1.3.2).

When a ratio channel has been computed to represent a larger range of ratios than is necessary for most pixels in the image, a suitable ratio range can be selected using the channel histogram then rescaled back to the actual ratio values for use in another processing run. This rescaling requires the expected ratio range (*a,b*) that was used to compute the ratio channel. Thus to rescale value *P* in the ratio channel to an actual ratio value we would compute:

$$\frac{P}{\text{image data range}} \times \left(b - a\right) + a$$

where

> *a* is the minimum ratio value used to rescale the ratio channel; and
>
> *b* is the maximum ratio value used to rescale the ratio channel.

The simple ratio transformation is computed quite rapidly and may be used to quickly determine ratio results for multiple channel pairs. The smoothed ratio, described in Section 10.1.3, is recommended for computation of the final ratio channel for feature analysis because the smoothed ratio process does not accentuate localised differences between channels as typically occurs in simple ratio channels.

Simple ratios are required however for processes which are attempting to reduce spatial 'noise' in an image. This noise may be due to uneven illumination effects, such as topographic shading. Since the latter has a similar effect on all image channels, shading differences are greatly reduced in ratioed channels (see Figure 10.7). This effectively means that shadow in one channel is being divided by shadow in another channel so that the resulting ratio channel retains significantly less 'shadow' information.

The ratio transformation effectively enhances differences between the channels and reduces similarities. Holben and Justice (1981) examined band ratioing as a technique to reduce topographic shading in Landsat MSS imagery, and reported a reduction in topographic effect by an average of 83%, with variations in this figure being due to Sun angle and surface slope.

Uneven illumination can be a problem with scan-digitised imagery. This may be compensated for in the same way as haze correction by using the scan of a blank target as the denominator or 'reference' channel (see Section 7.2.1). In this case, since the range of the reference channel is usually quite small, to preserve the data range of the image channels they should not be rescaled before ratioing and dark values should be given as zero.

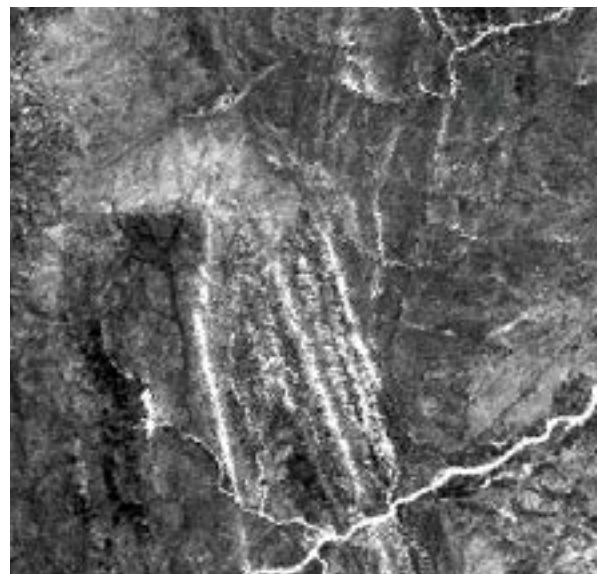**Figure 10.7** Impact of ratioing on topographic shading

Topographic shading in EO imagery is significantly reduced in ratio channels. This Landsat TM image was acquired in June 1993 over Fowlers Gap Research Station, NSW.

a. Landsat TM band 4 shows marked topographic shadowing.

b. Ratio of Landsat TM4/TM7 reduces topographic shadowing in this image.

Source: Megan Lewis, University of Adelaide

10.1.2 Computing Products and Ratios using Log Channels

Ratioing of values in a linear scale is equivalent to the differencing of the corresponding values in a logarithmic scale (see Section 1.2.3), that is:

$$a \times b = \log(a) + \log(b)$$

and

$$\frac{a}{b} = \log(a) - \log(b)$$

Thus ratioing of image channels may be computed as the difference between log-transformed channels (see Section 2.2). The differencing operation could be computed using an affine transformation to produce log-ratio channels (see Section 7). When channel ratios are computed this way, dark values are specified as negative scaling (offset) values in the logarithmic transformation (see Section 10.2.1). The products of image channels can be derived in a similar way by adding logarithmically-transformed channels.

Log-ratio channels provide a useful enhancement for reducing topographic shading effects, since the log transformation reduces variation in the brightest areas of the image and provides greater contrast in the darker features. For example, Ahmad *et al.* (1989) used log-ratio channels to reduce the effect of topographic shading in Landsat MSS imagery of a mountainous area in northeast Tasmania without a digital terrain model.

10.1.3 Smoothed ratios

As well as computing the simple ratio result of individual pixel values in a pair of channels, it is possible to apply a smoothing process during the ratio calculation. This operation reduces the 'noise-enhancement' effects, which can occur in ratio channels due to small variations between channels. Gordon and Philipson (1986) suggest that ratioing selected texture channels would reduce their characteristic edge effects. In this case, a smoothed ratio would avoid increasing within-feature texture.

Some image processing systems can directly compute a 'smoothed ratio' channel from a pair of image channels. For example, the 3×3 neighbourhood of each pixel may be used to estimate the ratio as the solution to the linear equation:

$$c_i - d_i = r_{ij} \times (c_j - d_j)$$

where

c_i and c_j are the channels i and j being ratioed;
d_i and d_j are dark values (see Section 10.2.1) in channels i and j ; and
r_{ij} is the estimated ratio value.

This operation is illustrated in Figure 10.8 (using the same minimum and maximum values as Figure 10.6 to enable comparison of results). The processing order, the non-linear nature of the ratioing operation, and the rescaling differences between processes, mean that this transformation produces slightly different results to either computing simple ratios of smoothed data channels or the smoothing of simple ratio channels using the same filter weights.

Some implementations of this process may also allow themes to be applied to the channels being ratioed. This option is a useful way of pre-segmenting the image into feature and non-feature areas and then computing a ratio value only for those pixels that are deemed to represent the selected feature. Pixels outside the specified theme are typically set to the default null value in the output image. These may be subsequently 'filled' using filter-based interpolation (see Section 4.6) if appropriate. However, it is usually recommended that broad, rather than narrow, themes be defined for this purpose to ensure all feature pixels are processed.

The additional computation required by the smoothing process means that this transformation is likely to be slower than the simple ratios described in Section 10.1.1. However, the results of this process provide a more 'stable' set of values, which highlight the major rather than minor variations between the channels. Thus, when ratio channels are being computed for density slicing or other interpretation purposes, it is highly recommended that they be produced as smoothed rather than simple ratios.

Figure 10.8 Smoothed ratio operation

a. Process—the 3x3 neighbourhood around each pixel pair is used to compute the smoothed ratio value, which is then rescaled in the output channel. Edge pixels have not been processed in this example.
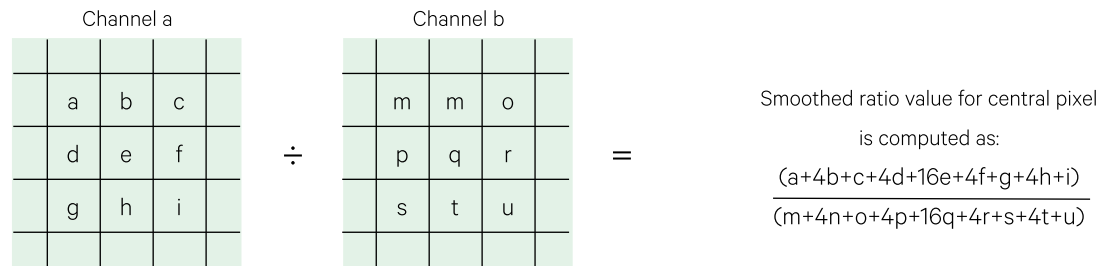
b. For this numeric example, the minimum ratio value ($\frac{1}{7}$) is rescaled to output channel value 0 and the maximum ratio value ($\frac{9}{3}$) is rescaled to 100. In the output channel, values are computed as $\dfrac{\left(\text{smoothed ratio} - \frac{1}{7}\right) \times 100}{\left(\frac{9}{3} - \frac{1}{7}\right)}$
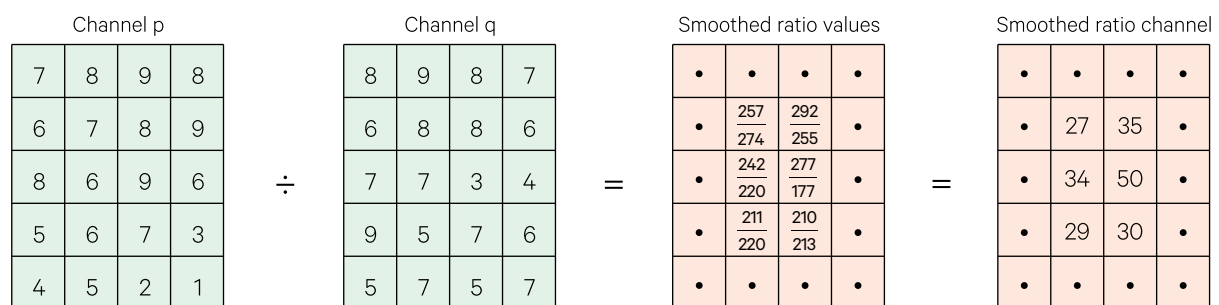
Source: Harrison and Jupp (1990) Figure 97

10.1.4 Directed band ratioing

Crippen *et al*. (1988) suggest a method called 'Directed Band Ratioing' (DBR) by which topographic and albedo information can be retained in colour composites formed from ratio channels. Various ratio channels are useful enhancements of geological information and the retention of geomorphic structure as relief shading can improve the interpretation of such imagery.

