UNIT 11 IMAGE CORRECTIONS

Structure

11.1 INTRODUCTION

You have studied the characteristics of remote sensing images in Unit 10 *Characteristics of Digital Remote Sensing Images* of MGY-002. You have also learnt that these images are commonly acquired by sensors mounted on remote sensing platforms. Such images are susceptible to a variety of distortions that have to be corrected before they are suitable for application. Distortions can affect the values recorded by the sensors in different wavelengths. The geometry of the pixel as well as alignment of an image with other images and also reference maps is also affected. It is important to know the effects of these distortions on the images and apply suitable corrections to minimise them.

In this unit, we will discuss the types and causes of major distortions in remote sensing images and methods of their corrections to make images able to represent the terrain more closely.

Objectives

After studying this unit, you should be able to:

- describe the concept of distortions suffered by remotely sensed images;
- discuss types and causes of radiometric and geometric distortions;
- explain the requirements and approaches of radiometric corrections; and
- outline the steps and methods for geometric correction as applied to the images.

11.2 CONCEPT OF IMAGE DISTORTION AND CORRECTION

Let us first get introduced to the concept of image distortion.

11.2.1 Image Distortions

Any kind of errors present in remote sensing images are known as *image distortions*. Any remote sensing images acquired from either spaceborne or airborne platforms are susceptible to a variety of distortions. These distortions occur due to data recording procedure, shape and rotation of the Earth and environmental conditions prevailing at the time of data acquisition.

Distortions occurring in remote sensing images can be categorised into two types:

- radiometric distortions and
- geometric distortions.

Sensors record the intensity of electromagnetic radiation (EMR) as digital numbers (DNs). These DNs are specific to the sensor and conditions under which they were measured. However, there are variations in the pixel intensities (digital numbers) irrespective of the object or scene being scanned. The recorded values get distorted due to one or more of the following factors:

- sensor ageing
- random malfunctioning of the sensor elements
- atmospheric interference at the time of image acquisition and
- topographic effects.

The above factors affect radiometry (variation in the pixel intensities) of the images and resultant distortions are known as *radiometric distortions*.

As you know, an image is composed of finite number of pixels. The positions of these pixels are initially referenced by their row and column numbers. However, if you want to use images, you should be able to relate these pixels to their positions on the Earth surface. Further, the distance, area, direction and shape properties vary across an image, thus these errors are known as *geometric errors/distortions*. This distortion is inherent in images because we attempt to represent three-dimensional Earth surface as a two-dimensional image. Geometric errors originate during the process of data collection and vary in type and magnitude. There are several factors causing geometric distortions such as:

- Earth's rotation
- Earth's curvature
- satellite platform instability and
- instrument error.

Let us now learn about the concept of image correction.

11.2.2 Image Corrections

As you now know that raw remote sensing images always contain significant amount of distortions, therefore, they cannot be used directly for further image analysis. The image correction involves image operations which normally precedes manipulation and analysis of image data to extract specific information. The primary aim of image correction operations is to correct distorted image data to create a more accurate representation of the original scene. Image corrections are also known as a *preprocessing* of remotely sensed images. It is a preparatory phase that improves quality of images and serves as a basis for further image analysis.

Modelling of the radiometric and geometric distortions and consequent corrections of distortions falls in the category of preprocessing of remotely sensed imagery.

Depending upon the kinds of errors which are present in images, the image correction functions are comprised of radiometric and geometric corrections. *Radiometric correction* attempts to improve the accuracy of measurements made by remote sensors pertaining to the spectral reflectance or emittance or back-scatter from the objects on the Earth surface.

Geometric correction is the process of correcting geometric distortions and assigning the properties of a map to an image.

Both of these corrections are made prior to actual use of remote sensing data in resource management, environmental monitoring and change detection applications by application scientists.

A complete chain of processing of remote sensing images is shown in Fig. 11.1. It becomes clear from this figure that image correction forms an integral part of processing of images.

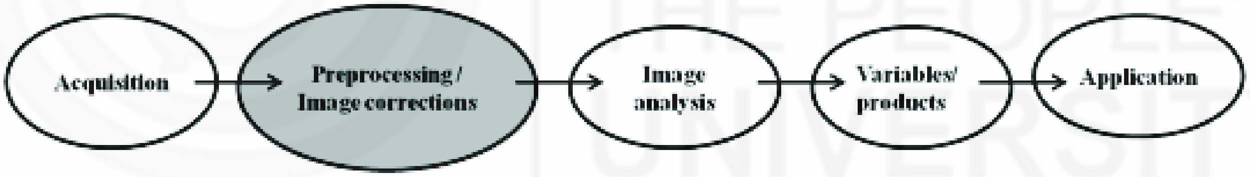


Fig. 11.1: Chain of broad steps in remote sensing data processing

11.3 RADIOMETRIC DISTORTIONS AND THEIR CORRECTIONS

You have read earlier that radiometric distortions relate to the distortions suffered by the images in recorded values at different pixel locations. Let us now discuss in detail about the types of radiometric errors and their correction procedures.

11.3.1 Nature of Radiometric Errors

The radiometric errors listed in subsection 11.2.1 can be broadly categorised into the following two categories:

- internal errors and
- external errors.

Internal errors are introduced by the electronics themselves. These kinds of errors are also known as *systematic errors* because of their systematic nature. These errors can be modelled, identified and corrected based on laboratory calibration or in-flight measurements. For example, if a single detector has become uncalibrated, the concerned row (in older satellites such as Landsat) or the column (in pushbroom scanners like SPOT, IRS, IKONOS, WorldView-1) would appear like a constant intensity stripe, that does not reflect the terrain changes on the ground.

