Chapter 4

Image Processing Methods

July 2, 03

Peng Gong, University of California, Berkeley

4.1 INTRODUCTION

Like any remote sensing project, a set of objectives is first identified in human settlement remote sensing. For instance, remote sensing of human settlements may be for the purpose of land use mapping, urban change detection, assessment of environmental pollution or health of urban forests, urban heat island studies, or population estimation. Once a particular project objective is identified, the investigator usually proceeds through three consecutive steps: (1) determination of imagery and data needs and acquisition from available sources; (2) determination of data and image processing to accomplish the objective; and (3) selection and/or development of a strategy for information extraction from the remotely sensed data. For example, in order to detect and map urban change, one needs to have multi-temporal data. The spatial and spectral resolution of the imagery determine the level of spatial and environmental detail that are detectable. The season of the year is an important consideration given the variation in vegetation, solar angle variation and other environmental features; consequently, images taken near anniversary dates are desirable for change detection. In order to map patterns of urban heat island, thermal images need to be collected. Different data types have different requirements for image processing. Finally, there are various information extraction approaches that can be followed to derive the meaningful information to meet specified project objectives. Figure ----- in Chapter ----- is a useful guide for identifying steps toward a particular objective and project. In this chapter, some of the commonly used *image processing* procedures will be reviewed. Strategies of *information extraction* will be introduced in Chapter 5.

4.1.1 Data Availability

Since the launch of the Earth Resource Technology Satellite-1 (ERTS-1, later became Landsat –1) in 1972, satellite imaging for earth observation has significantly changed our ways of monitoring and mapping human settlements. In particular, the constant improvement of the spatial resolution by various sensors from the original multispectral scanner (MSS) with 80 m resolution to the order of meter or sub meter levels makes detailed observation gradually possible to scientists, professionals, policy makers alike. These have been evidenced by the advent of the 30 m resolution Landsat Thematic Mapper (TM) in 1982, 10-20 m resolution SPOT High Resolution Visible (HRV) in 1986, the 5.8 m resolution India Remote Sensing Satellites in early 1990s, 1-4 m resolution Space Imaging's IKONOS in 1999 and the 0.67 – 2.6 m resolution Digital Globe's Quickbird in 2001. Since the launch of Earth-Observing –1 (EO-1) in 2000, hyperspectral images with a spectral resolution of approximately 10 nm and a spatial resolution of 30 m are also available from space. Satellite imaging is approaching the spatial quality of small-scale aerial photographs together with more spectral bands and better radiometric capabilities. Figure 4-1 shows example images acquired from some recently developed sensors with different resolution and spectral sampling. Additionally, the spectral imaging capabilities have been advancing both in spectral coverage and spectral resolution. Satellite images are not only widely available in the visible, near infrared, and shortwave infrared (1-2.5 μm) regions, but also in the thermal and microwave regions. These are exemplified by the increased thermal imaging capabilities with such sensors as Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Enhanced TM+ on board Terra and Landsat 7, respectively, and the ERS, JERS, and Radarsat radar imaging from space. All these developments have enormously increased the volume of data available to human settlement remote sensing. The diversity of data increases the demand for innovative image processing and information extraction.

4.1.2 Primary Components in Image Processing

Image processing consists of procedures that produce new images from old ones. It involves all operations applied to any form of images, e.g., raw images acquired freshly from remote sensors, images contaminated by noise such as clouds and shadow, or images distorted by sensor malfunctioning, sensor platform

instability, etc. The general purpose of image processing is to prepare image data that can be better utilized in subsequent image interpretation or information extraction.

Image processing procedures can be broken into three broad categories (Figure 4-2): *image restoration*, *image enhancement*, and *image compression*. Image restoration can be further divided into radiometric correction and geometric correction. These are presented in sections 4.2 and 4.3 respectively. Image enhancement is accomplished through global, local, and/or multispectral operations, and image fusion. A variety of processes can be employed such as histogram adjustment, filtering techniques, image arithmetic operations, and image transformations. Some of the techniques will be discussed in section 4.4. Employing various coding and transformation techniques to reduce data volume while preserving all or a major part of the image information, image compression serves the purpose for data storage and transmission. Some techniques can be applied in sequences for image enhancement and information extraction purposes. For example, Fourier transform can be applied to a single image and the transformed image in frequency domain can then be applied to image filtering and texture analysis. Wavelet transform has been widely used in image compression, texture analysis, and image fusion. In this chapter we will concentrate on some techniques of image restoration and image enhancement (section 4.4) that are important for remote sensing of human settlements.

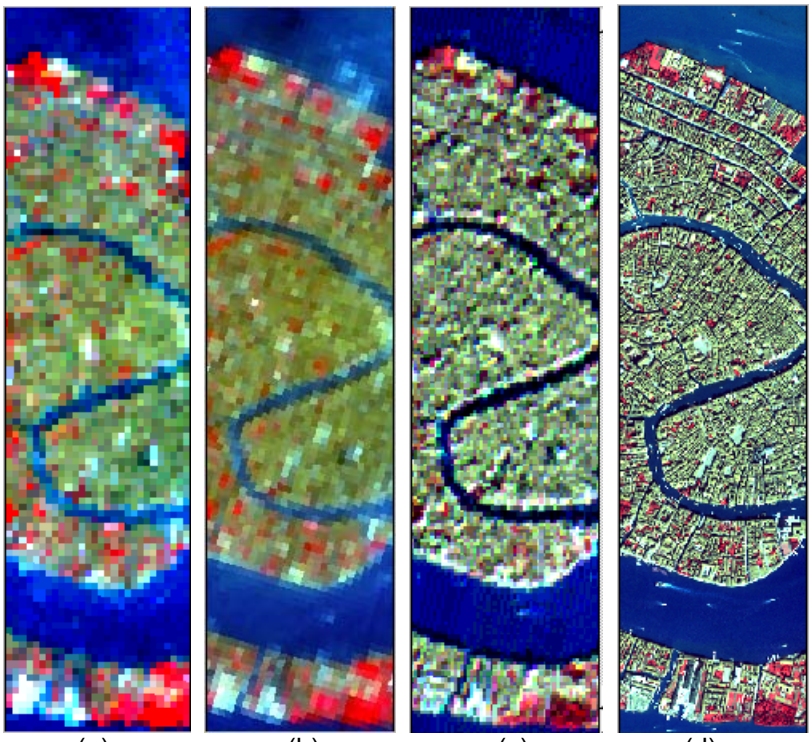


Figure 4-1. Examples of images acquired over Venice, Italy, by sensors on board of some recent satellites. a. Landsat ETM+ image (30 m, 6 spectral bands), b. EO-1 Hyperion image (30 m, 224 bands), c. Terra ASTER image (15m, 3 visible and NIR bands, 30 m, 6 shortwave infrared bands) and d. IKONOS image (4 m, 4 bands). (Image courtesy of Michael Abrams) [W]

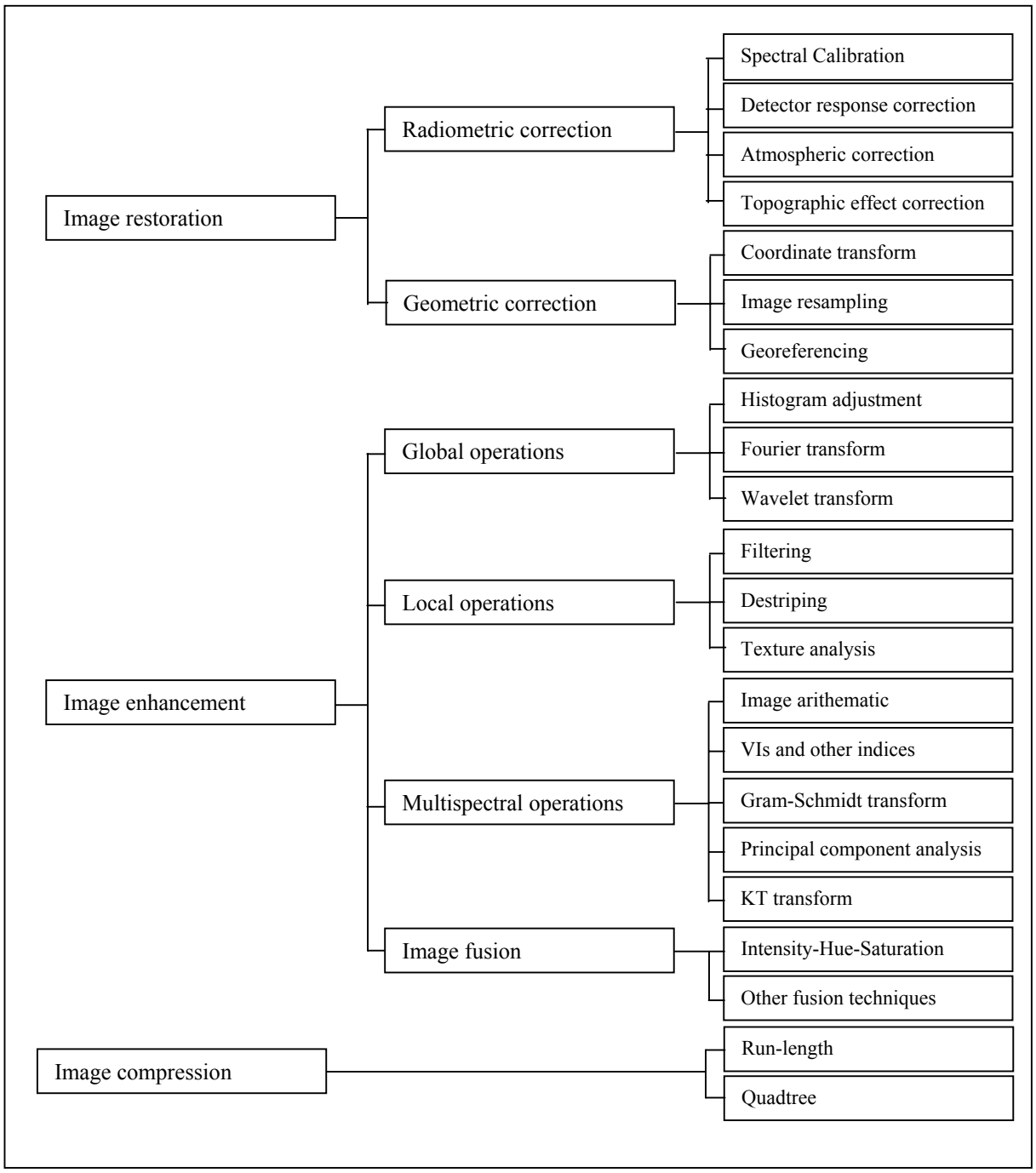


Figure 4-2. Primary components of image processing. [B/W]

4.2. RADIOMETRIC PROCESSING

Image radiometry is affected by factors such as system noise, inconsistent detector responses, sensor malfunction, atmospheric interference, and differences in illumination and viewing geometry caused both by differences in surface morphology (or topography) and time of image acquisition. Correspondingly,

radiometric processing, as part of the image processing task (Figure 4-2), aims at removing or reducing these effects. It includes radiometric calibration for removal of some of the system noise and rectification of detector responses, remediation of sensor malfunction, atmospheric correction for reduction of the atmospheric interference, illumination angle, and viewing angle correction for normalization (or standardization) of imaging conditions.