Directed band rationg involves retaining the illumination information as a consistent intensity variation in all ratio channels, so that, in a colour composite, the illumination components are correlated (channel invariant) and thus only vary as a grey-scale, while reflectance (possibly lithographic) variations are uncorrelated (channel variant) between ratios so are shown as colours. Thus, while topographic variations may mask reflectance information in individual channels, correlation of illumination variation between channels in a multi-channel composite separates topography as an achromatic feature from other coloured features. (This correlation could also be advantageous in multidimensional analyses such as classification or PCA to isolate topographic or albedo variations).

The procedure for directed band ratioing involves pre-processing the channels by:

- subtracting appropriate dark values (see Section 10.2.1);
- multiplying the denominator channel by a constant;
- adding a constant to both channels; then
- ratioing the results.

The principle of this processing is to effectively increase the ratio 'angle' of reflectance variations relative to illumination variations in a channel crossplot (see Section 1.3.4 and Figure 10.9). The required pre-processing may be undertaken using an affine transformation (see Section 7.1) before ratioing.

Figure 10.9 Directed band ratioing

This process increases the ratio 'angle' of reflectance variations in the image relative to the illumination variations. This preserves the topographic shading information, which is generally reduced by ratioing, but can provide useful information for visual interpretation, especially for geological applications.

a. Crossplot of original data channels
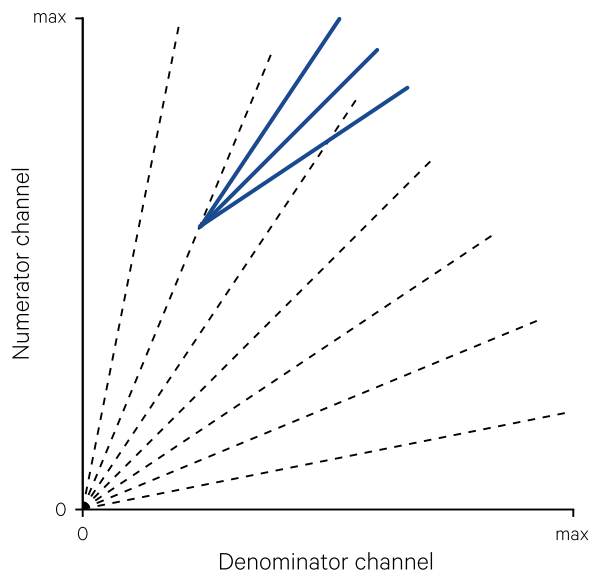
b. Ratios after dark value adjustment
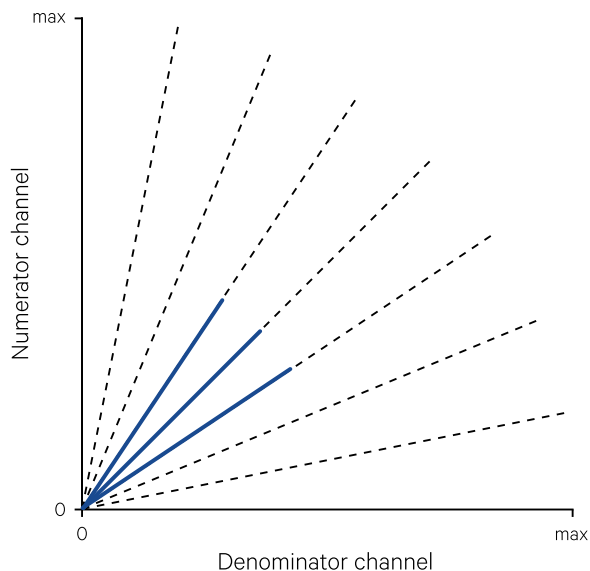
c. The slope of data lines is increased by multiplying the denominator channel by a constant value.
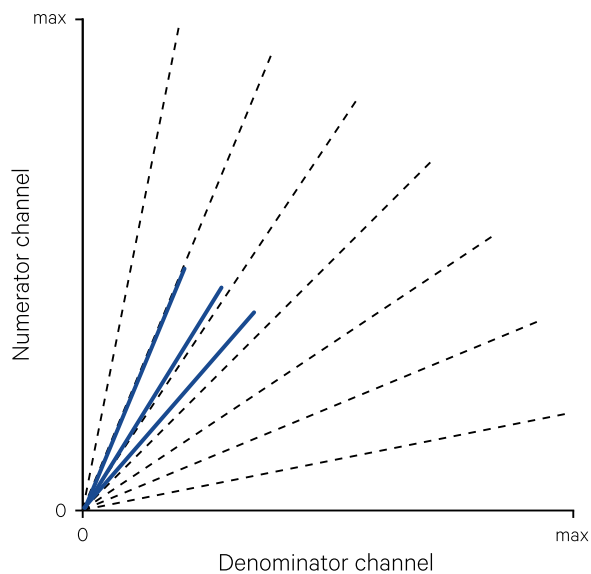
d. Equal constants are added to each channel before ratioing
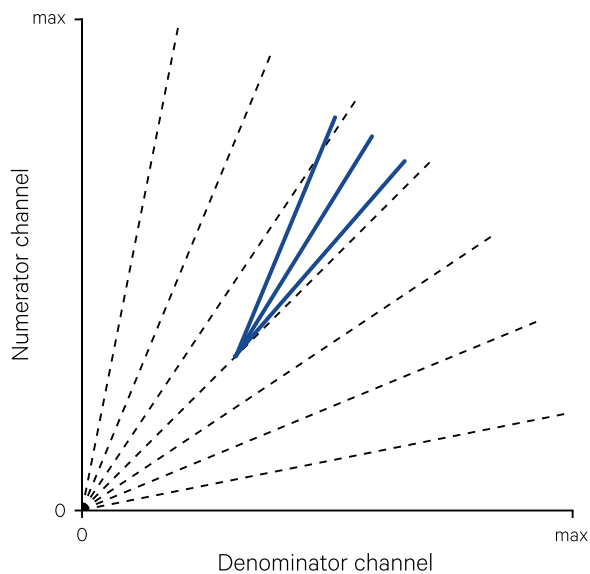
Source: Harrison and Jupp (1990) Figure 96 [Adapted from Crippen *et al.,* 1988]

10.2 Processing Parameters

Ratio variables are characterised by having a valid zero point (see Section 1.2.2). When ratioing a pair of channels in EO imagery, computations need to account for the true minimum value in each channel. This minimum value is commonly called the 'dark' value (see Section 10.2.1) as it effectively accounts for sensor 'dark current' (see Volume 1A—Section 13). When the channels in an image are not scaled consistently (due to variations in gains and/or calibration) the expected relationships between channels may not be observed in ratio results. Models to account for these variations are introduced in Section 10.2.2.

10.2.1 Dark values

Ratio transformations generally subtract a constant 'dark value' (d_j) from each channel before the ratioing process. In EO data, the dark value level is assumed to be the upper threshold of the response in a channel that is due to atmospheric effects and sensor calibration offsets. This represents the minimum level of information in the channel (see Volume 1A—Section 13).

Dark values are so named as the values corresponding to the darkest features in the image, such as shadows (Crane, 1971), and may be derived by interactive training on such features (see Volume 2A—Section 9.1.2). A simple approximation of the dark value is one value less than the absolute minimum value of a channel. More sophisticated methods for determining dark values to correct for atmospheric path radiance were proposed by Switzer *et al*. (1981), Crippen (1987) and Chavez (1988).

Crippen (1988) demonstrates the importance of appropriate selection of the dark value adjustment for producing useful ratio channels (see Figure 10.10). When image channels are not properly adjusted before ratioing, the resulting channel can contain undesirable overlap between the ratio ranges of different image features..