External errors are a result of phenomena that vary in nature through space and time and hence are also known as *non-systematic errors*. External variables such as atmospheric disturbances, steep terrain undulations can cause remote sensor data to exhibit radiometric and geometric errors.

Correction of radiometric errors requires knowledge about EMR principles and the interactions that take place during data acquisition process. The radiometric correction can benefit from the terrain information such as slope and aspect and advanced information like bi-directional reflectance distribution function (BRDF) characteristics of the scene. Radiometric correction procedures can be time consuming and at times problematic.

BRDF is a function which describes the magnitude of the upwelling radiance of the target in terms of illumination angle and the angle of view of the sensor.

We shall now discuss about the two types of radiometric errors and their correction procedures in the next two subsections.

11.3.2 Systematic Errors and their Corrections

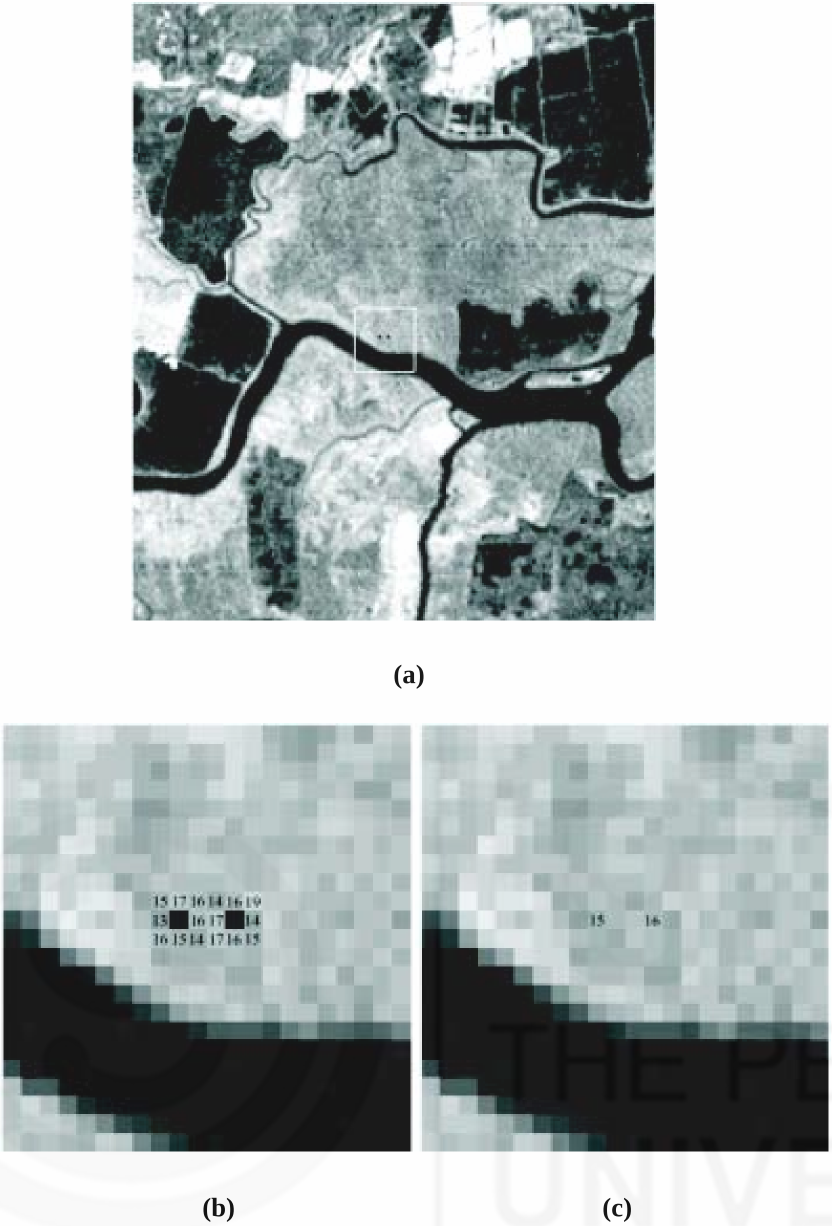
Some of the commonly observed systematic radiometric errors are:

- random bad pixels
- line or column drop-outs
- line start problems
- n-Line striping

We shall now discuss here about these errors and their corrections.

Random Bad Pixels (Shot Noise)

Sometimes, an individual detector does not record received signal for a pixel. This may result in *random bad pixels*. When there are numerous random bad pixels found within the scene, it is called ***shot noise***. Short noise gives the image an impression of having many dark poke marks. Generally, these bad pixels contain values in the range of 0 or 255 (in 8-bit data) in one or more bands. Shot noise is removed by identifying such pixels in a given band that are either 0 (black) or 255 (white) in the midst of neighbouring pixel values that are radically different. These noise pixels are then replaced by the average pixel values of their respective eight neighbouring pixels. For example, in Fig. 11.2ab, two of the pixels have zero gray levels, which is entirely different from their neighbouring pixels. These are marked as shot noise pixels and are replaced by the average of their eight neighbouring pixels.