Sensor system noise, detector responses, and spectral calibration are dealt with by sensor manufacturers or data suppliers. System noise is caused by dark currents of instrument when no external radiance is received by the sensor. Detector inconsistency can be calibrated by using standard light sources with known radiometric intensities. Known specific absorption features are used to calibrate any spectral wavelength displacement to achieve the goal of spectral calibration. Once these are processed, the radiance that reaches a sensor is expressed by

$$L_S = K \cdot DN + L_{min} \ (Wm^{-2} \ Sr^{-1}) \quad (4\text{-}1)$$

$$K = \frac{L_{max} - L_{min}}{DN_{range}} \quad (4\text{-}2)$$

Normally L_{max}, L_{min} and DN_{range} are known from the sensor manufacturer or operator. However, L_S is composed of contributions from the target, background and the atmosphere (Figure 4-3). In Figure 4-3, i is the solar incidence angle, e is the viewing angle.

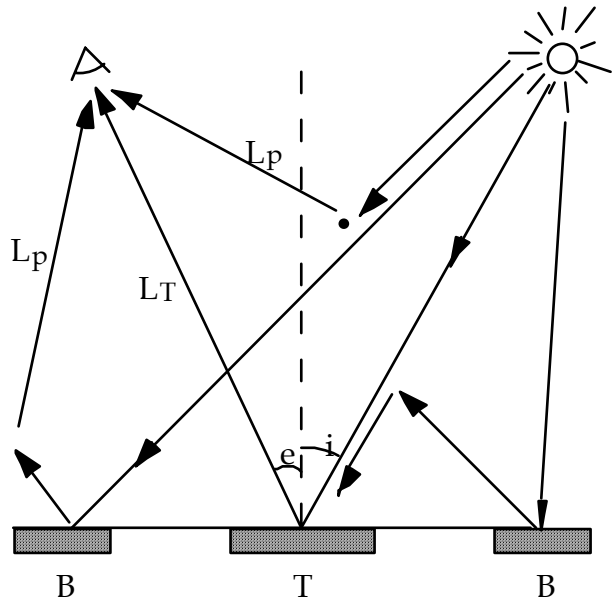


Figure 4-3. Light interaction with the atmosphere and the target (T) and background (B). Part of the scattered radiation is received by the sensor. [B/W]

The radiance recorded by the sensor contains two parts of the energy: one part carrying surface target signal traveling through the whole course from the illumination source, usually the sun, to earth surface and from the surface to the sensor, and the other part from the illumination source to the atmosphere and then to the sensor. The second part does not contain information about earth surface. During the process of energy transmission, the atmosphere plays two roles to attenuate the transmission of the energy: absorbing energy, a role of blocking, and *scattering*, a role of dispersing lights into different directions. The *absorption* role

varies with the change of atmospheric constituents and the scattering role is dependent of the size of particles in the atmosphere. Different constituents have different *absorption features* (strong light absorption at certain wavelengths). The abundance of certain constituents determines the strength of the absorption (low transmission valleys in Figure 4-4). It is well known that there is a relatively stable proportion of nitrogen, and oxygen in the atmosphere and their absorptions of light in the wavelengths of our concern are negligible. So does the CO2 in the atmosphere. The variable components of the atmosphere are water vapor, ozone and aerosols. Among these, the first plays a significant role in light absorption in certain wavelengths and the aerosols severely affect light transmission through scattering. Therefore, in order to estimate the atmospheric attenuation of light, we need a good knowledge of the amount of water vapor, ozone, and aerosols.

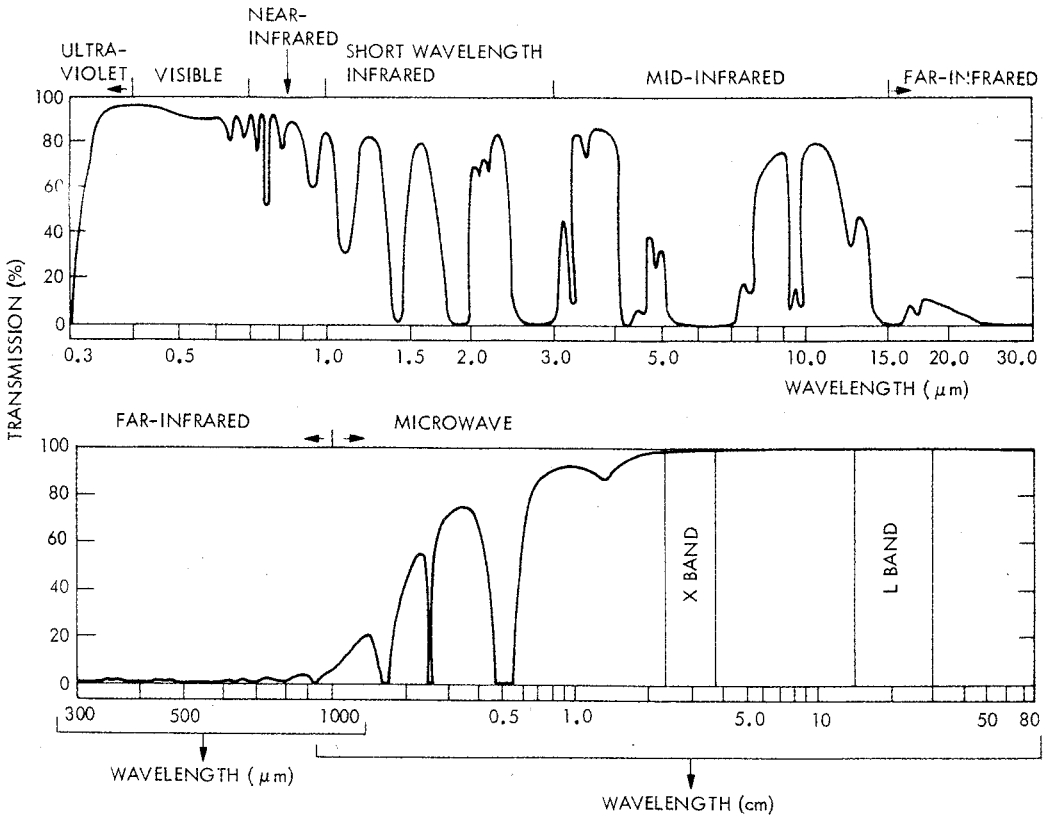


Figure 4-4. Spectral transformation of the atmosphere (after Colwell, 1983) **[B/W]**

Different ratios, r, between atmospheric particle size and wavelength of light have different scattering effects on light transmission. When $r << 1$, the particles have a strong wavelength-selective scattering effect, called *Rayleigh scattering*, i.e., the scattered amount of light is proportional to $1/\lambda^4$. It passes approximately 95% of the light in the forward directions. For wavelengths in the visible and near infrared, this is the case when the atmosphere is primarily composed of small particles at the molecular level, i.e., when the sky is crystal clear and the humidity is low. When r is close to one, the particles in the atmosphere has Mie scattering effect, i.e., the scattered amount of light is proportional to $1/\lambda^{1\sim2}$. It split the energy into 50% and 50% in the forward and backward directions. When $r >> 1$, the particles in the atmosphere have non-selective backward scattering.

In summary, the atmosphere affects the transmission of energy in three ways. Firstly, it modifies the spectral and spatial distribution of the radiation incident on the surface. Secondly, radiance being reflected

by ground target is attenuated. Thirdly, radiance scattered by the atmosphere called *path radiance*, L_p, is added to the transmitted radiance, L_T, toward the sensor. Assuming that L_S is the radiance received by a sensor, it can be divided into L_T and L_p

$$L_S = L_T + L_p \tag{4-3}$$

Clearly, L_T contains information about the target.

For a given spectral interval $[\lambda_1, \lambda_2]$ the solar irradiance reaching the earth's surface, E_G, is

$$E_G = \int_{\lambda_1}^{\lambda_2} (E_s T_i \cos i + E_d) d\lambda \tag{4-4}$$

where E_S is the solar irradiance outside the atmosphere, T_i the atmospheric transmittance along the incident direction, and i the incident angle, and E_d the diffuse sky irradiance.

Spectral properties of different surfaces are very complicated and can range from specular to diffusive. Conventionally, most surfaces are approximately considered as diffuse reflectors at high solar elevations, i.e. when i is small. If the surface is assumed to be a perfect diffuse reflector, i.e. the Lambertian case, which is rarely true in reality, then the surface reflectivity is a constant, ρ. The ratio of the radiation reflected in the viewing direction to the total radiation into the whole upper hemisphere is given by $\frac{1}{\pi}$. Based on the Lambertian assumption,

$$L_T = L_S - L_p = \frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \rho T_e (E_s T_i \cos i + E_d) d\lambda \tag{4-5}$$

where ρ is the target reflectance, T_e is the transmittance along the viewing direction. Therefore, in order to quantitatively analyze remotely sensed data, i.e. to find ρ, atmospheric transmittance T and path radiance L_p have to be known. One way to approximately estimate L_p is to use the radiance difference between shadow areas and sun-lit areas of the same type of surface. Another method is to use dark targets.

In reality, target spectral properties are a function of not only the physical properties of the surface cover but also the incident and viewing angles. This function is called bi-directional reflectance distribution function (BRDF). Because BRDF is more complicated than a Lambertian case, it has rarely been analyzed for most human settlement applications of remotely sensed data. Interested readers are referred to Hapke (1993) for detailed treatment of this subject.

Radiometric processing can be divided into absolute calibration and relative calibration. *Absolute calibration* includes radiometric calibration and correction. They convert image gray-level values (digital numbers) to absolute radiances. The goal is to reduce radiometric error, i.e., reduce the difference between radiance measurements by the sensor for targets of interests and the actual radiances for the same targets. *Relative calibration* includes procedures that can be used to balance the responses of one detector against others. Sometimes we can balance (standardize, or normalize) one image against another if some common parts of the images can be considered not changing over time. Gray-level value conversion to radiance may not be necessary. Obviously, radiometric processing is helpful to change detection because sources of radiometric distortion are reduced. The more accurate we can achieve for radiometric processing, the more universal can other analysis methods be used for information extraction purposes.

The requirement for accuracy in *radiometric correction* for remote sensing of settlements is not as rigorous as some other applications when daily images are required over a period of time longer than a year such as the analyses of annual trend of vegetation with NOAA AVHRR or Terra MODIS data. Therefore, in

applications of remote sensing for human settlements a thorough radiometric processing can hardly be found. Instead, simple and relative methods are often used. These methods include dark target correction, image radiometric calibration based on measurement from pseudo-invariant targets, solar angle and view angle normalization (when a DEM [digital elevation model] is used this reduces topographic effects as well). In the rest of this section we introduce some of the often-used radiometric processing methods.