Figure 10.10 Effect of dark value adjustment in channel ratios

a. The origin of ratios, and hence their stratification of the image data space, is determined by dark value selection.
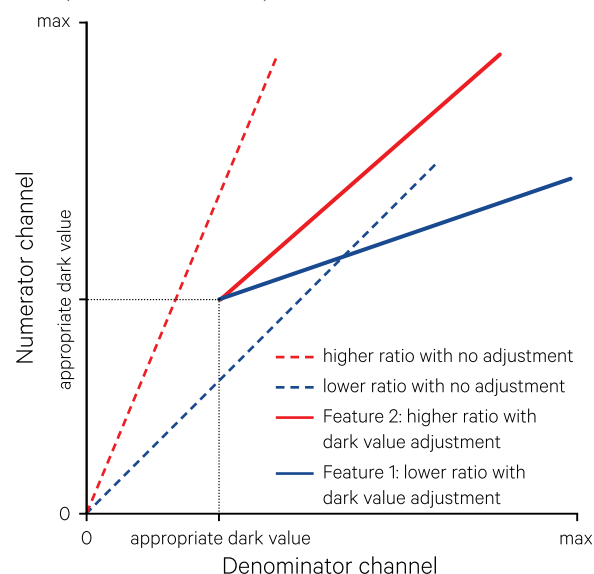
b. When the ratio range between features overlaps, they cannot be separated in the ratio channel.
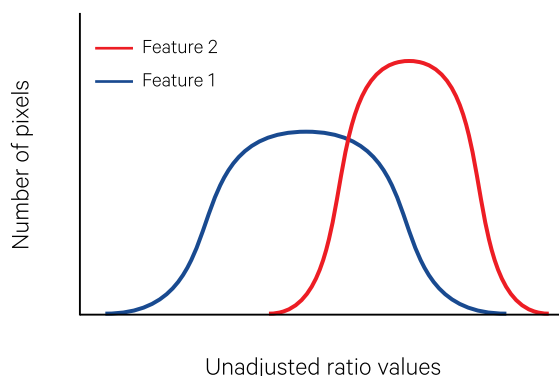
c. Using appropriate dark values, the ratio channel can be computed so that the features have differing ratio ranges.
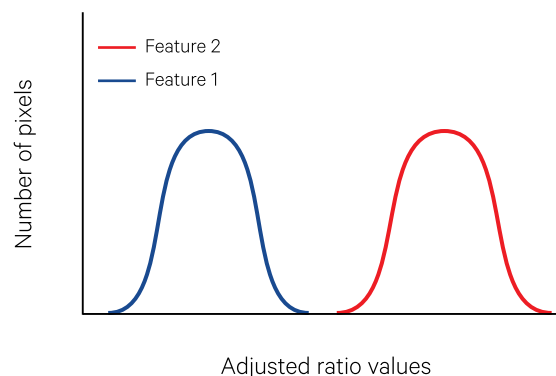
Source: Harrison and Jupp (1990) Figure 95 [Adapted from Crippen, 1988]

117

10.2.2 Calibration impact

The non-linear nature of the ratio operation accentuates any small variations in the spectral response between different channels so image channels are often smoothed before ratioing (see also Section 10.1.3). Conversely, the similarities between different channels are reduced by the ratioing operation. This effect may be used to correct for, or reduce, spatially variable 'noise' such as topographic shading in EO imagery or uneven illumination effects in scan-digitised imagery. These sources of noise (represented in the equation below by the term a_i) are related to ground reflectance by a multiplicative model of the form:

$$x_{ij} = a_i \times b_{ij} + d_j$$

where

x_{ij} is the image radiance value at pixel i in channel j ;

b_{ij} is reflectance (or emissivity) of pixel i in channel j ;

a_i is illumination (or temperature) at pixel i ; and

d_j is a constant atmosphere or sensor calibration offset factor in channel j (or the 'dark value' as explained in Section 10.2.1 above).

These sources of noise can be removed by the ratio operation using an appropriate reference channel (x_{ir}) that represents the noise pattern:

$$\frac{x_{ij} - d_j}{x_{ir} - d_r} = \frac{a_i \times b_{ij}}{a_i \times b_{ir}} = \frac{b_{ij}}{b_{ir}}$$

where subscript r refers to the reference channel.

Radar backscatter and scanner angular effects can similarly be corrected using ratioing techniques. A multiplicative model applies when the detected noise level is affected by target reflectance, that is, the noise factor is not simply added to the reflectance value of the target. (Where the spatial noise is related to ground reflectance by a linear or additive model, as is the case with atmospheric noise, it should be removed using an affine transformation—see Section 7.) For scan-digitised imagery, a 'reference' channel is usually scanned from a blank target and this is used as the denominator channel when computing channel ratios. Pre-processing and selection of suitable reference (denominator) channels for noise correction is discussed in Section 7.2.

10.3 Further Information

Gonzalez and Woods (2018)   Jensen (2016) Section 8

10.4 References

Ahmad, W., Jupp, D.L.B., and Nunez, M. (1989). Land cover mapping in a rugged terrain area using Landsat MSS data. *International Journal of Remote Sensing*, 13(4), 673–683. *https://doi.org/10.1080/01431169208904145*

Castleman, K.R. (1998). *Digital Image Processing*. 2nd edn. Prentice-Hall, Inc. 667 pp.

Chavez, P.S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24, 459–479.

Chavez, P.S., Berlin, G.L., and Sower, L.B. (1982). Statistical method for selecting Landsat MSS ratios. *Journal of Applied Photographic Engineering*, 8, 23–30.

Crane, R.B. (1971). Preprocessing techniques to reduce atmospheric and sensor variability in multispectral scanner data. *Proceedings of 7th International Symposium on Remote Sensing of Environment*, Ann Arbor, Michigan, 17–21 May, 1345–1355.

Crippen, R.E. (1987). The regression intersection method of adjusting image data for band ratioing. *International Journal of Remote Sensing*, 8, 137–155.

Crippen, R.E. (1988). The dangers of underestimating the importance of data adjustments in band ratioing. *International Journal of Remote Sensing*, 9, 767–776.

Crippen, R.E., Blom, R.G., and Heyada, J.R. (1988). Directed band ratioing for the retention of perceptually-independent topographic expression in chromaticity-enhanced imagery. *International Journal of Remote Sensing,* 9, 749–765.

Drury, S.A. (1987). *Image Interpretation in Geology.* Allen and Unwin (Publ) Ltd, London.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing.* Pearson Educational Inc., New York.

Gordon, D.K., and Philipson, W.R. (1986). A texture-enhancement procedure for separating orchard from forest in Thematic Mapper data. *International Journal of Remote Sensing*, 7, 301–304.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.

Holben, B., and Justice, C. (1981). An examination of spectral band ratioing to reduce the topographic effect on remotely sensed data. *International Journal of Remote Sensing*, 2,115–133.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective.* 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Levine, M.D. (1985). *Vision in Man and Machine*. McGraw-Hill Inc. USA.

Moik, J.G. (1980). *Digital Processing of Remotely Sensed Images*. NASA SP–431. Washington, DC. USA.

Openheim, A.V., and Schafer, R.W. (1975). *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, New Jersey.

Satterwhite, M.B., and Henley, J.P. (1987). Spectral characteristics of selected soils and vegetation in northern Nevada and their discrimination using band ratio techniques. *Remote Sensing of Environment*, 23, 155–175.

Switzer, P., Kowalik, W.S., and Lyon, R.J.P. (1981). Estimation of atmospheric path-radiance by the covariance matrix method. *Photogrammetric Engineering and Remote Sensing*, 47, 1469–1476.

Vogelmann, J.E., and Rock, B.N. (1988). Assessing forest damage in high elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sensing of Environment*, 24, 227–46.

11 Vegetation Indices

Various indices have been proposed to highlight and differentiate attributes of vegetation quantity and condition (Bannari *et al.,* 1995; Xue and Su, 2017). The following sections review the most commonly encountered indices used to measure the vegetation attributes of greenness (see Section 11.1) and water content (see Section 11.2). Other indices related to vegetation productivity, biochemistry or structure are described in Volume 3a.

11.1 Vegetation Greenness

Various indices have been proposed to highlight and differentiate vegetation 'greenness'. These vegetation indices have been well reviewed and compared by many authors, including Jackson (1983), Perry and Lautenschlager (1984), and Glenn *et al.* (2008).

The green colour of vegetation results from
- absorption of red and blue wavelengths by the chlorophyll pigment for photosynthesis; and
- reflection of green wavelengths.

These reflectance characteristics are illustrated in Figure 11.1 (see Volume 1B—Section 6.1 for details). However, it is the combined 'colour' resulting from the chlorophyll content of foliage, the area covered by leaves, and the density and structure of the canopy that determines the optical 'greenness' for a given patch of vegetation as measured by a remote sensor.