(a)

(b) (c)

Fig. 11.2: Illustration of shot noise and its removal. (a) A Landsat TM band 7 data with shot noise (two black dots in the region within the white box), (b) zoomed image of the box portion showing the bad pixels along with DNs of the neighbouring eight pixels for each bad pixel and (c) the same image portion after the shot noise removal showing the pixel values which has replaced the bad pixels (source: Lecture slides of Prof. J. R. Jensen, University of South Carolina; used with permission)

If pixel at location (m,n) is a shot noise pixel, then f(m,n) is corrected as given below:

$$f(m,n) = [f(m-1,n-1) + f(m-1,n) + f(m-1,n+1) + f(m,n-1) + f(m,n+1) + f(m+1,n-1) + f(m+1,n) + f(m+1,n+1)] / 8$$

where,

m and n are pixel locations in x and y coordinates, i.e. columns and rows, respectively.

By replacing with average of neighbouring pixels like (15+17+16+13+16+16+15+14) / 8 = 15 and (14+16+19+17+14+17+16+15) / 8 = 16), shot noise pixels disappear after the correction as seen in Fig. 11.2c.

Line or Column Dropouts

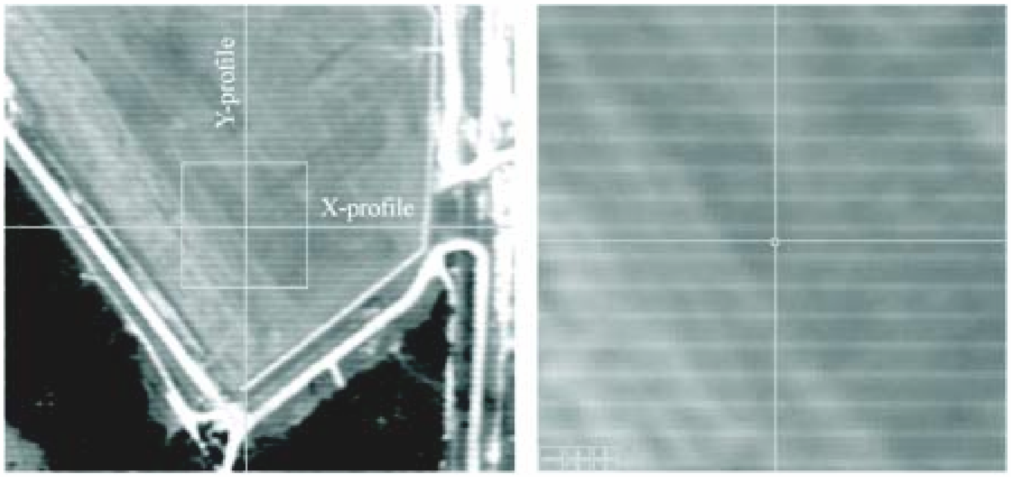
You may have noticed that a blank row containing no details of features on the ground may be seen if an individual detector in an electro-mechanical scanning system (e.g., Landsat MSS or Landsat 7 ETM^+) fails to function properly. If a detector in a pushbroom linear array (e.g., IRS-1C, QuickBird) fails to function, this can result in an entire column of data with no spectral information. The bad line or column is commonly called a *line* or *column drop-out* and contains brightness values equal to zero or some constant value, independent of the terrain changes. Generally, this is an irretrievable loss of information because there is no way to restore data that were never acquired. However, it is possible to improve visual interpretability of data by introducing estimated brightness values for each bad scan row (or column) by replacing it with the average of rows (or columns) above and below (or to the left and right). This concept works because adjacent pixels often have similar pixel values.

Line Start Problems

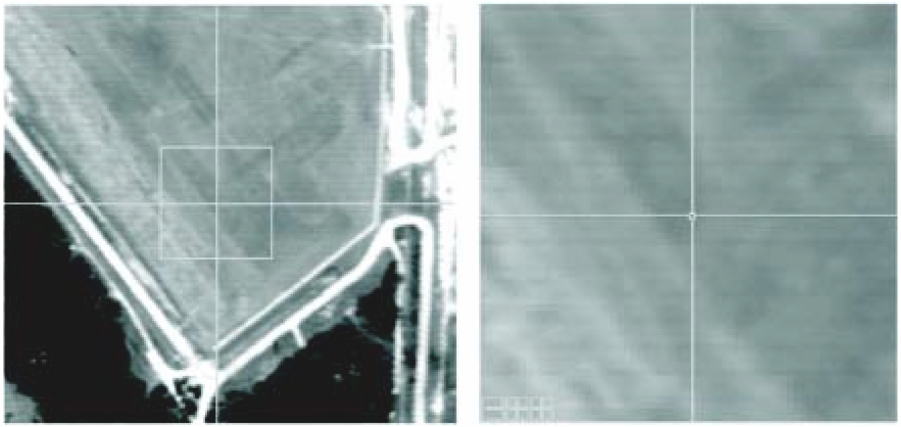
There is another kind of problem encountered in earlier satellites is that the scanner fails to start recording as soon as a new row starts. It may also happen that the sensors place pixel data at inappropriate locations (with shift) along the scan line. For example, all of the pixels in a scan line might be systematically shifted just one pixel to the right. This is called a *line-start* problem. If line start problem is always associated with a horizontal bias of a fixed number of columns, it can be corrected using a simple horizontal adjustment.

n-Line Striping

Sometimes, a detector does not fail completely but its calibration parameters (gain and offset/bias) are disturbed. For example, a detector might record signals over a dark, deep body of water that are almost uniformly 20 or 30 gray levels higher than the other detectors for same band. The result would be an image with systematic, noticeable lines that are brighter than adjacent lines. This is referred to as *n-line striping.* The affected line contains valuable information but should be corrected to have approximately the same radiometric quality as data collected by properly calibrated detectors associated with the same band (Fig. 11.3).