4.2.1 Detector Response Calibration

System noise refers to the white noise and a systematic noise in the electric-optical system of a sensor. While the white noise cannot be removed and it determines the signal to noise ratio (S/N, affecting the radiometric resolution) of a sensor system, the systematic noise caused by environmental conditions at the time of operation can be removed by subtracting the *dark current* of a sensor system measured with no external radiation. The Landsat TM sensor has 16 detectors for each of the visible, near infrared, and middle infrared bands, while the SPOT HRV sensor has 3000 to 6000 detectors in a band. The problem is that no detector functions the same way as another. If the problem becomes serious, the image acquired by such sensors will have banding or striping effects (Figure 4-5a). This can be corrected through a radiometric calibration procedure. By exposing different levels of known radiation to all detectors, their different responses are recorded. A regression between detector response, DN_r, and the known amount of radiation, DN_c, allows us to derive a response function (usually a linear one) for each detector. The linear regression takes the following form

$$DN_c = a \cdot DN_r + b \tag{4-6}$$

where coefficients a and b are obtained using a least squares method.

A *radiometric calibration* can then be made by applying the response function of each detector to its measured radiances to generate a better quality image (Figure 4-5b). Once, each detector is calibrated, the calibrated image data (digital numbers) can be converted into radiance or spectral reflectance. For the Landsat MSS and TM, these data can be found in Markham and Baker (1986). For other type of sensors, users can obtain calibration information from the sensor manufacturer or conduct their own calibration using standard light sources. Usually this will remove the inconsistencies of detector response. If after radiometric calibration, the image still has striping effects, a relative balancing of the detectors through histogram matching (Horn and Woodham, 1979; see section 4.4.1) or balancing the mean and standard deviation (Richard and Jia, 1999) can be used.

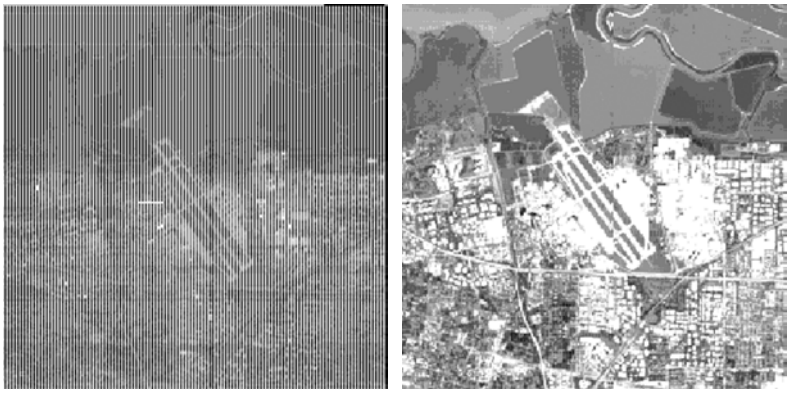


Figure 4-5. a. A raw Advanced Land Imager (ALI, 30 m resolution) Band 2 (0.525-0.605 μm) image (Level 0 product) of Moffett Field, California, acquired on board EO-1 showing the striping caused by detector response inconsistency; b. The radiometrically calibrated ALI Band 2 image (Level 1 product). (Image courtesy: GSFC, NASA) [B/W]

4.2.2 Gray-level Value to Radiance or Reflectance Transformation – Simple Atmospheric Correction

This is usually done after detector response correction in order to reduce atmospheric effects. One of the simplest techniques is a linear transformation between image gray-level value and radiance or reflectance. Therefore, it requires the measurements of spectral radiances or reflectances in the field to be imaged at the same time of sensor overpass. However, it is usually impossible to do simultaneous spectral measurements at the time of imaging because a large number of parallel operations in the field may be required. A *pseudo-invariant target* approach is often adopted. Spectral reflectance of targets such as parking lot, roof tops, pavements, water bodies, gravel pits, sandy lands, may be considered not changing over time. Therefore, their spectral reflectance can be measured before or after an image is taken. Even so, similar type of weather condition and time of the day should be chosen. To assure a good linear relationship between image gray-level values and spectral reflectances to be derived, targets with larger ranges of spectral reflectance should be chosen. Spectral data collected in the field with a field spectrometer should be processed to have similar spectral wavelength ranges to those of the sensor (this is usually done through averaging and merging of spectral bands of the spectrometer). A linear regression can be made in the same manner as in detector response calibration. When a linear regression function is established, it can be applied to convert gray-level values into spectral reflectances.

This empirical method has been widely used in remote sensing (Forster, 1980; 1985, Gong et al., 1994). This method can be considered as a simple method for *atmospheric correction* in which the offset and gain of a linear equation account for the path irradiance and multiplicative factors of the atmosphere, respectively.

4.2.3 Radiometric Normalization

A problem associated with the use of multi-temporal remote sensing data for change detection is that data are usually acquired under different sun angle, atmospheric, and soil moisture conditions. Such data can be adjusted through a process called *radiometric normalization*, so that the effects of those undesirable conditions can be minimized or eliminated (Hall *et al*., 1991). This process normalizes an image according to another reference image. The linear equation can be built based on a scatterplot between the image to be normalized and the reference image (Elvidge and Ding, 1995). Although the same linear procedure is used as in the previous sections, radiometric normalization is a relative calibration as it does not convert image gray-level values to their actual reflectances. The reference image should be chosen to contain the smallest amount of external effects. For example, it should be acquired when the solar zenith is high and the atmospheric condition is good to reduce shadow and minimize atmospheric interference.

For any regression, an R^2 value can be used to evaluate the adequacy of the regression equation. In general, an R^2 of 0.85 or greater is considered good for radiometric normalization, but a higher value is preferable for the previous methods.

4.2.4 Atmospheric Path Radiance Determination

Under clear sky condition when Rayleigh atmosphere dominates, a dark target may be used to determine the atmospheric path irradiance. Since Rayleigh scattering affects strongly in short wavelength particularly visible and we know that clear-deep water has a very low spectral reflectance in the visible and near and middle infrared wavelength region. We can use such water bodies in a *dark target* approach to estimate atmospheric path radiance. If a water body greater than 10-20 pixels in diameter can be found in an image, we can use the radiance of water derived from the image as L_W and the radiance of water, L, measured on site or assumed to estimate L_p.

$$L_W = K \cdot D_N \text{ (water dody)} + L_{min} \qquad (4\text{-}7)$$

$$L_p = L_w - L \quad (4\text{-}8)$$

L_p can then be subtracted from other radiances in an image for the visible channels. For the infrared channels, water completely absorbs radiance incident to it. What is recorded in an image of water bodies can be considered as the upward path radiance of L_p.

Sometimes, the effects of thin clouds on the visible and shorter than 1.0 μm images can be largely removed using information available in the longer wavelength if such information is available. Gao et al (1998) used an AVIRIS image over Bowie, Maryland to reduce cloud effects. On a 1.38 μm water vapor absorption image showing only cirrus clouds, the correction of cirrus scattering effects is done in a simple way. Using a visible band image at 0.66 μm and the 1.38 μm image (Figure 4-6), they built the following linear model to restore the radiance in 0.66 um:

$$R\_cir\_corr\ (0.66\ \mu m) = R(0.66\ \mu m) - c * R(1.38\ \mu m) \quad (4\text{-}9)$$

where R_cir_corr (0.66 μm), R(0.66 μm) and R(1.38 μm) are respectively the corrected image, and the raw images in the corresponding bands; c is derived from the scatter plot of R(1.38 μm) vs R(0.66 μm). Because of strong water vapor absorption below cirrus clouds, the sensor at the 1.38-μm channel only received solar radiation scattered by cirrus clouds. The solar radiation on the downward path transmitted through cirrus layer is absorbed by water vapor below cirrus. The cirrus path radiance in the visible is proportional to the 1.38 μm radiance.

On the top of Figure 4-6, a color composite of the AVIRIS image acquired on July 7, 1996 is shown (red: 0.66 μm; green: 0.55 μm; and blue: 0.47 μm). From this image, the residential areas can be seen contaminated by puffy cirrus clouds. In the middle is the 1.38-μm channel AVIRIS image. This channel detected only cirrus cloud in the upper part of the troposphere. The solar radiation transmitted through cirrus clouds on the downward path was absorbed by water vapor below cirrus clouds. As a result, this channel did not see surface background. The bottom of Figure 4-6 shows the cirrus-corrected image using the method described above.

Using the *aerosol optical path* information derived from TM band 7 (2.10 – 2.35 μm) based on a "dark object" method, Liang et al (1997) developed an operation atmospheric correction method to correct for the atmospheric effect on Landsat TM imagery over the land. Here the dark objects are dense vegetation or even dark soil in TM band 7. The reflectances of dark objects are used to estimate the aerosol optical depth. Other atmospheric parameters are estimated following assumptions of atmospheric conditions. Atmospheric correction can then be applied to all the bands.

4.2.5. Application of Atmospheric Radiative Transfer Models

Atmospheric correction procedures introduced earlier are simple and straightforward. They are most widely used in the remote sensing of human settlement. However, they can only be used to reduce the first order atmospheric attenuation. As illustrated in Figure 4-3, the atmospheric interference with light is complicated and contains multiple scattering that cannot be fully addressed by those approaches mentioned above. The alternative is to model the radiative transfer process in the atmosphere. This can be done through the use of some well-established radiative transfer models such as MODTRAN and 5S. Both models require as their inputs parameters about atmospheric constituents and abundances. Although it is desirable to measure