Healthy green vegetation typically has high reflectance in NIR and low reflectance in red wavelengths (see Figure 10.1 and Volume 1). Other image features may have either high NIR or low red values so it becomes difficult to identify vegetation pixels using only one of these channels. Most vegetation indices are based on highlighting the difference between pixel values in the red and NIR channels.

Reflectance values in NIR and red can be either differenced or ratioed to emphasise those pixels that are likely to contain a high proportion of photosynthetic material. Transformation results may be normalised for comparison between multiple images or differenced between two images to show temporal changes in greenness (see Volume 2D). Variations in vegetation index values over time have been used to monitor a range of changes in vegetation cover and condition (see Volume 3A), including:
- crop growth and senescence—greenness indices are commonly used to monitor crop phenology;
- seasonal changes in plant growth—the annual cycle of vegetation change can be tracked using greenness indices;
- land clearing and regeneration—greenness indices emphasise the difference between vegetated and cleared landscapes;
- fire impact—since fire reduces the amount of green vegetation, greenness indices have been used to map fire footprints, infer burn severity and monitor recovery; and
- fire risk—greenness indices are also used to determine fire risk by mapping the curing status of grasslands and the fire-prone nature of woody vegetation (Chuvieco *et al.*, 2003; Lasaponara, 2005).

Figure 11.1 Reflectance characteristics of typical green leaf structure

Chloroplasts reflect green wavelengths but absorb blue and red wavelengths for use in photosynthesis. Spongy mesophyll cells strongly reflect infrared wavelengths.
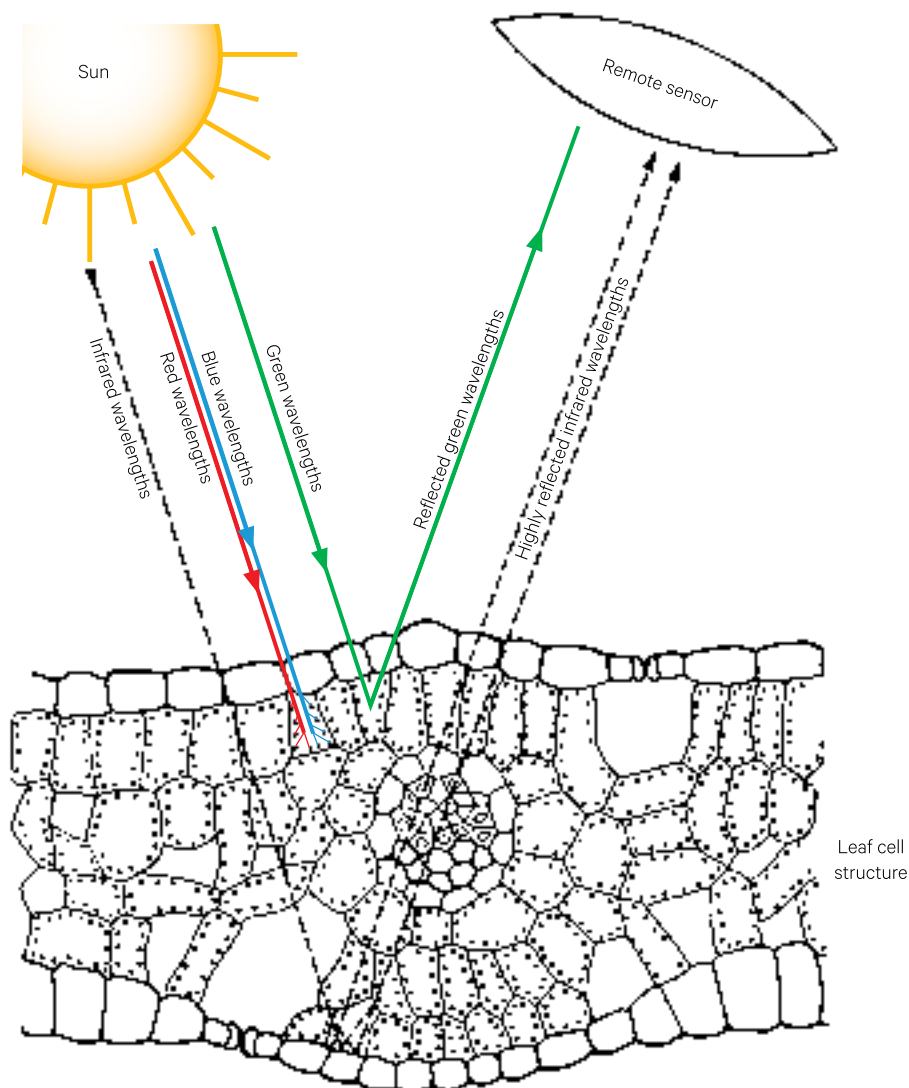
Source: Harrison and Jupp (1989) Figure 5

The following three vegetation indices are commonly encountered in EO analyses:

- Simple Ratio (SR—see Section 11.1.1);
- Normalised Difference Vegetation Index (NDVI— see Section 11.1.2); and
- Kauth-Thomas greenness transformation (or 'Tasselled Cap' Greenness, TCG; see Section 11.1.3).

Applications using vegetation indices, including non-green leaf attributes, are further considered in Volume 3A. Some of the most commonly used vegetation indices are compared in Excursus 11.1.

Excursus 11.1—Comparison of Vegetation Indices

Source: Norman Mueller, Geoscience Australia

Two Landsat-8 OLI images over the diverse agricultural district surrounding the rural township of Kerang in northern Victoria are shown in Figure 11.2. These images contrast the variation in regional vegetation cover and condition between summer and winter in 2018 and will be used to compare four popular vegetation indices. The individual data channels for these example images are shown in Figure 11.3. In this climate, vegetation generally browns off in summer, but predictably the irrigated sections of this district show less seasonal difference than the non-irrigated.

Figure 11.2 Summer and winter example images

These example images were acquired by Landsat-8 OLI over Kerang, northern Victoria, and are displayed as natural colour composites using bands 4, 3, 2 as RGB. Some parts of this agricultural district are irrigated, so show less seasonal difference than others, but in this climate vegetation generally browns off in summer.

a. Summer (7 January 2018)

b. Winter (2 July 2018)

123

Figure 11.3 Individual channels for example images

The most obvious differences between summer and winter images occur in the green and NIR channels, corresponding to changes in vegetative cover and condition.

a. Summer (7 January 2018)   Band   b. Winter (2 July 2018)

blue

green

red

NIR

SWIR1

SWIR2
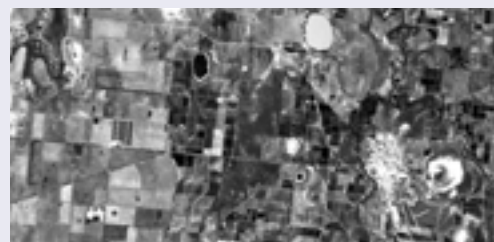
Four different vegetation indices were computed for these two images. The resulting transformations are shown in Figure 11.4. All indices clearly highlight the differences between vegetated and non-vegetated pixels, but visible differences occur both within patches of green vegetation and between green vegetation and other land covers. The extent of green vegetation in the two seasons is clearly indicated by all indices.

Figure 11.4 Four vegetation indices for example images

Four vegetation indices are shown for the pair of example images in Figure 11.2: the simple ratio (NIR/red), Normalised Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and 'Tasselled Cap' Greenness (TCG). While all indices clearly highlight pixels containing green vegetation, there are differences in the way the index values are scaled.

a. Summer (7 January 2018)   Band   b. Winter (2 July 2018)