(a) **(b)**



(c) (d)

Fig. 11.3: Illustration of image destriping on a sample remote sensing data. (a) Original image data containing bad lines (stripes), (b) zoomed portion of the box region of the image (a) showing the stripes, (c) sample data after destriping and (d) zoomed portion of the box region of image (c) (source: Lecture slides of Prof. J.R. Jensen, University of South Carolina; used with permission)

To repair systematic n-line striping, it is necessary to identify mis-calibrated scan lines in the scene. This is usually accomplished by computing a histogram of values for each of the *n* detectors that collected data over the entire scene (ideally, this would take place over a homogeneous area likes a water body). If one detector's mean or median DN value is significantly different from the others, there is a probability that this detector is out of adjustment. Consequently, every line and pixel in the scene recorded by the maladjusted detector may require a correction. This type of n-line striping correction:

- adjusts all the bad scan lines so that they have approximately the same radiometric scale as the correctly collected data and
- improves visual interpretability of the data.

Let us now discuss about non-systematic errors and their correction in the next subsection.

11.3.3 Non-Systematic Errors and their Corrections

It is essential to carry out corrections for non-systematic errors in the following circumstances:

- if you are trying to compare remote sensing images which have been acquired at different times
- if you are modelling interactions between EMR and a surface feature, or
- using band ratios for image analysis.

Correction of non-systematic errors include following three steps (Fig. 11.4):

Step 1: Involves the conversion of DNs to at sensor spectral radiance. This step requires information on the 'gain' and 'bias' of the sensor in each image band. The 'gain' and 'bias' are the sensor calibration information. *Bias* is the spectral radiance of the sensor for a DN of zero and *gain* represents the gradient of the calibration. The sensor calibration is carried out before the launch of the sensor.

Sensor calibration includes procedures that convert digital numbers to physical values of radiance. The relationship between DN recorded at a given location and reflectance of the material making up the surface of pixel area changes with time hence the coefficient values that are used to calibrate image data also vary with time.

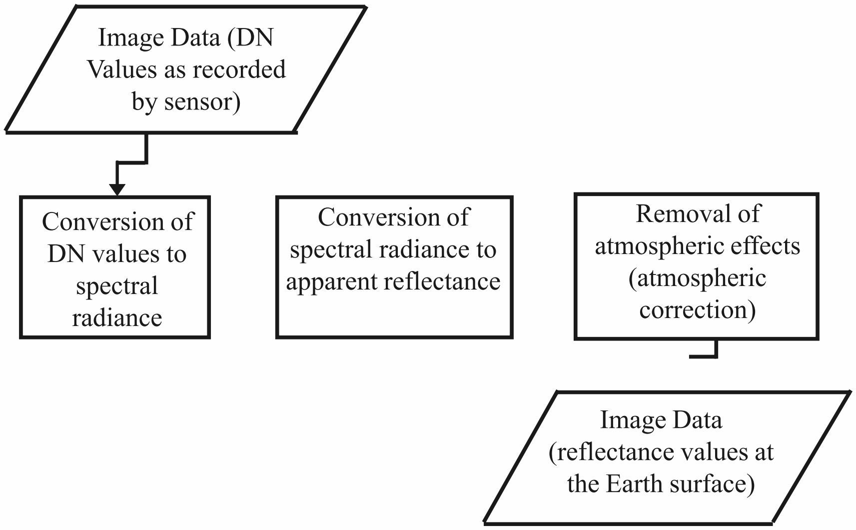


Fig. 11.4: Steps in non-systematic radiometric error correction process

Step 2: Is the conversion of spectral radiance to apparent reflectance. It converts DN values from radiance units to apparent reflectance. Apparent reflectance defines as the reflectance at the top of the atmosphere.

Step 3: Involves the removal of atmospheric effects due to the absorption and scattering of light (atmospheric correction). There are several methods for atmospheric correction of the remotely sensed data. Some methods are relatively straight forward while others are based on the physical principles of interaction of radiation with atmosphere and require a significant amount of information pertaining to the atmospheric conditions to be effective. For our convenience we can categorise atmospheric correction procedures into the following three:

- atmospheric modelling
- image based methods and
- direct calibration using field-derived reflectance (empirical line method).

Atmospheric Modelling

Ideally, this approach is used when scene specific atmosphere data (such as aerosol content, atmospheric visibility) are available.

In this approach, atmospheric radiative transfer codes (models) are used that can provide realistic estimates of the effects of atmospheric scattering and absorption on satellite imagery. Once these effects have been identified for a specific date of imagery, each band and/or pixel in the scene can be adjusted to remove the effects of scattering and/or absorption. The image is then considered to be *atmospherically corrected*. Unfortunately, application of these codes to a specific scene and date also requires knowledge of both the sensor spectral profile and atmospheric properties at the same time. And, for most of the historic satellite data, they are not available.

Most current radiative transfer based atmospheric correction algorithms can compute much of the required information, if

- the user provides fundamental atmospheric characteristic information to the programme or
- certain atmospheric absorption bands are present in the remote sensing dataset.

For example, most radiative transfer based atmospheric correction algorithms require that the user provides

- latitude and longitude of the image scene
- date and exact time of the image
- image acquisition altitude (e.g., 600 km above the ground level)
- mean elevation of the scene (e.g., 450 m above the mean sea level)
- an atmospheric model (e.g., polar summer, mid-latitude winter, tropical)
- radiometrically calibrated image radiance data
- data about each specific band (mean and full width at half-maximum (FWHM)) and
- local atmospheric visibility at the time of remote sensing data collection.