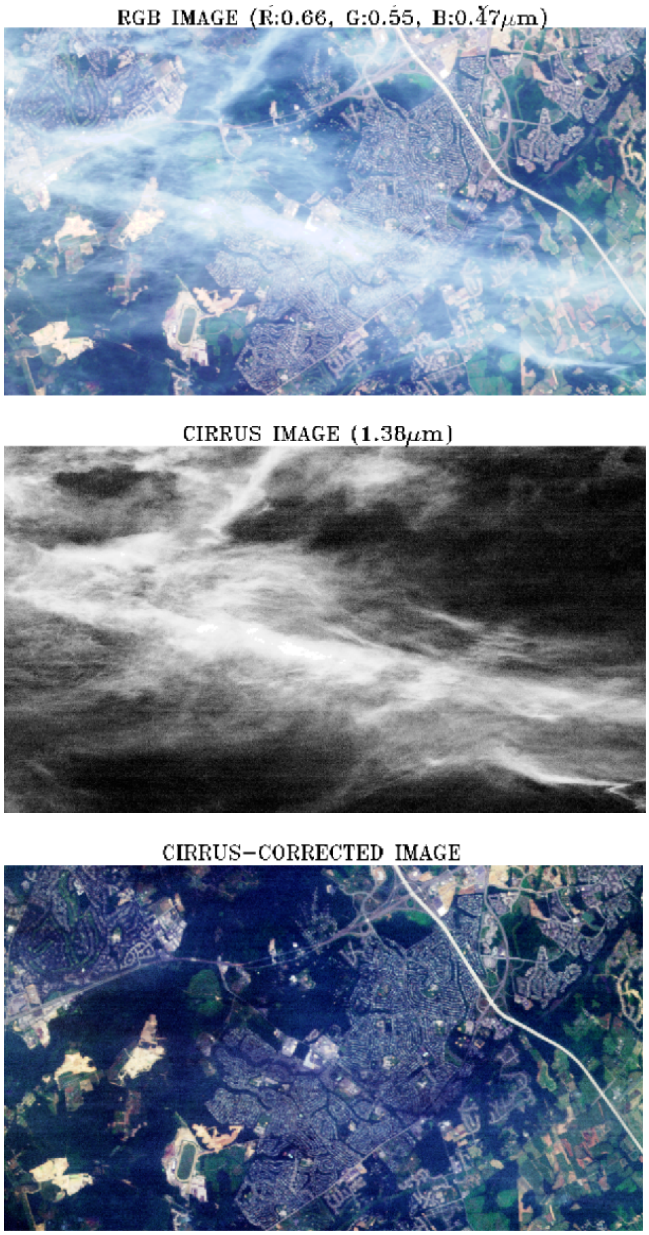


Figure 4-6. Cirrus cloud effect reduction using water absorption images. (Courtesy of Rong Rong Li and Bo-Cai Gao) [B/W]

the atmospheric parameters at the time of image acquisition, this is usually impossible to do as atmospheric condition changes from time to time and from location to location. Therefore, previous application of those models approximated atmospheric conditions through the use of one or two parameters such as the atmospheric visibility and humidity obtainable from weather stations in the area of interest.

A number of radiative transfer codes are available. These include MODTRAN, 5S (Simulation of the Satellite Signal in the Solar Spectrum 5S) and 6S (Second Simulation - aircraft, altitude of target). There are FORTRAN codes available for these algorithms (Air Force Research Laboratory, 2001). The 5S and 6S are proposed by Tanre and his colleagues (e.g. Tanre et al., 1990). Other atmospheric correction procedures include ATREM, FLASH, HATCH and ACORN (EO-1 Special Issue). Anyone of these can be used with various levels of complication. They can be used to carry out sophisticated atmospheric

corrections of remotely sensed data. The key to the successful application of these models is the good knowledge of the atmospheric conditions at the time of image acquisition. As the increase of the remote sensing capabilities, sensors such as MODIS and those on-board EO-1 are designed to include spectral bands allowing simultaneous extraction of parameters about the surface and atmospheric conditions. It is hopeful that in the future, parameters interpreted from those atmospheric bands can be used in the atmospheric correction of those bands for surface information extraction. Atmospheric correction must be done with great care. Since there are many factors to be accounted for, if the estimations are not properly made, the atmospheric correction might introduce new errors to the corrected image.

4.2.6 Summary

Radiometric calibration usually converts DN in a raw image into radiance or reflectance. This process includes sensor calibration, atmospheric correction, and *topographic correction* (see Figure 4-2; Gong, http://camfer.cnr.berkeley.edu/~gong/textbook). Among these, sensor calibration is usually done by data providers while the remaining two types of correction are done by users. For remote sensing of urban settlements, topographic correction is not very applicable due to the severe modifications of smoothly varying terrain by human settlements. Correction for atmospheric effects can either be done using information obtained from on flight measurements and field measurements of atmospheric conditions or using information available in the remotely sensed image as for the cases of cirrus cloud effect correction using the 1.38 water absorption band (Gao et al, 1998) and the aerosol optical path determination (Liang et al, 1997). On the other hand, it is not always necessary to carry out radiometric calibration. For instance, it is unnecessary to apply atmospheric correction to single image classification as long as the training data are specified for each single-date image. For Landsat TM data where the primary atmospheric distortion is due to scattering which is additive to the original signal, atmospheric correction is essentially translating each image band without changing the internal data structure. Such operations are not necessary for multidate image classification, post-classification comparison, change detection based on image differencing, and linear transformations of multiple images such as KT transform and PCA (Song et al, 2001). It is necessary to perform atmospheric correction if multiband ratioing (such as the normalized difference vegetation index NDVI) is used in change detection. Absolute correction such as the use of radiative transfer models requires knowledge about atmospheric optical path and abundance of various atmospheric particles. Relative calibration does not require any extra information and it balances radiances from one image against those of another.

4.3 GEOMETRIC PROCESSING

Geometric processing is part of image processing (Figure 4-2). There are three major forms of imaging geometry. The first one is central perspective. It is the simplest because the entire image frame is defined by the same set of geometric parameters. Images acquired with cameras take this form. The second imaging geometry is multi-central perspective where each pixel makes its own central perspective. Images obtained with Landsat TM, AVIRIS, and other rotating mirror scanners take this form. This is the most complicated because each pixel has to be corrected separately if there exists geometrical distortion. The third one is also multi-central perspective but each line of an image has a central perspective. Sensors such as the SPOT HRV, IRS, and EO-1 Hyperion acquire images in this form.

There are a number of factors that could introduce errors or distortions to image geometry. These include sensor lens distortion, platform instability (roll, yaw, and pitch), platform velocity, surface relief, earth rotation, and earth surface curvature (when a large area is imaged). Atmospheric turbulence affects the stability of airborne platform. Image view angle differences may cause different image distortion. Procedures developed to reduce the distortions caused by those factors is *geometric correction or rectification*.

Geometric processing has been traditionally dealt with in photogrammetry where techniques have been developed to address problems related to images acquired with cameras. Each photograph taken from a camera has a single perspective projection. Photogrammetric techniques developed to process central perspective images include geometric rectification, stereomodel building, 3D measurements, orthoimaging,

and multiple image mosaicing (Slama, 1980). The other two types of imaging geometry acquired by scanner type of sensors are rarely used in stereo modeling and 3D measurements. They are primarily used to produce plane maps. *Geometric processing* for the scanner type of sensors involves image geometric rectification, image to image registration, georeferencing and image mosaicing. Image to image *registration* is a procedure that transforms the coordinate system of one image to that of another so that multiple images taken at different times or even with different sensors can be compared spatially. Technically, image to image registration is similar to georeferencing. *Georeferencing* is a procedure that transforms the coordinate system of an image into a coordinate system of a specific map projection. While there are many different map projections designed for various mapping purposes, the universal transverse Mercator (UTM) projection (Snyder, 1982) is the dominant projection used for remote sensing of human settlements. With a known map projection, the geographic coordinates (i.e., longitude and latitude (λ, φ)) for any pixel in an image can be calculated in real time. Therefore, georeferencing makes it possible to analyze images acquired at different times, by different sensors and with different resolutions. Furthermore, georeferenced images can be analyzed with spatial data from other sources such as map data and field survey data. Image georeferencing is similar to geocoding of thematic data in a GIS. *Image mosaicking* is a procedure for joining multiple frames of images that have been georeferenced in the same coordinate system.

Presently, geometric processing at different levels can be done by image providers with extra expenses. In this section, we first introduce the procedures and principles for georeferencing of images acquired with non-camera sensors. It consists of three steps: selection of *ground control points* (GCPs) and determination of their coordinates in both image and map coordinate systems, determination of an optimal geometric transformation model between the image coordinate system and that of a map, and *image resampling* to re-construct the image in the map coordinate system. Image to image registration will then be briefly introduced with reference to georeferencing. Finally we will briefly explain the procedures involved in image mosaicking.

4.3.1 Ground Control Point Selection and Coordinate Determination

The 3D location of any real world phenomenon is referenced in a 2D map through a map projection. In fact, a remotely sensed image is a 2D representation (model) of the 3D world. However, the image coordinates are established arbitrarily and they must be transformed into a standard coordinate system. Ideally, the correspondence between each image pixel and its ground counterpart is clearly known in a simple mathematical form. However, because it is difficult to track the exact behavior of each distorting factor during an imaging process, an analytical relationship between an image pixel and its real world counterpart cannot be easily established. Therefore, the correspondence between the real world and its image is usually done through approximation using polynomial equations. To establish the polynomial equations a number of tie points, i.e., GCPs, that can be located both on the image and on the reference source must be selected and their coordinates in both systems must be determined. The reference source can be maps, images already georeferenced, orthophotos, or points in the field whose coordinates can be measured through surveying methods. Nowadays, field survey is primarily done using global positioning systems (GPS) technology. Positioning accuracies of various reference sources vary largely (Table 4-1). The topographic map accuracy standard has been modified in 1987 (Merchant, 1987) to be measured by a root-mean-square (rms) error between measured point coordinates and its true values. The rms error should be less than 0.00025 m X scale factor. Therefore, for the 1:24,000 scale map, this should be 6 m. Topographic maps produced before the adoption of the new accuracy standard were evaluated according to the following specifications: for scales less than 1:20,000, no more than 10% of well defined points should be greater than 1/50 inch in horizontal error while for scales greater than 1:20,000 no more than 10% of those test points should be greater than 1/30 inch in error (Thompson, 1988). Neither standard states what could be the worst case in map coordinate determination. We suggest that a factor of 2 is multiplied to the error limits as specified in the Thompson standard, due to other errors that may be introduced during map measurement and digitization (e.g., Gong et al, 1995). Orthophotos should have an accuracy as good as topographic maps. The rms errors for most image georeferencing tasks range between 0.5 and 1 pixel of the raw image. In the case of Landsat TM image whose orginal resolution is 30 m, the positional accuracy for georeferenced TM data would be 15-30 m. However, if the georefernced TM image is resampled to substantially smaller than its original pixel size, say 3 m, positioning accuracy may be improved a bit. This

is because detailed resampling allows the image operator to more precisely position GCPs on the image. The most accurate reference source is from GPS survey in the field.

GCPs taken from reference sources will be divided into two groups with one group used to determine the coefficients in the polynomial transformation, and other group will be used to check the accuracy of the transformations.

Table 4-1. Positioning Accuracies of Selected Various Reference Sources

Reference Source for Georeferencing	Positional Accuracy
Topographic Map* at 1:24,000 scale	90% or more test points within 12 m
Topographic Map* at 1:50,000 scale	90% or more test points within 25 m
Topographic Map** at 1:10,000 scale	Limiting root mean square error is 2.5 m
Topographic Map** at 1:20,000	Limiting root mean square error is 5 m
Georeferenced images	0.5-1 pixel size of the raw image
Raw GPS data	5 m in horizontal directions, 10 m in vertical direction
GPS with WAAS capability	Better than 1 m
Differential GPS at geodetic quality	Better than 0.1 m

* Old national mapping accuracy standard in the US (Thompson, 1988)
** New ASPRS mapping standard (Merchant, 1987)

GCPs should be selected with care. They should be permanent characteristic points that can be easily identified on the image and on the reference source. For human settlements, road intersections, large buildings, bridges, sport facilities, large parking lots of major commercial centers and harbor facilities. In the rural areas, there may not be a sufficient number of human structures to be used as GCPs. Stable turning points of stream and river systems, natural landscape components that have large contrasts with their surrounding areas can also be selected. Locations such as natural beaches where the water and land boundary changes with tide are examples of unsuitable locations for GCP selection. The distribution of GCPs should be related to the severity of image distortions. If no *a priori* knowledge is known about image distortion, GCPs should be evenly scattered around the area of interest. If the patterns of image distortions are known to some extent, more points should be selected in areas where distortion is severe. Depending on the accuracy requirement, different methods may be used to obtain the coordinates of GCPs. In a single georeferencing task, GCPs may be obtained from multiple sources.