Simple ratio

NDVI

EVI

TCG
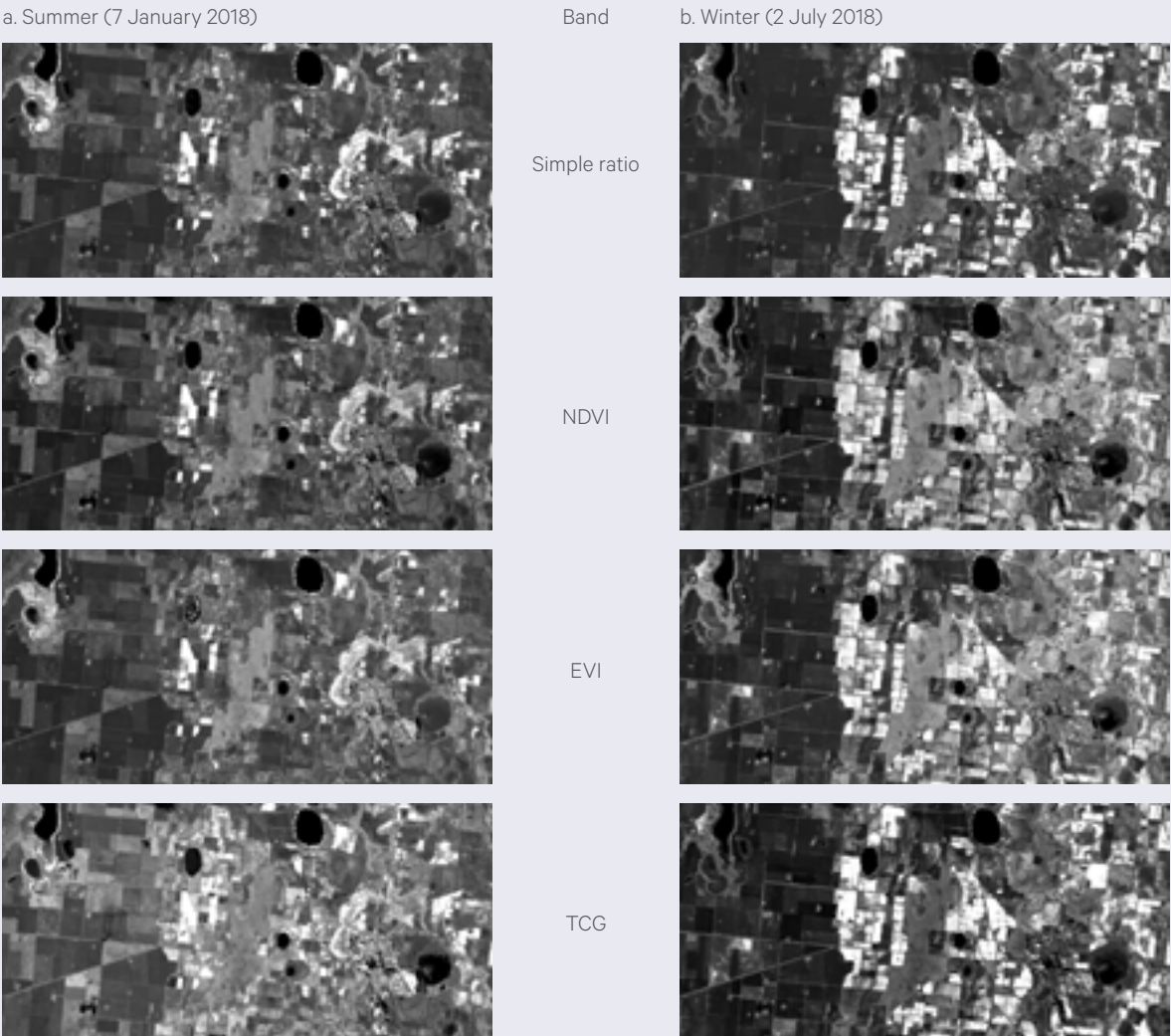
Vegetation is the basic instrument the creator uses to set all of nature in motion.
(Antoine Lavoisier)

11.1.1 Simple Ratio

Dividing a NIR channel value by the red channel value for each pixel in the image gives us a ratio result in which high values only occur in pixels containing vegetation (see Figure 11.4). For vegetation, the strengths of both the NIR reflectance and the red absorption are indicative of plant vigour or type (see Volumes 1 and 3A). Consequently the ratioed values computed from these wavelength bands not only identify vegetative features but also allow us to stratify condition classes within those features. By summarising the ratio result as a single channel we can subsequently use interactive density slicing to define and highlight these classes (see Volume 2A—Section 9.2.1 and Volume 2E).

The direct ratio of NIR to red reflectance (NIR/red) is often referred to as the simple ratio (SR) vegetation index. It is commonly used as a measure of green biomass, but can also indicate vegetation water content (Tucker *et al.* 1980a) and crop condition (Tucker *et al.* 1980b, Thompson and Wehmanen 1979). As well as vegetation 'greenness' or condition mapping, this index has been related to yield of agricultural crops using either logarithmic or linear regression analysis (see Volume 3A).

11.1.2 Normalised Difference Vegetation Index (NDVI) and variants

One of the most commonly used vegetation indices is the Normalised Difference Vegetation Index (NDVI), which is computed as:

$$\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}}$$

where ρ_{NIR} and ρ_{red} indicate the reflectance values for each pixel in the NIR and red bands respectively. As illustrated in Figure 11.5, the normalising of NDVI values results in a controlled range of raw ratio values between 1 and -1 and thus greatly simplifies the scaling of output data and comparison of multi-temporal imagery (Mather, 1987).

NDVI has been related to a range of vegetation characteristics, including cover and condition, which may indicate crop yield in agricultural systems (see Volume 3A). Studies have reported that NDVI saturates for dense vegetation (Carlson and Ripley, 1997; Huete *et al*, 1985). It has been observed to reach its maximum value for fractional covers as low as 60% (Jiang *et al*., 2006) and is adversely affected by bare soil backgrounds when mapping sparse vegetation in arid and semi-arid environments (see also Volume 3A).

Figure 11.5 Normalisation effect in 'Normalised Difference Vegetation Index' (NDVI)

a. Simple ratio (y/x)
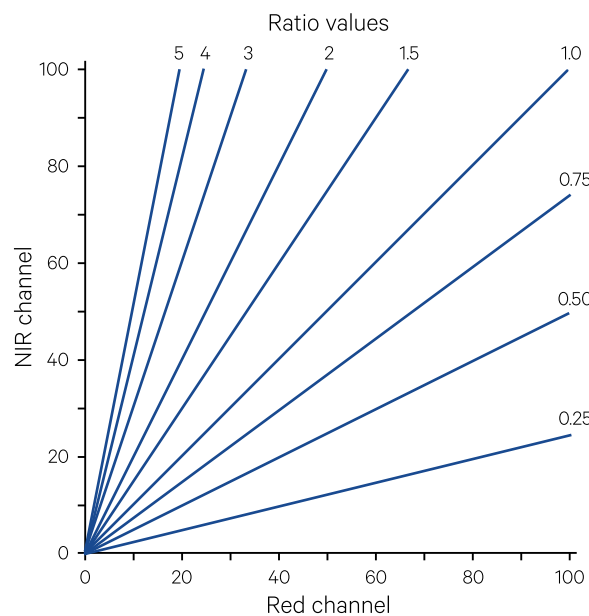
b. Normalised ratio ((y-x)/(y+x))
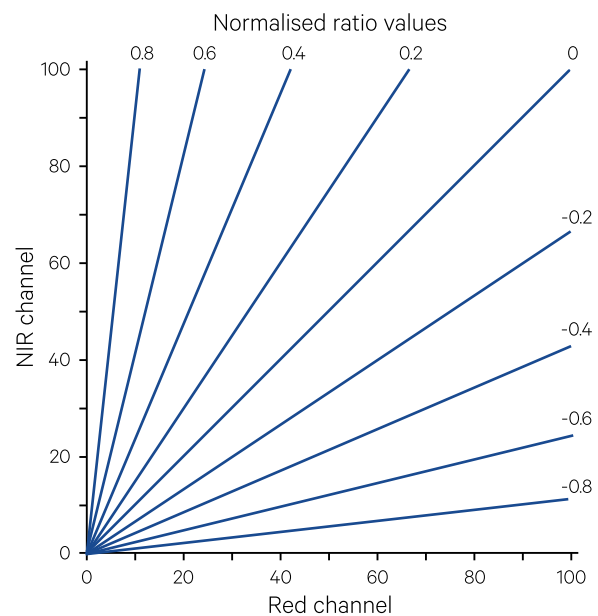
Source: Harrison and Jupp (1990) Figure 94 [Adapted from Mather 1987]

Various modifications to the basic index have been proposed to correct for atmospheric or soil-induced variations (such as Huete 1988; Singh and Saull 1988). Some NDVI variants are summarised in Table 11.1 and detailed in Volume 3A. In particular, Enhanced Vegetation Index (EVI; Huete *et al.*, 1999; 2002; Liu and Huete, 1995; Jiang *et al.*, 2008) attempts to reduce saturation by accounting for effects due to atmosphere and background, though not necessarily topographic effects (Matsushita *et al.*, 2007):

$$\text{EVI} = G \times \left(\frac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + C_1 \times \rho_{\text{red}} + C_2 \times \rho_{\text{blue}} + L} \right)$$

where

- ρ_{red} and ρ_{NIR} are corrected values for MODIS bands 1 (red) and 2 (NIR) respectively;
- ρ_{blue} is the corrected value for MODIS band 3 (blue);
- G is a gain factor (G=2.5; Huete *et al.*, 1994; 1997);
- L is an adjustment for the canopy background (L=1); and
- C_1 and C_2 are coefficients for aerosol resistance (C_1=6, C_2=7.5).