FWHM (Full Width at Half-Maximum) is the wavelength range defined by the two points at which the intensity level is 50% of its peak value.

These parameters are then input to the atmospheric model selected and used to compute the absorption and scattering characteristics of the atmosphere at the time of remote sensing data collection. These atmospheric characteristics are then used to invert the remote sensing radiance to *scaled surface reflectance.* Many of these atmospheric correction programmes derive the scattering and absorption information they require from robust atmosphere radiative transfer code such as

- MODTRAN 4
- ACORN
- ATCOR
- ATREM and
- FLAASH

Image Based Methods

One of the commonly used methods is known as *dark pixel subtraction method* (Chavez, 1988). This method of radiometric correction is based on simple heuristics, used to reduce the effect of haze in the image. The underlying assumption is that some image pixels have a reflectance of zero and DN values of these "zero" pixels recorded by the sensor result from atmospheric scattering (path radiance; see Fig. 11.5). To remove path radiance, minimum pixel value for each band is subtracted from all other pixels in different bands. DNs in pixels representing deep clear water in near-infrared (NIR) band and dark shadows in visible bands are assumed to result from atmospheric path radiance. Since histogram offset can be used as a measure of path radiance, it is also known as *histogram minimum method*. Bi-plots of NIR band against the other bands are generated for pixels of dark regions and then regression techniques are used to calculate the y-intercept which represents path radiance in other bands. The y-intercept value is then subtracted from all pixels in the image.

This method is generally used prior to band-ratioing a single image and is not generally employed for image-to-image comparisons.

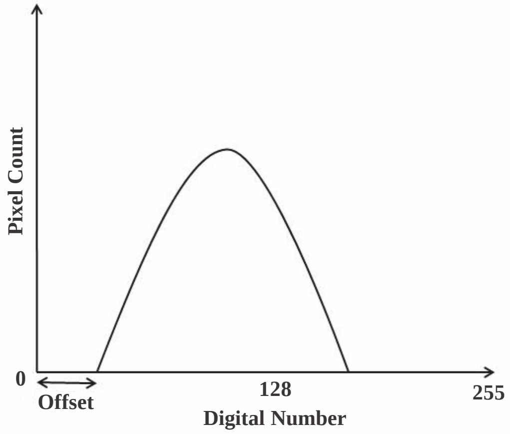


Fig. 11.5: Histogram with an offset (DN = 31) in brightness values

When multi-temporal imagery is being used for studying changes taking place in a study area, common radiometric quality is required for quantitative analysis of multiple satellite images of a scene acquired on different dates with same or different sensors. The radiometric quality could also be rectified using one image as a reference image. Transformed images appear to have been acquired with the reference image sensor, under atmospheric and illumination conditions nearly identical to those in the reference scene.

In order to achieve this, a few sets of scene landscape elements with a mean reflectance which is (almost) time invariant are identified. These elements are also known as *pseudoinvariant features*. The average gray level values of these reference sets are used to calculate a mathematical mapping relating the gray levels between the reference and the remaining images.

Direct Calibration Using Field-Derived Reflectance

This method is based on the assumption that reflectance measured in one region for a particular feature is directly applicable to the same feature occurring in other regions. The method requires some field work to measure true ground reflectance of at least two regions/targets of area covered by the image. The ground measurements are made using a spectral radiometer. Sometimes, large areas on the ground are painted in white and black as seen in Fig. 11.6 and recorded values over these areas are examined in different bands and calibration values are computed based on the relation between expected and recorded values in different bands.

Fig. 11.6: Preparation of test site for calibration purpose (source: Lecture slides of Prof. J.R. Jensen, University of South Carolina; used with permission)

Besides atmosphere, topography of the Earth surface also induces errors, which requires correction. The effects of topographic slope are:

- local variation in view and illumination angles and
- identical surface objects might be represented by totally different intensity values.

The goal of topographic correction is to remove all topographically caused variance, so that areas with same reflectance have same radiance or reflectance. Ideal slope-aspect correction removes all topographically induced illumination variation so that two objects having the same reflectance properties show the same gray levels despite their different orientation to the Sun's position. This requires digital elevation data of the area covered by the entire image as well as satellite heading along with Sun elevation and azimuth details.

Check Your Progress I

Spend 5 mins

1) List out the radiometric errors occurring in a satellite image.

....

....

....

....

11.4 GEOMETRIC DISTORTIONS AND THEIR CORRECTIONS

You have been briefly introduced to the causes of geometric distortions in subsection 11.2.1 of this unit. It is usually necessary to preprocess remotely sensed data and remove geometric distortion so that individual pixels are in their proper map locations. This allows remote sensing derived information to be related to other thematic information in Geographical Information System (GIS) or Spatial Decision Support Systems (SDSS). Geometrically corrected imagery can be used to extract accurate distance, polygon area and direction (bearing) information.

Let us further discuss about the nature of geometric errors.

11.4.1 Nature of Geometric Errors

Geometric errors present in remote sensing images can be categorised into the following two types:

- internal geometric errors, and
- external geometric errors.

Ephemeris refers to the geometry of the sensor system at the time of imaging as well as that of the Earth surface.

It is important to recognise the source of internal and external error and whether it is *systematic* (predictable) or *non-systematic* (random). As like systematic radiometric errors, systematic geometric error is generally easier to identify and correct than non-systematic or random geometric error.

Internal geometric errors are introduced by the sensor system itself and/or by the effects of Earth's rotation and curvature. These errors are predictable or computable and often referred to as *systematic* that can be identified and corrected using pre-launch or platform ephemeris.