In the following section, we will introduce the polynomial transformation method and explain the minimum number of GCPs required in building various orders of polynomial models.

4.3.2 The Geometric Transformation Model between Image and Reference Source

The purpose of georeferencing is to transform the image coordinates (u,v), to coordinates in a specific map projection (x,y) that can then be used or converted into other standard coordinate systems. Naturally one would like to know the relationship between the two coordinate systems. For a given image point (u_i, v_i), we can find (x_i, y_i) through a forward transformation, T (Figure 4-7).

$$\begin{pmatrix} x \\ y \end{pmatrix} = T \begin{pmatrix} u \\ v \end{pmatrix} \tag{4-10}$$

and its inverse process is

$$\begin{pmatrix} u \\ v \end{pmatrix} = T^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \tag{4-11}$$

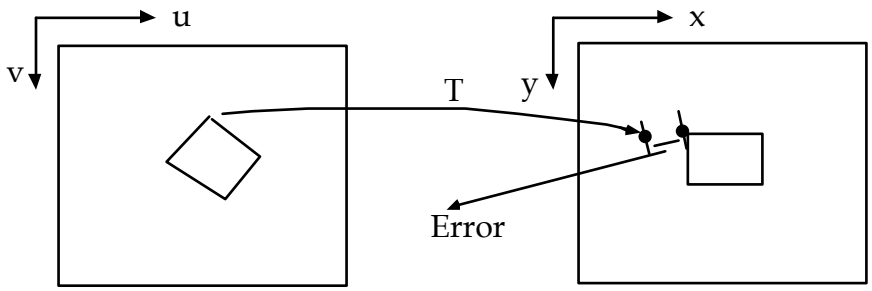


Figure 4-7. Georeferencing requires a transformation between the image space and the reference coordinate space. [B /W]

As mentioned earlier that not every step of the imaging process is exactly known, we adopt a simpler and widely-used alternative: polynomial approximation. It takes the following form:

$$u = \sum_{p=0}^{n} \sum_{q=0}^{n} a_{pq} x^p \cdot y^q \tag{4-12}$$

$$v = \sum_{p=0}^{n} \sum_{q=0}^{n} b_{pq} x^p \cdot y^q \tag{4-13}$$

where coefficients a's and b's are determined through least squares using GCPs. For example, we can use very low order polynomials such as the affine transformation

$$\begin{aligned} u &= a_0 + a_1 x + a_2 y \\ v &= b_0 + b_1 x + b_2 y \end{aligned} \tag{4-14}$$

or a bilinear transfomraiton

$$\begin{aligned} u &= a_0 + a_1 x + a_2 y + a_3 xy \\ v &= b_0 + b_1 x + b_2 y + b_3 xy \end{aligned} \tag{4-15}$$

The affine transformation is a first-order transformation that uses a linear model to describe the relationship between the two coordinate systems. It takes care of translation, scaling and rotation between the image and the reference coordinate systems. A minimum of three GCPs is needed to determine the coefficients in the affine transformation. The bilinear system is the simplest nonlinear system that requires a minimum of four GCPs to solve. If an image is divided into blocks whose four corners are composed of GCPs then the

bilinear transformation established by each of the four corner GCPs guarantees a continuity from block to block during a transformation. This is an important property for image mosaicking.

Most commercial software implements a "geometric correction" procedure that allows for establishment of polynomial transformations at different orders. A complete second order polynomial is in the following forms:

$$\begin{aligned} u &= a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2 \\ v &= b_0 + b_1 x + b_2 y + b_3 xy + b_4 x^2 + b_5 y^2 \end{aligned} \qquad (4\text{-}16)$$

This obviously requires a minimum of six GCPs to solve. In order to make the coefficients representative to the whole image, one must make sure that a sufficient number of GCPs are well distributed all over the image. Therefore, the number of GCPs is usually more than the number of unknowns in the polynomials making the problem an over determined one requiring a least squares approach to solve.

Suppose that we are solving for the first order polynomial with more than three, say *n*, GCPs,

$$\begin{aligned} &(u_1, v_1), \quad (x_1, y_1) \\ &(u_2, v_2), \quad (x_2, y_2) \\ &\quad . \\ &\quad . \\ &\quad . \\ &(u_n, v_n), \quad (x_n, y_n) \end{aligned}$$

by substituting the *n* GCPs coordinates into an affine transformation we obtain

$$\begin{aligned} u_1 &= a_0 + a_1 x_1 + a_2 y_1 \\ u_2 &= a_0 + a_1 x_2 + a_2 y_2 \\ &\quad . \\ &\quad . \\ &\quad . \\ u_n &= a_0 + a_1 x_n + a_2 y_n \end{aligned} \qquad \text{and} \qquad \begin{aligned} v_1 &= b_0 + b_1 x_1 + b_2 y_1 \\ v_2 &= b_0 + b_1 x_2 + b_2 y_2 \\ &\quad . \\ &\quad . \\ &\quad . \\ v_n &= b_0 + b_1 x_n + b_2 y_n \end{aligned}$$

The least squares solution in matrix form:

For **U**

$$\begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ . & & \\ . & & \\ . & & \\ 1 & x_n & y_n \end{bmatrix} \bullet \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

In symbolic form

$$U_{nx1} = M_{nx3} \quad A_{3x1}$$

(4-17)

Through some matrix arrangement, the solution for A is

$$A = [M^T \bullet M]^{-1} \bullet M^T \bullet U \tag{4-18}$$

where M^T is the transpose of M. Similarly, we can solve b_0, b_1, and b_2. This method can be applied to bilinear and higher order polynomial transformations. Once the coefficients are obtained, the n GCPs used in calculating the coefficients in the (x,y) system can be substituted into the polynomial equations to calculate a set of estimated coordinates (u', v') in the (u, v) system. The set of estimated coordinates can be compared with the original coordinates of the GCPs point by point to analyze the error characteristics. If |u' – u| or |v' – v| for a certain GCP is large, then the (x, y) and (u, v) coordinates for this particular GCP are re-examined to check if error has been introduced during GCP selection. If this is not the case, then this may imply a large amount of distortion for this point or an area around this point. More GCPs around this point should be taken. A new set of polynomial coefficients can then be calculated. The adjustment of the number and distribution of GCPs should be done with a consideration of the order of polynomials to be used.

Higher order polynomials can account for local variability provided that a sufficient number of GCPs are used. If the required number of GCPs is not available or their distribution is not even, then higher order polynomials will introduce a large amount of distortion in the georeferenced image at places where the number of GCPs is sparse. Usually, for images with a narrow view angle such as in the case of Landsat TM, SPOT HRV, IRS, and IKONOS, if their viewing direction is close to nadir, the level of geometric distortion on those images is not very series. Under such circumstances, 10-20 well distributed GCPs applied in an affine transformation would be sufficient. However, when images are acquired from an oblique direction with large view angle for mountainous areas, the geometric distortion in an image may vary considerably. A denser distribution of GCPs and higher order polynomials may be necessary. The actual number of GCPs applied to derive the polynomial transformation is determined empirically by evaluating the accuracy of transformed points.

We suggest that accuracy checking be done using a sufficient number of check points (GCPs not used in calculation of polynomial coefficients). The question is how many check points are sufficient? This should be the smallest number of check points that can give a relatively stable estimation of the average error. The coordinates of the m GCP check points in (x, y) system can be substituted in the final polynomial transformation equations to generate the estimated (u', v'). A root mean squared error (RMS) can be calculated through the following

$$rms = (1/m \bullet \sum_{i=1}^{m} (u_i^{'} - u_i)^2)^{0.5} \tag{4-19}$$

It can be used to assess the accuracy of the geometric transformation for u. This can also be done for v.

From the procedure presented above, we can see that the forward transformation allows us to find out the coordinates of a pixel in the (x, y) system. However, we introduced the determination of the inverse transformation. This is because we will use this inverse process in image resampling. The forward process can be determined in the same manner.

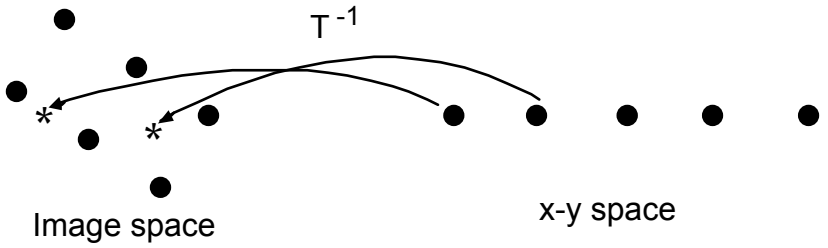
4.3.3 Image Resampling

With a forward geometric transform, each pixel (i, j) from image space (u, v) can be converted to a desirable reference space (x, y). A possible result for pixel located at (1, 1) in the image space transformed to a reference space at (485030.672 m, 4421625.341 m) in UTM projection. Because of the distortions

involved in the image, it is impossible to have a regularly gridded georefernced image with a forward transformation. The georeferenced image obtained with a forward transformation must be resampled to a regular grid. This is conveniently done using the inverse transformation as introduced in the previous section (Figure 4-8). This process starts from one pixel at a time in the x-y space, determines its coordinates in the u-v space. Obviously, the transformed (u, v) coordinates for the pixel in the (x, y) space may not fall right onto a pixel in the (u, v) space of the raw image. Image resampling can then be done in the u-v space by estimating a gray-level value from its surrounding gray-level values in the original image. This process is repeated for each pixel at a time until the gray-level values for all the pixels in the (x, y) space of the newly specified image area are estimated through interpolation methods.

Although there exist some other interpolation methods such as inverse distance weight based interpolation (Jensen, 1996), the use of sinc function, spline function, or Kriging methods, they are primarily used in spatial data whose locations are irregular. There are three most widely used ways of interpolation for image resampling:

- Nearest neighbour interpolation
- Bilinear (linear in each dimension)
- Cubic – a special case of spline



* in u-v space corresponds to the regular grid (pixel) location in the reference coordinate space.