Table 11.1 NDVI variations

Name	Equation	Advantages	Reference
Transformed Vegetation Index (TVI)	$\text{TVI} = \sqrt{\dfrac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + \rho_{\text{red}}}} + 0.5$	avoids operating with negative NDVI values and introduces normal distribution	Deering *et al.* (1975)
Enhanced Vegetation Index (EVI)	$\text{EVI} = G \times \left(\dfrac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + C_1 \times \rho_{\text{red}} - C_2 \times \rho_{\text{blue}} + L} \right)$ where G is a gain factor; C_1 and C_2 are aerosol resistance coefficients; and L is the canopy background adjustment	reduces saturation by accounting for effects due to atmosphere and background	Huete *et al.* (2002)
Atmospherically Resistant Vegetation Index (ARVI)	$\text{ARVI} = \dfrac{\rho_{\text{NIR}} - \left(\rho_{\text{red}} + \gamma \left(\rho_{\text{blue}} - \rho_{\text{red}} \right) \right)}{\rho_{\text{NIR}} + \left(\rho_{\text{red}} + \gamma \left(\rho_{\text{blue}} - \rho_{\text{red}} \right) \right)}$ where usually $\gamma \sim 1$	reduces atmospheric effects	Kaufman and Tanre (1992)
Soil Adjusted Vegetation Index (SAVI)	$\text{SAVI} = \left(\dfrac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + \rho_{\text{red}} + L} \right) (1 + L)$ where L is a soil brightness correction factor (0 for high vegetation cover, 1 for low vegetation cover)	reduces the effects of soil background	Huete (1988)
Optimised Soil Adjusted Vegetation Index (OSAVI)	$\text{OSAVI} = (1 + L) \left(\dfrac{\rho_{800} - \rho_{670}}{\rho_{800} + \rho_{670} + L} \right)$ where $L = 0.16$	minimises soil effects	Rondeaux *et al.* (1996)
Visible Atmospherically Resistant Index (VARI)	$\text{VARI} = \left(\dfrac{\rho_{\text{green}} - \rho_{\text{red}}}{\rho_{\text{green}} + \rho_{\text{red}} - \rho_{\text{blue}}} \right)$	minimally sensitive to atmospheric effects	Gitelson *et al.* (2002)
Absolute Greenness index (Gabs)	$\text{Gabs} = \left(\dfrac{\text{observed NDVI} - \text{min NDVI}}{\text{max NDVI}} \right)$	distinguishes the weather-related variations in NDVI from longer term temporal variability (such as curing) within a selected time period	Eidenshink *et al.* (1990)
Relative Greenness index (RGI)	$\text{RG} = \left(\dfrac{\text{observed NDVI} - \text{min NDVI}}{\text{max NDVI} - \text{min NDVI}} \right) \times 100$	ranks the current NDVI of each pixel relative to its minimum and maximum NDVI over a specified period of time and better distinguishes the volume of dead fuel from live fuel	Burgan *et al.* (1998)
Photochemical Reflectance index (PRI)	$\text{PRI} = \left(\dfrac{\rho_{570} - \rho_{531}}{\rho_{570} + \rho_{531}} \right)$	assesses photosynthetic radiation use efficiency	Peñuelas *et al.* (1997)

127

EVI uses data from MODIS band 3 (500 nm) to remove the effects of smoke, aerosols and thin clouds, and masks out cloud, cloud shadows, marine (as opposed to inland) waters, and aerosols (Huete *et al.*, 2011). Accordingly, it is more sensitive to variations in canopy structure than NDVI (Gao *et al.*, 2000), and thus has been more responsive for land covers with high biomass, such as tropical forests (Huete *et al.*, 2002) and crops (Wardlow *et al.*, 2007). Both NDVI and EVI products are routinely and freely available at global and regional scales (see Volume 2D). Their use for phenological studies is discussed in Volume 3A.

The narrow bandwidths in hyperspectral imagery have enabled a range of indices that highlight subtle changes in vegetation type, health, and density, such as the Red Edge Normalised Difference Vegetation Index (RENDVI; Sims and Gamon, 2002). This index replaces the red and NIR reflectances in NDVI with the red edge peak and red edge vale respectively.

Vegetation indices have also been tailored for fuel modelling. The Fire Potential Index (FPI), for example, used RGI of AVHRR imagery for regional scale fire danger assessment (Burgan *et al.*, 1998) and demonstrated strong correlation with actual fire occurrence and predicting fire potential (see Volume 3A).

11.1.3 Kauth-Thomas Greenness Transformation

Another vegetation index, commonly known as the Kauth-Thomas or Tasselled Cap transformation, can be computed from a linear combination of visible and NIR channels. This transformation is sensor-specific and was originally defined for Landsat MSS (Kauth and Thomas, 1976). A range of similar transformations now exist for other sensors (Crist and Cicone 1984; Cicone and Metzler 1984; Roberts *et al.*, 2018) and can be used to analyse specific feature classes in imagery from multiple sensors. Jackson (1983) provides a detailed description for calculating the coefficients for '*n*-space' indices (that is, linear combinations of *n* spectral channels in *n*-dimensional space) using the Gram-Schmidt process with a minimum of data points.

For mixed agricultural scenes, the Kauth-Thomas transformation describes the essential three dimensions of the Landsat MSS data space as forming the shape of a tasselled cap (see Figure 11.6a). This cap 'sits' on the principal data axis, which has been described as the 'plane of soils'. The initial change in crop reflectance during the crop growing cycle is described as moving down the soil line due to shadowing (though the extent of shadowing varies with the presence and orientation of crop rows). Once maximum soil shadowing has been achieved, the trajectory of reflectance moves away from the soil line. This trajectory curves up to the region of 'green stuff' as the crop canopy becomes denser, then folds over to converge on the region of 'yellow stuff' as the vegetation hays off. The shadowed nature of tree canopies renders them a special position in this model, which is fancifully described as the 'badge' on the 'Tasselled Cap'. The point of shadow represents the minimum reflectance level and can be considered as the equivalent of dark values in image ratioing (see Section 10.2.1). Figure 11.6b illustrates the way changes in vegetation greenness affect its reflectance of red and NIR wavelengths. This dynamic crossplot represents one plane of the Tasselled Cap space.


When a wall is slowly covered over by earth,
the materials it's made from decay and become part of the soils around and above it,
sometimes causing vegetation above and next to the wall to grow faster or slower.
Satellite imagery helps archaeologists to pick up these subtle changes.
*(Sarah Parcak)*

Figure 11.6 The 'Tasselled Cap' feature in Landsat MSS image data

a. For vegetated landscapes, the major dimensions in Landsat MSS imagery can be described in terms of Brightness, Greenness and Yellowness. The Tasselled Cap feature graphically maps different stages of vegetation growth relative to these dimensions.
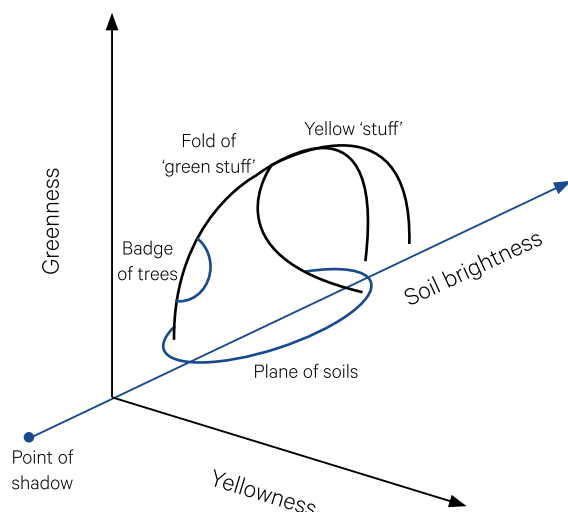
b. One plane through the Tasselled Cap shows the effect of changes in vegetation 'greenness' on red and NIR reflectance.
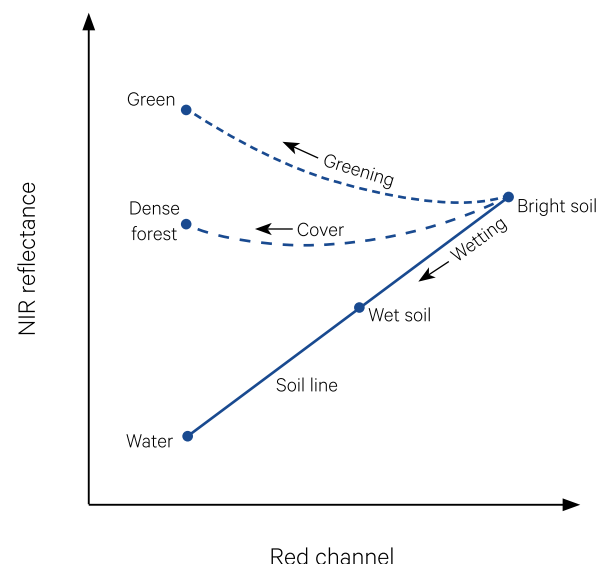
Source: Harrison and Jupp (1990) Figure 92

Using broad land cover categories such as vegetation and soils, a generalised affine transformation was defined to rotate Landsat MSS data channels into dimensions referred to as Brightness, Greenness, Yellowness and Non Such (Kauth and Thomas, 1976). The plane of Brightness versus Greenness contains at least 95% of the total variation for agricultural imagery, with Brightness defining the direction of soil reflectance variation and the perpendicular direction, Greenness, being associated with the reflectance characteristics of green vegetation (Crist and Kauth, 1986). The matrix defined for this rotation is shown in Figure 11.7.