Reasons of geometric distortions causing internal geometric errors in remote sensing images include the following:

- skew caused by the Earth's rotation effects
- scanning system induced variation in ground resolution cell size and dimensional relief displacement and
- scanning system tangential scale distortion.

Earth Rotation Effect: You know that Earth rotates on its axis from west to east. Earth observing sunsynchronous satellites are normally launched in fixed orbits that collect a path (or swath) of imagery as the satellite makes its way from the north to the south in descending mode. As a result of the relative motion between the fixed orbital path of satellites and the Earth's rotation on its axis, the start of each scan line is slightly to the west of its predecessor which causes overall effect of skewed geometry in the image.

Variation in Ground Resolution Cell Size and Dimensional Relief Displacement: You have read in Unit 4 of MGY-002 that an orbital multispectral scanning system scans through just a few degrees off-nadir as it collects data hundreds of kilometers above the Earth's surface (between 600 and 700 km above the ground level). This configuration minimises the amount of distortion introduced by the scanning system. In case of low altitude multispectral scanning systems, numerous types of geometric distortion may introduce that can be difficult to correct.

Tangential Scale Distortion: It occurs due to the rotation of the scanning system itself. Because when a scanner scans across each scan line, the distance from scanner to ground increases further away from the centre of the ground swath. Although scanning mirror rotates at a constant speed but the instantaneous field of view of the scanner moves faster and scans a larger area as it moves closer to the edges. It causes in the compression of image features at points away from the nadir as shown in Fig. 11.7. This distortion is known as tangential scale distortion.

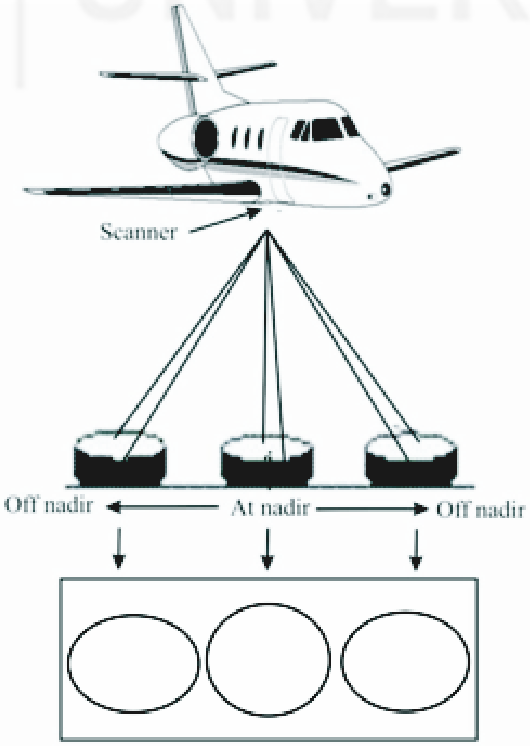


Fig. 11.7: Tangential scale distortion

External geometric errors are usually introduced by phenomena that vary in nature through space and time. The most important external variables that can cause geometric error in remote sensor data are random movements by the spacecraft at exact time of data collection, which usually involve:

- altitude changes and/or
- attitude changes (yaw, roll and pitch).

11.4.2 Geometric Correction

There are various terms which are used to describe geometric correction of remote sensing images and it is worthwhile to understand their definitions before proceeding.

In a 3-dimensional object, *yaw* refers to the direction of its orientation within xy plane, or rotating around z-axis

Roll of the object refers to its orientation within yz plane, or rotating around x-axis

Pitch of the object refers to its orientation within xz plane, or rotating around y-axis

Geometric correction is the process of correction of raw remotely sensed data for errors of skew, rotation and perspective.

Rectification is the process of alignment of an image to a map (map projection system). In many cases, the image must also be oriented so that the north direction corresponds to the top of the image. It is also known as *georeferencing*.

Registration is the process of alignment of one image to another image of the same area not necessarily involving a map coordinate system.

Geocoding is a special case of rectification that includes geographical registration or coding of pixels in an image. Geocoded data are images that have been rectified to a particular map projection and pixel size. The use of standard pixel sizes and coordinates permits convenient overlaying of images from different sensors and maps in a GIS.

Orthorectification is the process of pixel-by-pixel correction of an image for topographic distortion. Every pixel in an orthorectified image appears to view the Earth from directly above, i.e., the image is in an orthographic projection.

It is essential to remove geometric errors because non-removal of geometric distortions in an image may not be able to:

- relate features of the image to field data
- compare two images taken at different times and carry out change analysis
- obtain accurate estimates of the area of different regions in the image and
- relate, compare and integrate the image with any other spatial data.

Geometric correction is usually necessary but it is not required if the purpose of the study is not concerned with the precise positional information and rather with the relative estimates of areas.

You now realise that a geometrically uncorrected image is not of much use. There are following two common geometric correction procedures which are often used:

- image-to-map rectification
- image-to-image registration

Image-to-map rectification is the process by which the geometry of an image is made planimetric. Whenever accurate area, direction and distance measurements are required, image-to-map geometric rectification should be performed. It may not, however, remove all distortions caused by highly undulating terrain heights, leading to what are known as *relief displacement* in images. This process normally involves selecting some image pixel coordinates (both row and column) with their map coordinate counterparts (e.g., meters northing and easting in a UTM map projection).

Ground control point (GCP) is a specific position which consists of two pairs of known coordinates (reference and source coordinates).