Figure 4-8. The transformed image coordinates of pixels regularly arranged in the reference space. [B/W]

Nearest neighbor interpolation simply assigns the value of a pixel that is closest to the transformed coordinates. The bilinear interpolation uses the gray-level values of the four nearest pixels while the cubic interpolation requires 16 pixels. Figure 4-9 shows an example of the implementation of the bilinear interpolation. The cubic convolution can be done similarly. All three interpolation methods can be implemented through convolutions in 1D space. In general, this can be achieved using the following convolution operation:

$$Z(u) = Z_d(i) * w(i-u) = \sum_{|i-u| \le h} z_d(i) \bullet w(i-u) \qquad (4\text{-}20)$$

where Z(u) is the interpolated gray-level value at transformed coordinate u, $Z_d(i)$ is the known gray-level values for pixels in the original image at pixel location i, w(i- u) is the weight function and, h is the neiborhood size for a specific interpolation method. This is 0.5 for nearest neighbor, 1 for bilinear and 2 for cubic (Shlien, 1979). Clearly, the performance of a particular interpolation is controlled by the weight function.

For the linear case, the weight function is

$$w(\delta) = \begin{cases} 1 - |\delta| & for \quad |\delta| \le 1 \\ 0 & otherwise \end{cases}$$

where δ = i-u, For the cubic convolution, the weight function is

$$w(\delta) = \begin{cases} 1 - 2\delta^2 + |\delta|^3 & for \quad |\delta| \le 1 \\ 4 - 8|\delta| + 5\delta^2 - |\delta|^3 & for \quad 1 \le |\delta| \le 2 \\ 0 & otherwise \end{cases}$$

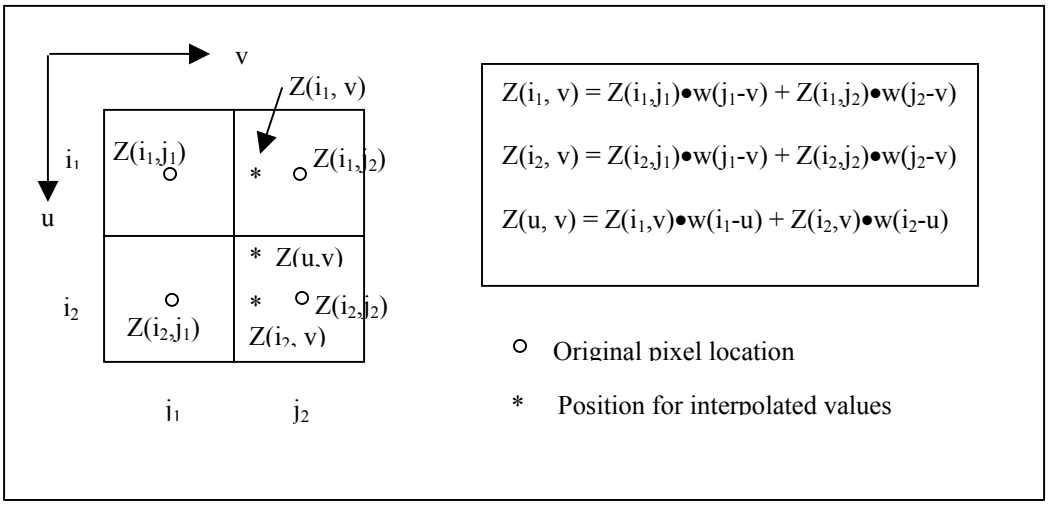


Figure 4-9. Procedures involved in the bilinear interpolation of one image pixel. [B/W]

The convolution is extended to 2D interpolation to obtain Z(u, v) by applying a sequence of 1D interpolations. For the case of applying a bilinear convolution, first Z(u, j_1) and Z(u, j_2) are calculated along each line of the image. Then, Z(u, v) is calculated from the two interpolated values in each line along a column (Figure 4-8). Therefore, the four points will be used in the bilinear convolution. Similarly, four interpolations along the row direction and one interpolation along the column direction are needed to complete a cubic convolution.

An example of georeferencing is shown in Figure 4-10, where a raw TM image was registered to a UTM projection. Such data can then be input to a GIS or to be overlaid with other layers of information for analysis.

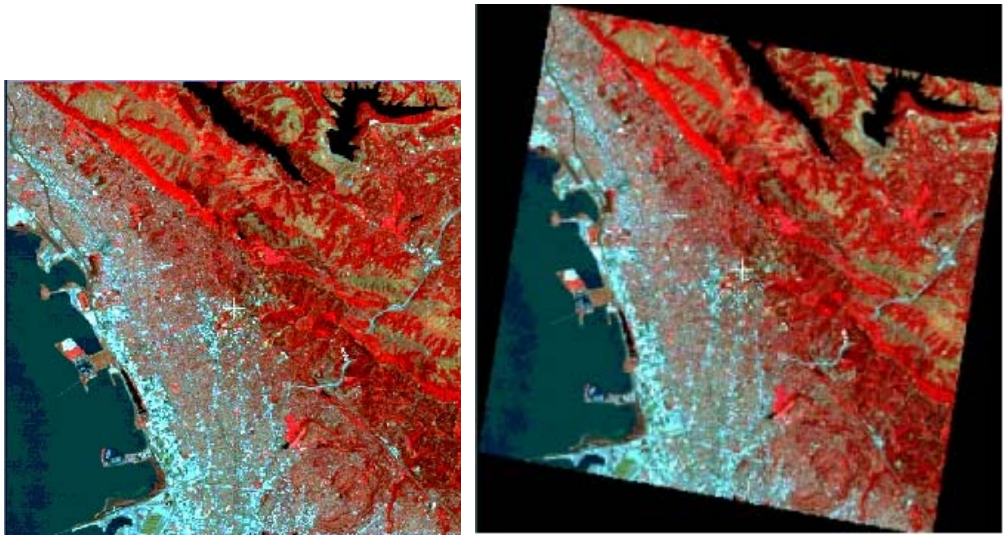


Figure 4-10. Georeferencing of a TM image of City of Berkeley and its surroundings. Left: the original image false color composite of Bands 4, 3, and 2 displayed in red, green and blue respectively. Right: the georeferenced version using cubic interpolation with 8 GCPs. The image in the right is already in UTM coordinates. Thus the north is along the vertical direction. The orientational change is caused by the oblique orbit of Landsat satellites. [W]

4.3.4 Image Mosaicking

In some applications, the entire study area requires multiple frames of images. Such images may be taken in the same day as in aerial photography, or on different days as satellite images acquired from polar orbits. Images need to be joined together with land features well matched from one image to another. This process is *image mosaickng* and the product is an image mosaic. Currently, image mosaicing is a function provided by a number of commercially available image analysis software systems as part of their georeferencing packages.

There are two pre-requisites for image mosaicing: accurate georeferencing of each of the images and sufficient overlap among images to be joined. The overlap area between two images is useful in reducing the radiometric differences between the two images and determination of a cut line between the two images. To avoid large radiometric differences from one image to the other on the final image mosaic, it is desirable that images to be mosaicked are taken on dates close to each other under similar illumination conditions. With images properly georeferenced, overlapped areas are examined to select a cut line between each pair of images to be mosaiced. Although georeferencing and radiometric correction may have been carefully done, linear features such as roads, streams, water boundaries, and other structurally prominent targets running from one image across the cut line to the other may still not matching well with each other (Figure 4-11a). Therefore, in cut line selection we should avoid structurally dense areas. In particular, we should avoid running cut lines through areas whose structural orientations are at a right angle with the cut lines (Figure 4-11b). Sometimes the gray-level values between the two images may have a contrast along the cut line (Figure 4-11c). Therefore, the portions of the images to be preserved on each side of the cut line need to be further processed. Using the overlapped portion of the images, we can reduce the radiometric difference through normalization (see section 4.2.3) and histogram adjustment (see section 4.4.1). Finally, *image blending* may be used to further remove local brightness differences and structural mismatches along the cut line (Figure 4-11d). Image blending is done through a process similar to image interpolation (see section 4.3.3) and image smoothing (see section 4.4.2). A buffer (usually 10-20 pixels wide) on each side of the cut line will be singled out in which the original graylevel of each pixel will be replaced by a new one resulting from a weighted averaging. The weight may be determined purely based on inverse distance in a local neighborhood (see Jensen, 1996, for inverse distance based weighting) or in combination with the distance of the pixel, whose gray level is to be altered, to the cut line. Figure 4-

11 shows examples of cut lines and the effect of image blending in removing radiometric differences between the two sides of the cut line.

(a) (b) (c) (d)

Figure 4-11. a. An improper selection of cut line running across a building. b. a proper selection of a cut line to avoid the leaning of the building and gray-level contrast of road surface. c. radiometric differences from one side vs. the other. d. removal of radiometric difference around the cut line through image blending. (Courtesy of H. Cao and Z. Zhang) [W]

4.3.5 Summary and Additional Remarks

At the beginning of this section we mentioned that there are two broad categories of geometric processing techniques, photogrammetric techniques developed for processing central perspective images acquired by cameras and, georeferencing techniques introduced in this section primarily developed for processing of satellite images. These include the selection of GCPs, establishing the geometric transformation model (polynomial models), image resmapling, and sometimes image mosaicing. Although we have not explicitly described the procedures for image to image registration, the principles are the same as image georeferencing with the reference source being another image.

There are a number of practical issues to be considered when applying georeferencing techniques for human settlement applications. Firstly, what is the acceptable georeferencing accuracy in human settlement remote sensing? Georeferencing is primarily done for the purpose of image to image comparison, or integrated use of image data with other spatial data such as maps. For urban change detection involving the comparison of only two Landsat MSS images taken in different years, Jensen (1981) suggested that the accuracy for geometric registration (the rms error in polynomial transformation measured by GCPs) be within half a pixel. Townshend et al (1992) found that a misregistration of one pixel could cause greater than 50% of false NDVI differences. They suggested that in order to maintain a less than 10% error in NDVI differences a 0.2 pixel RMSE is required for image registration over densely vegetated areas. Dai and Khorram (1998) examined the effects of misregistration to the accuracy of change detection from subpixel level of error to several pixels and obtained similar conclusions as Townshend et al (1992). In addition, they found that the near infrared channel of Landsat Thematic Mapper (TM) data was most sensitive to registration errors. Verbyla and Boles (2000) report the over-estimation of change due to misregistration with real examples. It should be noted that an RMSE of less than one pixel is hard to obtain. Most image registration procedures report an RMSE derived from coordinates of GCPs that have been used in building the image to image coordinate transformation model. Therefore, RMSE calculated from GCPs would usually underestimate registration error. An unbiased estimate of RMSE can only be calculated from independent check points separately selected. Gong et al (1992) proposed an adaptive filter technique to reduce the image differencing errors caused by misregistration that is greater than one pixel. From an economic point of view, a compromise is sought when the minimum cost meets with a maximum tolerable error in each specific application.

Secondly, what map projection should be used as the coordinate system for georeferencing? As mentioned earlier, the most commonly used map projection is UTM. Once image data is transformed into one map projection system, the coordinate systems can be easily transformed into other map projection systems.