Figure 11.7 Kauth Thomas greenness transformation matrix

$$
\begin{bmatrix}
\text{Brightness} \\
\text{Greenness} \\
\text{Yellowness} \\
\text{Non such}
\end{bmatrix}
=
\begin{bmatrix}
0.433 & 0.632 & 0.586 & 0.264 \\
-0.290 & -0.562 & 0.600 & 0.491 \\
-0.829 & 0.522 & -0.039 & 0.194 \\
0.223 & 0.012 & -0.543 & 0.810
\end{bmatrix}
\times
\begin{bmatrix}
\text{MSS4} \\
\text{MSS5} \\
\text{MSS6} \\
\text{MSS7}
\end{bmatrix}
+
\begin{bmatrix}
32 \\
32 \\
32 \\
32
\end{bmatrix}
$$

Here we see Brightness as a weighted average of the four original MSS channels while Greenness is essentially a weighted difference of NIR and red, which is typical for vegetation indices. The Greenness feature can also be normalised by the 'Green Brightness', that is:

$$\text{Unnormalised Greenness} = -0.29 \times \text{MSS4}$$
$$-0.562 \times \text{MSS5}$$
$$+0.60 \times \text{MSS6}$$
$$+0.491 \times \text{MSS7}$$

and

$$\text{Normalised Greenness}$$
$$= \frac{\text{Unnormalised Greenness}}{\text{Green Brightness}}$$

where

$$\text{Green Brightness} = 0.29 \times \text{MSS4}$$
$$+0.562 \times \text{MSS5}$$
$$+0.60 \times \text{MSS6}$$
$$+0.491 \times \text{MSS7}$$

This normalisation partially compensates for variations in solar elevation, so allows more valid comparisons of greenness changes in multi-temporal images.

A similar transformation was defined for six channels of Landsat TM data, which produced three significant features called Brightness, Greenness and Wetness (Crist and Cicone, 1984):

$$\text{Brightness} = 0.33183 \times \text{TM1} + 0.33121 \times \text{TM2}$$
$$+0.55177 \times \text{TM3} + 0.42514 \times \text{TM4}$$
$$+0.48087 \times \text{TM5} + 0.25252 \times \text{TM7}$$
$$\text{Greenness} = -0.24717 \times \text{TM1} - 0.16263 \times \text{TM2}$$
$$-0.40639 \times \text{TM3} + 0.85468 \times \text{TM4}$$
$$+0.05493 \times \text{TM5} - 0.11749 \times \text{TM7}$$
$$\text{Wetness} = 0.13929 \times \text{TM1} + 0.22490 \times \text{TM2}$$
$$+0.40359 \times \text{TM3} + 0.25178 \times \text{TM4}$$
$$-0.70133 \times \text{TM5} - 0.45732 \times \text{TM7}$$

Equivalent coefficients have been derived for Landsat-8 OLI imagery, calibrated to surface reflectance (Roberts *et al.*, 2018):

$$\text{Brightness} = 0.3567 \times \text{blue} + 0.3567 \times \text{green}$$
$$+0.3567 \times \text{red} + 0.5350 \times \text{NIR}$$
$$+0.5350 \times \text{SWIR1} + 0.2140 \times \text{SWIR2}$$
$$\text{Greenness} = 0.1987 \times \text{blue} - 0.2826 \times \text{green}$$
$$-0.2724 \times \text{red} + 0.5357 \times \text{NIR}$$
$$+0.2388 \times \text{SWIR1} - 0.6800 \times \text{SWIR2}$$
$$\text{Wetness} = 0.5702 \times \text{blue} + 0.1584 \times \text{green}$$
$$+0.2627 \times \text{red} - 0.3959 \times \text{NIR}$$
$$-0.0045 \times \text{SWIR1} - 0.6511 \times \text{SWIR2}$$

Variations in these (land cover related) features are represented by changes in two or more channels in the original imagery. As with PCA, rotation of the original data space to axes that are aligned with the major data variation in land cover features allows them to be more simply analysed and related to ground-based measurements.

Dark values can generally be specified for each channel to represent the level of atmospheric or sensor noise. In most image processing systems, these values can be defined by interactive training on shadowed or deep water areas, or defaulted to image minimum values (see Section 10.2.1).

This transformation has been used in various agricultural models (see Volume 3A). As a single channel it both allows simpler relationships to be developed between image radiance and ground-based measurements, and permits vegetation categories to be analysed easily using interactive image painting (see Volume 2A—Section 9.2.1).

130

Earth Observation: Data, Processing and Applications. Volume 2: Processing

11.2 Vegetation Water Content

EO data can be used to quantify the water content of vegetation by relying on the liquid water absorption features in optical wavelengths (see Volume 3A for details). Both NIR and SWIR reflectance data are needed to determine water content since:

- NIR indicates the internal leaf structure and dry matter content (see Figure 11.1); and
- SWIR indicates the vegetation water content, internal structure and dry matter content (see Figure 11.8).

Some of the ratios and normalised ratios computed from these wavelengths are summarised in Table 11.2. These water content indices have shown good correlation with empirical measures of foliar water content, such as Equivalent Water Thickness (EWT) and have been used to infer fuel moisture for fire risk prediction (see Volume 3A).

Table 11.2 Vegetation water content indices

Name	Formulation	Reference
Water Index (WI)	$WI = \left(\dfrac{\rho_{900}}{\rho_{970}} \right)$	Peñuelas et al. (1997b)
Normalised Difference Water Index (NDWI)	$NDWI = \left(\dfrac{\rho_{860} - \rho_{1240}}{\rho_{860} + \rho_{1240}} \right)$	Gao (1996)
Simple Ratio Water Index (SRWI)	$SRWI = \left(\dfrac{\rho_{860}}{\rho_{1240}} \right)$	Zarco-Tejada et al. (2003)
Normalised Difference Infrared Index (NDII)	$NDII = \left(\dfrac{\rho_{819} - \rho_{1600}}{\rho_{819} + \rho_{1600}} \right)$	Hardinsky et al. (1983)
Moisture Stress Index (MSI)	$MSI = \left(\dfrac{\rho_{1600}}{\rho_{819}} \right)$	Hunt and Rock (1989)
Datt Water Index (DWI)	$DWI = \left(\dfrac{\rho_{895} - \rho_{675}}{\rho_{895} + \rho_{675}} \right)$	Datt (1999)

Figure 11.8 Typical green vegetation spectrum

Reflectance properties of green vegetation primarily derive from leaf pigments in visible wavelengths, cell structure in NIR wavelengths, and water content in SWIR wavelengths. Characteristic absorption features occur in visible wavelengths (due to chlorophyll) and SWIR wavelengths (due to water content).
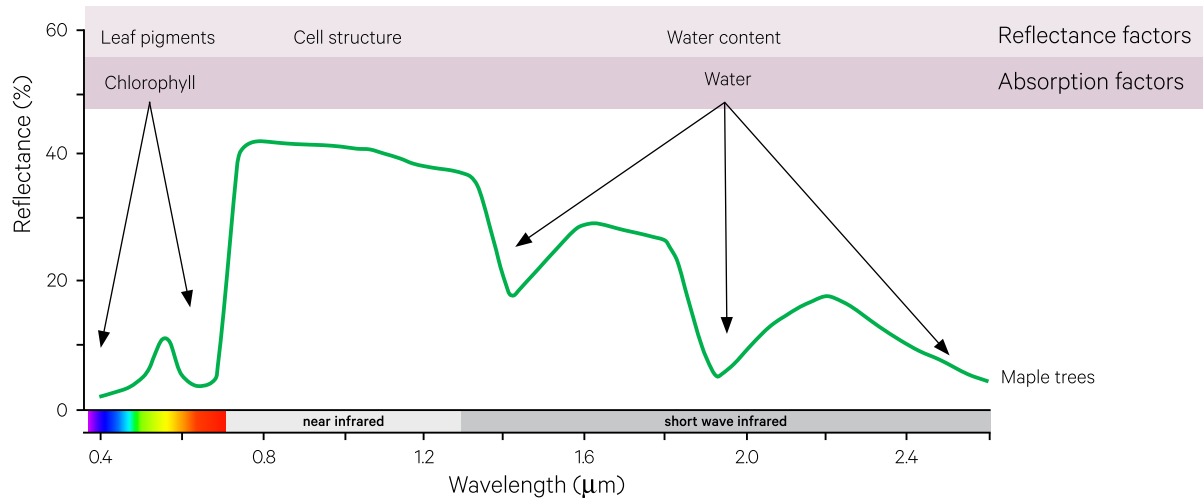
Adapted from: Harrison and Jupp (1989) Figure 30

11.3 Further Information

Jensen (2016) Section 8

Index Database (a database for remote sensing indices): *https://www.indexdatabase.de*

11.4 References

Bannari, A., Morin, D., Bonn, F. and Huete, A. R.(1995). A review of vegetation indices. *Remote Sensing Reviews*, 13, 1, 95–120. doi: *10.1080/02757259509532298*

Burgan, R., Klaver, R., and Klaver, J. (1998). Fuel models and fire potential from satellite and surface observations. *International Journal of Wildland Fire*, 8, 159–170.