Image-to-image registration is the translation and rotation alignment process by which one image is aligned to be coincident with respect to another image, thereby allowing the user to select a pixel (i.e. Ground control point - GCP) in one image and its positionally exact counterpart from the other image. The same general image processing principles are used in both image rectification and image registration. In case, an image that is already rectified to a map reference system is used as base image and second image also retains all geometric errors present in the base image. However, this approach is more appropriate when images of multiple dates are used for observing changes on the ground. This is because if two images are separately rectified to the map reference system each may have the same overall error but may be of a different nature, resulting in twice the individual errors when two rectified images are used together.

The general rule of thumb is to rectify remotely sensed data to a standard map projection whereby it may be used in conjunction with other spatial data in a GIS to solve problems. Therefore, most of our discussion here will focus on image-to-map rectification. However, irrespective of the procedure used, the process of geometric correction involves five steps as shown in Fig. 11.8.

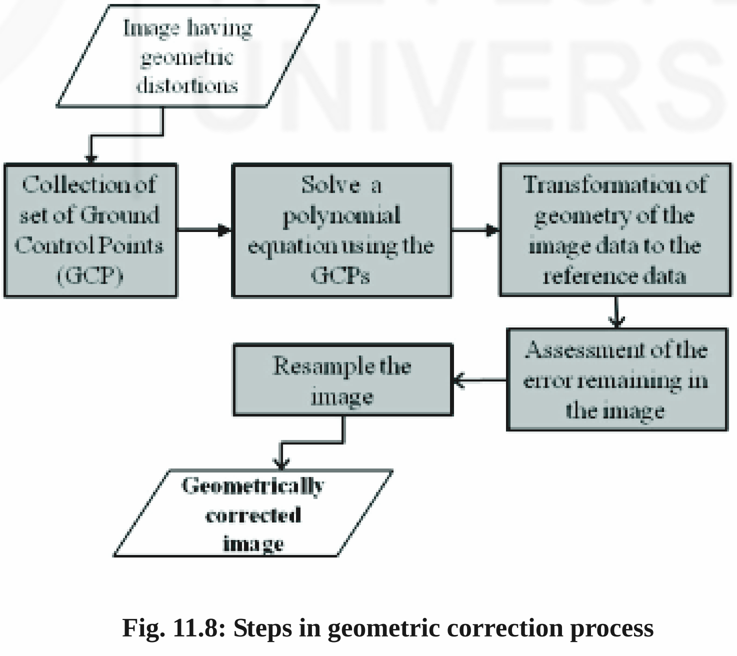


Fig. 11.8: Steps in geometric correction process

The best GCPs are the intersections of roads, airport runways, large buildings, edges of dams and other stable features.

Let us now discuss in detail about the steps in geometric correction process.

Step 1: Collection of Ground Control Points (GCPs)

You have read that geometric distortions introduced by sensor system attitude (roll, pitch and yaw) and/or altitude changes can be corrected using GCPs and appropriate mathematical models. You would take a simple approach to achieve this in order to align the image with a national reference map. You should learn that an important concept here is GCP which is a location of features on the surface of the Earth (e.g., a road intersection) that can be identified on the imagery and also located accurately on a map. The image analyst must be able to obtain following two distinct sets of coordinates associated with each GCP:

- image coordinates specified in i rows and j columns, and
- map coordinates (e.g., *x, y* measured in degrees of latitude and longitude, or meters in a polyconic or Lambert conformal conic projection or a more widely adopted Universal Transverse Mercator (UTM) projection.

Step 2: Solving a Polynomial Equation Using GCPs

After identifying GCPs, the next step is to transform row and column coordinates of each pixel into coordinates of the reference map projection. This is achieved using a transformation matrix which is derived from GCPs. The transformation matrix contains coefficients calculated from a polynomial equation.

The paired coordinates from many GCPs (e.g., 20) can be modelled to derive *geometric transformation coefficients* relating to geometry of the image coordinate system and the reference map coordinate system. These coefficients may be used to geometrically rectify remotely sensed image to a standard datum and map projection.

Image-to-map rectification involves computing coefficients of transformation between the image and map coordinate system using GCPs. Each GCP contains row-column coordinates of a clearly identified point in the image and its corresponding location in latitude-longitude or metres (East) and metres (North) with reference to some origin on map. The transformation is normally represented by a polynomial equation whose coefficients are determined by solving the system equations formed using GCPs.

Step 3: Transformation of the Image to the Geometry of the Reference Map/Image (Spatial Interpolation)

An important consequence of correcting images for geometric distortions is that the raster grid for master dataset (map or image) and that of slave (image) would not match. In such a case, gray level for each cell needs to be recomputed for slave image after that it is geometrically rectified. The determination of locations of slave pixels in the reference grid is a spatial transformation from one coordinate system to another.

This type of transformation can model following kinds of distortions in the remote sensor data:

- translation in x and y
- scale changes in x and y
- skew and
- rotation.

Using six coordinate transform coefficients that model distortions in the original scene, it is possible to use the output-to-input (inverse) mapping logic to transfer (relocate) pixel values from the original distorted image (x', y') to the grid of the rectified output image, (x, y).

Where, x and y are positions in the *output*-rectified image or map and x' and y' represent corresponding positions in the original *input* image.

Step 4: Assessment of Error

It is important to remember that the units of RMS (root-mean-square) error are the units of the source coordinate system.