Therefore, it is unnecessary to prefer one map projection to another during georeferencing. Selecting the appropriate map projection to work with should be based on the popularity of projection systems being used by other sources of data and the availability of most accurate maps. The reason for UTM being the most common projection system used is because it is used as a standard projection system in topographic maps that are usually used as the base reference for other mapping projects. Map projections are produced based on a specific model of the earth, called datum. The datum is calculated with geodetic methods. It defines the long and short axes of the earth and a surface model of the average sea level covering the entire earth. Different datums may be used for different parts of the world and geodecists are refining the parameters using advanced techniques. For example, Canada and USA used to adopt NAD27 (North America Datum 1927) as a standard in map making. This has been modified in the 1990s by NAD83 (North America Datum 1983). The same type of map projection derived from different datums will have different coordinates for the same point. Therefore, it is important to know what type of map projection based on which datum is used in image georeferencing.

Thirdly, what type of image resampling algorithm should be applied? This depends on the application purpose. In the past, the computation time involved in different image resampling methods is a big concern. Image resampling is now a trial task for even a low-end PC. Usually, cubic convolution results in the most accurate image resampling results. Therefore, unless there is a special reason to preserve the original gray-level values, we recommend the use of cubic convolution.

Fourthly, when should georeferencing be done? In image classification applications (see section 4.5.1), some image analysts choose to do geometric correction before image classification while others prefer the other way around. Other tasks in image georeferencing such as GCP selection and polynomial transformation model development can be done during the early stage of image processing, but not image resampling. Unless images are analyzed jointly with other sources of data, we recommend that image analysis tasks be done before image resampling. This is especially so in cases where preserving radiometric details in each band of image is essential for subsequent analysis. If maps and other spatial data are needed altogether, image resampling may have to be made before subsequent analysis. In urban areas, if maps such as thematic maps and topographic variables are needed, image resampling should be done first. No matter how little an image may be altered during georeferencing, certain structural and radiometric information that are critical to image analysts may be lost during the image resampling process. The order placement of image resampling should not be limited to image classification. After some assessment of texture information contents, Roy and Dikshit (1994) suggested that texture based analysis be made prior to image resampling.

Finally, for urban applications, permanent GCPs may be marked on buildings or selected locations with their coordinates surveyed by GPSs (Forster, 1980). The coordinates of those GCPs can be recorded in a database and small image chips for each GCP can also be stored in the database. When a new image is acquired, image chips can be used to match with the corresponding GCPs in the newly obtained image for automatic image georeferencing. This saves time and labor.

4.4 Image Enhancement

The goal of image enhancement is to make it easier to extract meaningful information from images. Different from the procedures discussed earlier about image radiometric processing whose goal is to reduce radiometric errors introduced during image acquisition, image enhancement produces new images with image radiometry altered, or images with new indices such as vegetation index and wetness index, or new feature images such as edge density and texture. Frequently used image enhancement procedures include histogram adjustment, filtering, image arithmetic, Fourier transforms and wavelet transforms of single images, transformation of multispectral images, and image fusion (see Figure 4-2). Among these techniques, some operate with data from the entire image and therefore they are considered as *global operators*, while others are based on local masks and are treated as *local operators*.

Histogram adjustment, spatial filtering, and image arithmetic are simple operations that have been implemented in most commercial image analysis software packages. *Histogram adjustment* includes linear

stretching and compression, nonlinear operations such as histogram equalization, and taking the square root of image gray-level values and then linearly rescaling them to the full gray-level range. The purpose is to improve image displays for visual image analysis. The search for valleys along a histogram is helpful in image segmentation of image that have similar gray-level values. It can also be used as an image compression tool for subsequent image processing (e.g., texture feature extraction). A rarely used function of nonlinear histogram adjustment is histogram matching. *Image filtering* can be done in two domains, the spatial domain and frequency domain. In the spatial domain we apply spatial filters such as averaging, edge-enhancement, and morphological filters. Images can be transformed into the spectral domain using Fourier transform for various filtering operations such as low pass and high pass filters. The purpose of image filtering is to suppress noises (using averaging or low pass filters) or sharpening structural features (using edge filters and high pass filters). Figure 4-12 summarizes three primary histogram adjustment schemes. The gray histogram in the background is the histogram of the original image. Clearly, the linear stretch does not alter the histogram shape other than simply stretching it to a wider range of gray levels. The square root adjustment first takes square roots of the image gray-level values and then linearly stretch them to the full gray-level range. As a result, ranges of the brighter gray levels are compressed while darker gray values are stretched. The histogram equalization attempts to adjust the histogram distribution so that an equal distribution is spread over the entire gray-level range (see the following section).

Figure 4-13 shows five typical kernel filters. Their effects ranges from smoothing, to high-frequency enhanced (edge-sharpening) to high-frequency preservation only.

Image arithmetic is usually a middle step that employs various arithmetic operators for deriving useful features from multispectral images. Among various image arithmetic operations, the most widely used methods are *image ratioing* and *differencing*. Taking the image ratio between two image bands can suppress illumination differences in the image which tends to normalize the data. Shadows of cloud and high-rise buildings can be reduced on ratio images. This is why images taken from hilly environment are normalized by dividing each band of image by the sum of images of all bands. Calculating a difference between two images of different bands is helpful in reducing the effect of atmospheric path radiance. Both *image differencing* and ratioing can be specially designed based on spectral characteristics of land cover types. For vegetation, because of the large reflectance contrast between the red and near infrared bands, it is often useful to enhance vegetation by generating a ratio or a difference image between near infrared and red bands.

Transformation of multispectral images includes generating linear combinations of multispectral bands using *principal component transformation* (PCT, also referred to as *principal component analysis*, PCA) or *Kauth-Thomas transform*. PCA is a procedure that redistributes the total variability in the multispectral space established from a number of image bands to a new set of mutually uncorrelated principal component (PC) images. The majority of the total variance in the original images is loaded in the first few PC images so that the remaining PC images can be treated as noise and discarded in later analysis. This effectively combines information from multiple image bands and reduces data dimensionality. The procedure is statistically based and therefore loses the original physical meaning of each image bands (e.g., Jensen, 1996; Richard and Jia, 1999). Consequently the interpreter is faced with a subjective judgment as to the physical meaning of the respective components.

In contrast, Kauth-Thomas transform (KT Transform, also known as *Tasselled Cap Transform*) produces a new set of physically meaningful and uncorrelated images from the multispectral image (Kauth and Thomas, 1976). This is done through the application of *Gram-Schmidt orthogonalization* (Jackson, 1983). Some KT transformed images contain information about landscape components, particularly for soil (brightness image) and vegetation (greenness image). Each new image is a linear combination of the original multispectral images. Because greenness and brightness images defined in KT transform are derived based on spectral characteristics of vegetation and soil objects, they may be more directly applicable in quantitative analysis of vegetation and soil conditions. However, the coefficients for KT Transform were only established for Landsat MSS and TM imagery. With TM imagery a third physically meaningful image, wetness image, has been added. The derived KT coefficients for MSS and TM imagery will be given in section 4.4.4 along with the introduction of Gram-Schmidt orthogonalization.

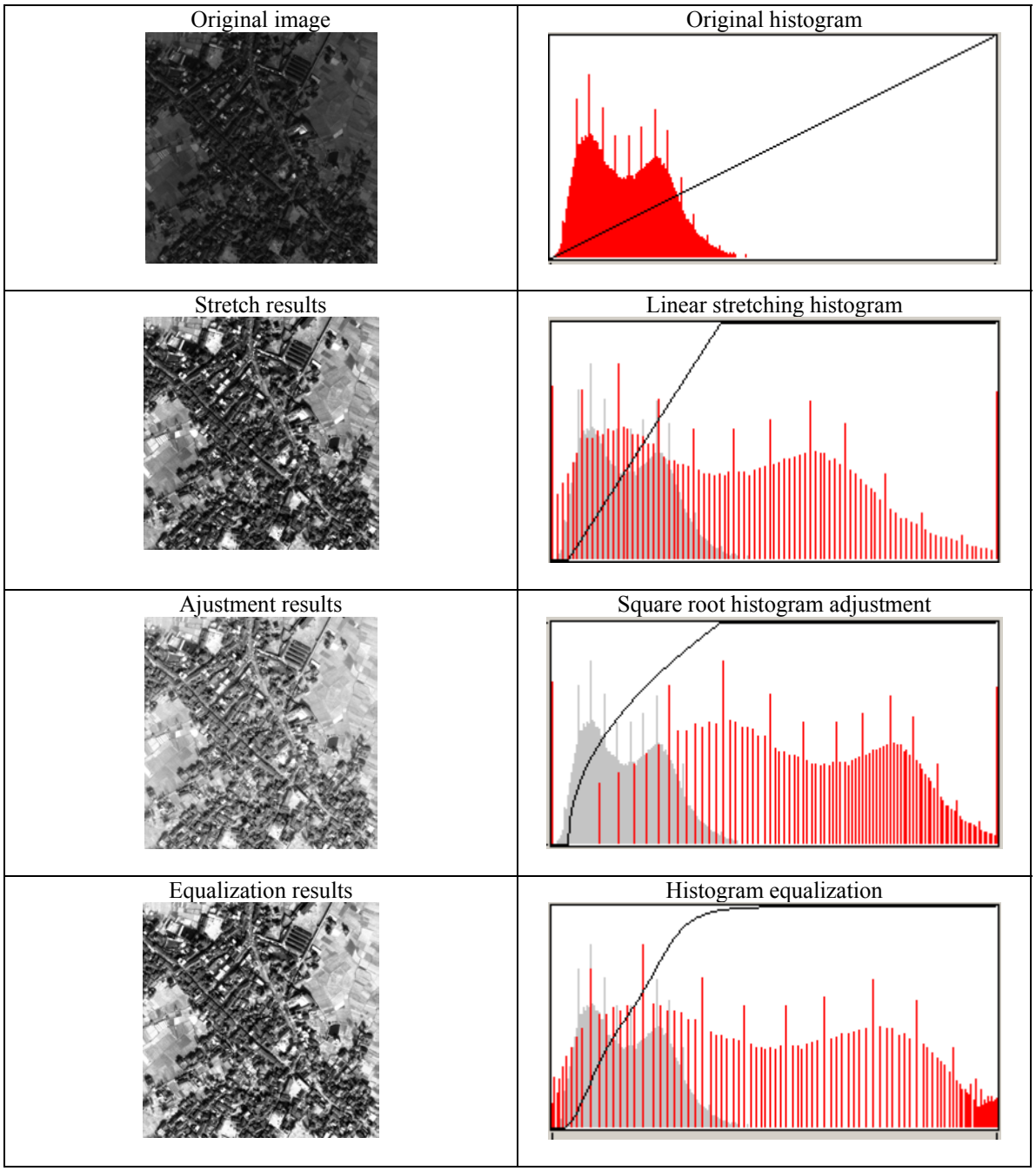


Figure 4-12. Examples of Three Typical Histogram Adjustments. Gray histograms in the background are the original histogram. Red histograms are the adjusted histogram. The curve in the histogram is the transfer function for the input brightness in the horizontal axis and the output brightness in the vertical axis. [W]