Carlson, T., and Ripley, D. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, 62, 241–252.

Chuvieco, E.,I. Aguado, D. Cocero, D. and, and Rian, O. (2003). Design of an empirical index to estimate fuel moisture content from NOAA-AVHRR images in forest fire danger studies. *International Journal of Remote Sensing*, 24, 1621–1637.

Cicone, B.C., and Metzler, M.D. (1984). Comparison of Landsat MSS, Nimbus–7, CZCS, and NOAA–7 AVHRR features for land-use analysis. *Remote Sensing of Environment*, 14, 257–265.

Crippen, R.E. (1987). The regression intersection method of adjusting image data for band ratioing. *International Journal of Remote Sensing,* 8, 137–155.

Crippen, R.E. (1988). The dangers of underestimating the importance of data adjustments in band ratioing. *International Journal of Remote Sensing*, 9, 767–776.

Crippen, R.E., Blom, R.G., and Heyada, J.R. (1988). Directed band ratioing for the retention of perceptually-independent topographic expression in chromaticity-enhanced imagery. *International Journal of Remote Sensing*, 9, 749–765.

Crist, E.P., and Cicone, R.C. (1984). Application of the Tasselled Cap concept to simulated Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, 50, 343–352.

Crist, E.P., and Kauth, R.J. (1986). The Tasselled Cap de-mystified. *Photogrammetric Engineering and Remote Sensing*, 52, 81–86.

Datt, B. (1999) Remote Sensing of Water Content in *Eucalyptus* Leaves. *Australian Journal of Bot*any, 47, 909–923.

Deering, D.W., Rouse, J.W., Haas, R.H., and Schell, J.A. (1975). Measuring Forage Production of Grazing Units from Landsat MSS Data. *10th International Symposium on Remote Sensing of Environment*, 2, 1169–1178.

Eidenshink J.C., Burgan, R.E., and Haas, R.H. (1990). Monitoring fire fuels condition by using time series composites of Advanced Very High Resolution Radiometer (AVHRR) data. *Proceedings of Resource Technology 90*, Washington, DC, ASPRS, 68–82.

Gao, B.C. (1996). NDWI—a normalised difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58, 257–266.

Gao, X., Huete, A. R., Ni, W., and Miura, T. (2000). Optical–biophysical relationships of vegetation spectra without background contamination, *Remote Sensing of Environment*, 74, 609–620.

Gitelson, A.A., Kaufman, Y.J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment,* 80(1), 76–87.

Glenn, E.P., Huete, A.R., Nagler, P.L., and Nelson, S.G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*, 8, 2136–2160.

Hardisky, M. A., Klemas, V., and Smart, R. M. (1983). The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of *Spartina alterniflora* canopies. *Photogrammetric Engineering and Remote Sensing*, 49, 77–83.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Huete, A, Jackson, R., and Post, D. (1985). Spectral response of a plant canopy with different soil backgrounds. *Remote Sensing of Environment*, 17, 37–53.

Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295–309.

Huete, A., Justice, C. and van Leeuwen, W. (1999). *MODIS vegetation index (MOD 13) algorithm theoretical basis document (ATBD)* Version 3. 0. EOS Project Office, NASA Goddard Space Flight Center, Greenbelt, MD.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. and Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83, 195–213.

Huete, A., Didan, R., van Leeuwen, W., Miura, T., and Glenn, E. (2011). MODIS Vegetation Indices. Ch 26 in *Land Remote Sensing and Global Environmental Change, Remote Sensing and Digital Image Processing* Volume 11, Springer Science+Business Media, LLC. ISBN 978-1-4419-6748-0.

Hunt, E.R., Jr, and Rock, B.N. (1989). Detection of changes in leaf water content using near- and middle-infrared reflectances. *Remote Sensing of Environment*, 30, 43–54.

Jackson, R.D. (1983). Spectral Indices in n-Space. *Remote Sensing of Environment*, 13, 409–421.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Jiang, Z., Huete, A., Chen, J., Chen, Y., Li, J., Yan, G., and Zhang, X. (2006). Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sensing of Environment*, 101, 366–378.

Jiang, Z., Huete, A.R., Didan, K., and Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112, 3833–3845.

Kaufman, Y.J., and Tanre, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS- MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2), 261–270.

Kauth, R.J., and Thomas, G.S. (1976). The Tasselled Cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proc. Symposium on Machine Processing of Remotely Sensed Data*. Purdue University, West Lafayette, Indiana, 4B41–4B51.

Lasaponara, R. (2005). Inter-comparison of AHVRR-based fire susceptibility indicators for the Mediterranean ecosystems of Southern Italy. *International Journal of Remote Sensing*, 26, 853–870.

Liu, H.Q., and Huete, A.R. (1995). A feedback based modification of the NDV I to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing*, 33, 457–465.

Matsushita, B., Yang, W., Chen, J., Onda, Y., and Qiu, G. (2007). Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-Density Cypress Forest. *Sensors*, 7, 2636–2651.

Peñuelas, J., Llusia, J., Piñol, J., and Filella, I. (1997a) Photochemical reflectance index and leaf photosynthetic radiation-use efficiency assessment in Mediterranean trees. *International Journal of Remote Sensing*, 18, 2863–2868.

Peñuelas, J. Pinol, J., Ogaya, R., and Filella, I. (1997b) Estimation of plant water concentration by the reflectance Water Index WI (R900/R970). *International Journal of Remote Sensing,* 18(13), 2869–2875. doi: *10.1080/014311697217396*

Perry, C.R.Jr., and Lautenshlager, L.F. (1984). Functional equivalent of spectral vegetation indices. *Remote Sensing of Environment*, 14, 169–182.

Roberts, D., Dunn, B., and Mueller, N. (2018). Open Data Cube Products Using High-Dimensional Statistics of Time Series. *Proceedings 2018 IGARSS*, 22–27 July, Valencia Spain.

Rondeaux, G, Steven, M. and Baret, F. (1996). Optimisation of Soil-Adjusted Vegetation Indices. *Remote Sensing of Environment*, 55, 95–107.

Sims, DA., and Gamon, J.A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, 81, 337–354.

Singh, S.M., and Saull, R.J. (1988). The effect of atmospheric correction on the interpretation of multitemporal AVHRR-derived vegetation index dynamics. *Remote Sensing of Environment*, 25, 37–51.

Thompson, D.R., and Wehman, O.A. (1979). Using Landsat digital data to detect moisture stress. *Photogrammetric Engineering and Remote Sensing*, 45, 201–207.

Tucker, C.J., Elgin, J.H.Jr., and McMurtney, J.E.III. (1980a). Relationship of crop radiance to alfalfa agronomic values. *International Journal of Remote Sensing*, 1, 69–75.

Tucker, C.J., Holben, B.N., Elgin, J.H.Jr., and McMurtney, J.E.III. (1980b). Relationship of spectral data to grain yield variation. *Photogrammetric Engineering and Remote Sensing*, 46, 657–666.

Wardlow, B.D., Egbert, S.L., and Kastens, J.H. (2007) Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment*, 108, 290–310.

Xue, J., and Su, B. (2017). Review Article: Significant remote sensing vegetation indices: a review of developments and applications. *Journal of Sensors*, ID 1353691. *https://doi.org/10.1155/2017/1353691*

Zarco-Tejada, P.J., Rueda, C.A., and Ustin, S.L. (2003). Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sensing of Environment*, 85, 109–124.

134

Earth Observation: Data, Processing and Applications. Volume 2: Processing