Though geometric correction process corrects much of the geometric error present in the original image, not all of the errors are removed. Hence, prior to creating output rectified image, it is essential to test quality of fit between the two coordinate systems based on coefficients. In other words, it is required to assess the accuracy of the polynomial transformation. Each GCP has influence in value of coefficient in the transformation matrix and some GCPs would introduce large amounts of error. The method used most often for calculating total error involves the computation of *root-mean-square error* (RMSE or RMS_{error}) for each of GCP. RMS_{error} is the distance between GCP original source and re-transformed coordinates. It is calculated using Pythagoras's theorem as

$$RMS_{error} = \sqrt{\left(x' - x_{orig}\right)^2 + \left(y' - y_{orig}\right)^2}$$

where, x_{orig} and y_{orig} are the original row and column coordinates of GCP in the image and x' and y' are the computed or estimated re-transformed coordinates in the original image when we utilise six coefficients. The square root of the squared deviations represents a measure of accuracy of each GCP. By computing RMS_{error} for all GCPs, it is possible to

- see which GCPs contribute the greatest error and
- sum all RMS_{error}.

You may note that not all of the original GCPs selected are used to compute final six-parameter coefficients to rectify input image. This involves several cycles, wherein all GCPs are initially used to find coefficients. RMS_{error} associated with each of these initial GCPs is computed and summed. Then, the individual GCPs that contributed the greatest amount of error are determined and deleted. After the first iteration, this would leave you with a reduced number of GCPs. A new set of coefficients is then computed using these GCPs. The process continues until RMS_{error} reaches an acceptably low value (e.g., ≤ 1 pixel error in the x-direction and ≤ 1 pixel error in the y-direction). When the acceptable threshold is reached, the final coefficients and constants are used to rectify input image to an output image in a standard map projection as previously discussed.

Pixels are resampled so that new pixel values for the transformed image can be calculated.

Step 5: Resampling (Intensity Interpolation)

The next step is to create output image by assigning values to pixels on new grid of the transformed image by interpolation from the unrectified image. This process is called *resampling* or *intensity interpolation*. You have read earlier that when pixel positions are recomputed through transformations controlled by GCPs, a pixel in the reference image would overlap with several pixels in the unrectified input image. Therefore, it is necessary to find pixel values in the transformed image through intensity interpolation.

Intensity interpolation (resampling) involves the computation of pixel value at a pixel position x′, y′ in the original (distorted) input image and its relocation to the appropriate x, y coordinate location in the rectified output image. This process is used to produce output image row-by-row and pixel-by-pixel. Most of the time, x′ and y′ coordinates to be sampled in the input image are real numbers (floating point), which do not coincide with pixel positions. In such a case, you can employ one of the methods of intensity interpolation (resampling) mentioned below:

- nearest neighbour (zero order) method
- bilinear interpolation (first order) method and
- cubic convolution (second order) method

Nearest neighbour method uses the closest pixel's pixel value to assign to output pixel value. In the bilinear interpolation method, values of four pixels in a 2 x 2 window are used to calculate an output pixel value with a bilinear function. And, in the cubic convolution method, values of sixteen pixels in a 4 x 4 window are used to calculate an output pixel value with a cubic function. Of the three methods, the nearest neighbour method is the least computationally intensive whereas cubic convolution method is the most computationally intensive method. Nearest neighbour method is used for a majority of applications involving digital interpretation because it transfers original pixel values without averaging them whereas in other two methods pixel values are altered.

In all the three methods, number of rows and columns of pixels in the output image is calculated from the dimensions of output image, which is determined by the geometric transformation and cell size.

Check Your Progress II

Spend 5 mins

1) Name the sources of geometric distortions in a satellite image.

....

....

....

....

....

11.5 SUMMARY

- Radiometric and geometric corrections are the most important operations that help to utilise information contained in the remotely sensed images in a standard reference framework along with other geospatial datasets.
- Image processing and classification operations may not depend on the geometric rectification step but radiometric correction is extremely important for data processing and analysis steps.

- Geometric correction places result on its correct location on a geographic reference system whereby it can be used, for instance, to detect changes between the content of a map and image, and also to update earlier date maps with new information as observed in the latest image.

Spend 30 mins

11.6 UNIT END QUESTIONS

1) How is shot noise eliminated from an image?
2) What is atmospheric correction? List the methods of atmospheric correction.
3) What is GCP?
4) What are the different methods of intensity interpolation (resampling)?

11.7 REFERENCES

- Chavez, P. S. Jr, (1988), An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, Vol 24, pp 459-479.
- Jensen, J. R., (2005), *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice Hall, New Jersey (for lecture slides of Prof. J. R. Jensen, University of South Carolina).

11.8 FURTHER/SUGGESTED READING

- Jensen, J. R., (2005), *Digital Image Processing: A Remote Sensing Perspective*, Prentice-Hall, New Jersey.
- Mather, P. M., (2004), *Computer Processing of Remotely Sensed Images-An Introduction*, John Wiley & Sons Ltd, New Delhi.
- Schowengerdt, R., (2007), *Remote Sensing: Models and Methods for Image Processing*, Academic Press, Burlington.

11.9 ANSWERS

Check Your Progress I

Radiometric errors are mainly divided into two types such as internal and external errors. Internal errors are caused by the electronics themselves. But the external errors are a result of phenomena such as atmospheric disturbances and steep terrain undulations.

Check Your Progress II

Earth rotation and its curvature, satellite altitude and perturbations in satellite nominal altitude are the main sources of geometric distortions in a satellite image.

Unit End Questions

1) Refer to subsection 11.3.2.
2) Refer to subsection 11.3.3.
3) Refer to subsection 11.4.2.
4) Refer to subsection 11.4.2.