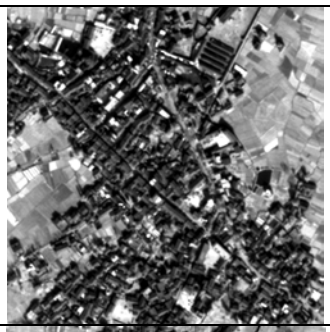
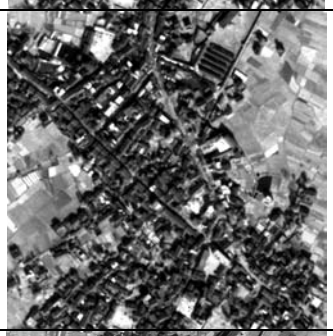
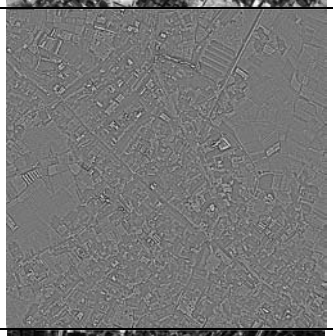
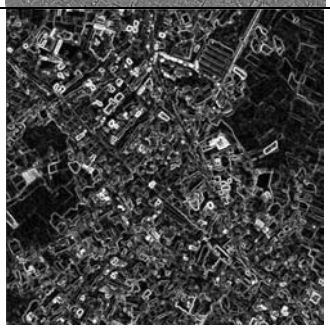
3 X 3 smooth filter Kernel $1/9\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$	
3 X 3 median filter	
3 X 3 Edge sharpening Kernel $\begin{bmatrix}-1 & -1 & -1\\ -1 & 9 & -1\\ -1 & -1 & -1\end{bmatrix}$	
3 X 3 8-neighbor Laplacian Kernel $\begin{bmatrix}1 & 1 & 1\\ 1 & -8 & 1\\ 1 & 1 & 1\end{bmatrix}$	
3 X 3 Sobel filter Kernels $H=\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix}$ V=$\begin{bmatrix}-1 & -2 & -1\\ 0 & 0 & 0\\ 1 & 2 & 1\end{bmatrix}$ $Gradient=\sqrt{H^2+V^2}$	

Figure 4-13. Some Typical Kernel-based Filters [B/W]

There are various types of *image fusion* such as fusion of images of different spatial resolution, fusion of images from different types, fusion of information extracted from images of various sources. Fusion of images can be achieved in a number of ways including the use of principal component analysis and wavelet transform among other approaches. The goal is to make integrated use of spatial and spectral data from multiple sources (Gong, 1994). For remote sensing of human settlements, deriving texture and other spatial features from images are important components of image enhancement. In this section, we will introduce the principles of histogram adjustment, morphological filtering, Gram-Schmidt orthogonalization, texture feature generation, and image fusion (see Figure 4-2).

4.4.1 Histogram Adjustment

The most straightforward histogram matching is *histogram equalization* that transforms a histogram of any shape to a histogram with the same frequency along the whole range of digital numbers (DN) (Figure 4-12).

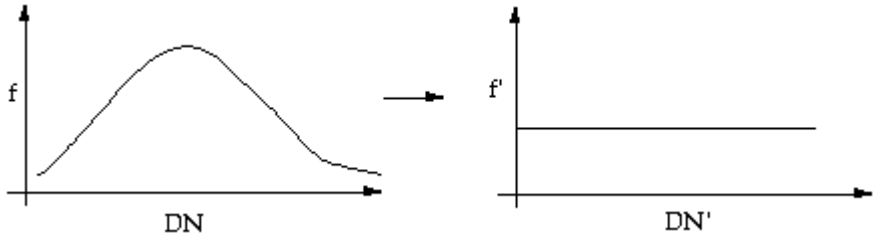


Figure 4-14. In the continuous case, equalizing the left histogram (f) to the one on the right (f') by reshaping f. [B/W]

The equalization process can be realized through histogram matching (Horn and Woodham, 1979). The cumulative histogram Fc1 of the original image is matched to the new cumulative histogram Fc2 (Figure 4-15).

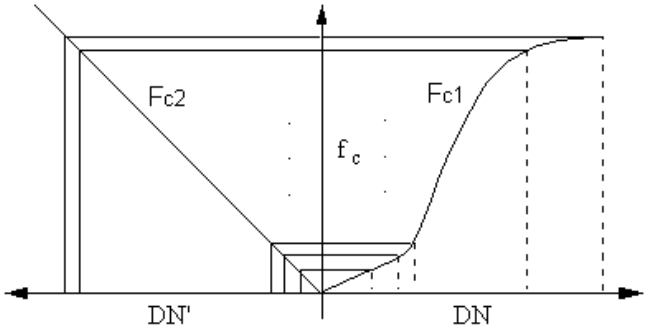


Figure 4-15 Matching the cumulative histogram on the right Fc1 to the Fc2 on the left. Any greylevels from the right (DN) whose cumulative histogram corresponds to a greylevel on the left (DN') will be assigned as DN'. [B/W]

The following example shows how an equalization made in discrete digital form through histogram matching (Table 4-2) with a 3-bit image (only 2^3, 8 gray levels). It starts with the generation of image histogram (first two columns in Table X). Then probability, Pi is calculated from frequency, f(vi) (third column). A cumulative histogram Fc can be calculated from frequencies. Similarly, the cumulative distribution function (CDF) can be derived from probabilities. Based on the cumulative distribution function we can convert the original gray levels into gray levels of the equalized image.

The assumption for histogram matching is that each detector (sensor) has the same probability of seeing the scene and, therefore, the gray-level distribution function should be the same. Thus if two detectors have different histograms (a discrete version of gray-level distribution function), they should be corrected to have the same histogram. Figure 4-16 shows two histograms to be matched.

Table 4-2 Histogram, Cumulative Histogram, Cumulative Distribution Function (CDF), and the Output in Histogram Equalization

Grey Level DN	Frequency f(vi)	Probability Pi	Cumulative histogram Fc	CDF	$(2^3 - 1)$ * CDF	Output DN' (Round off)
0	4	0.04	4	0.04	0.28	0
1	17	0.17	21	0.21	1.47	1
2	15	0.15	36	0.36	2.52	3
3	18	0.18	54	0.54	3.78	4
4	24	0.24	78	0.78	4.46	4
5	12	0.12	90	0.90	6.3	6
6	0	0	90	0.90	6.3	6
7	10	0.10	100	1.00	7	7
	100	1.00				

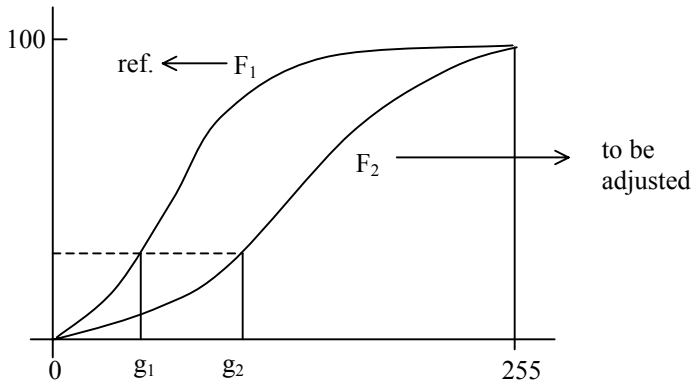


Figure 4-16. Matching the histogram F2 to the reference histogram F1. [B/W]

This process is done for each given gray-level, g2, to find its cumulative frequencies fc2(g2) in F2. Then in F1 find the gray-level value, g1, such that its cumulative frequency fc1(g1) = fc2(g2). Then assign g1 to g2 in the histogram to be adjusted. Histogram matching can be used to balance detector responses as a relative detector calibration technique. In image mosaicking, when the overlapped images are different in radiometry, it can be used to balance the differences. When two images acquired by the same sensor over the same location but at different times and atmospheric conditions, it can be used to normalize the two images.

4.4.2 Multispectral Operations and Simple Indices

The multispectral or vector nature of most remote sensing data makes it possible to use spectral transformations that generate new sets of images. A transformed image may make evident features not discernible in the original images or alternatively it may be possible to preserve the essential information content of the original images in a reduced number of dimensions. The preservation is useful for data display and presentation as we only have three primary colors available for display and printing. It is also useful for data transmission and storage.

Arithmetic operators such as addition, subtraction, multiplication, division, and powering or their combinations can be used to pixel graylevel values from two or more bands of image data to form new images. Among simple image arithmetic operators, multiplication is rarely used. Multi-band differencing and ratioing allows for reduction of atmospheric effects in different bands. Band ratioing allows for shade effect suppression. Among these, the ratio between the near infrared and red bands has been commonly regarded as an index for vegetation status. Some additional *vegetation indices* (VI) are listed below:

Normalized Difference Vegetation Index (NDVI)

$$NDVI = \frac{DN_{NIR} - DN_R}{DN_{NIR} + DN_R} \tag{4-21}$$

This is calculated from the raw data. For convenience of multidate data comparison, NDVI is calculated using reflectance data converted from digital numbers (through various radiometric calibration methods)

$$NDVI = \frac{SR_{NIR} - SR_R}{SR_{NIR} + SR_R} \tag{4-22}$$

To suppress the effect of different soil backgrounds on the NDVI, Huete (1988) recommended to use a soil-adjusted vegetation index:

$$SAVI = \frac{SR_{NIR} - SR_R}{SR_{NIR} + SR_R + 0.5} \bullet 1.5 \tag{4-23}$$

Transformed Vegetation Index

$$TVI = \{(DN_{NIR} - DN_R)/(DN_{NIR} + DN_R)\}^{1/2} \tag{4-24}$$

Perpendicular Vegetation Index

$$PVI = [(b_{NIR} - B_{NIR})^2 + (b_R - B_R)^2]^{1/2} \tag{4-25}$$

where $b_{NIR} = 0.851\, B_R + 0.355\, B_{NIR}$

$b_R = 0.355\, B_R + 0.148\, B_{NIR}$

Normalized difference water index (NDWI) has been proposed separately to enhance different matters. McFeeters (1996) proposed the use of green and near infrared channels to highlight surface water bodies while Gao (1996) proposed the use of spectral data in the 0.86 μm and 1.24 μm to quantify vegetation liquid water conditions. Therefore, the McFeeters' index takes the form of

$$NDWI_M = (Green - NIR)/(Green + NIR) \tag{4-26}$$

and the Gao's NDWI is

$$NDWI_G = (\rho_{0.86} - \rho_{1.24})/(\rho_{0.86} + \rho_{1.24}) \quad (4\text{-}27)$$

where ρ is the reflectance or radiance.

There are many indices proposed but they are mostly developed for quantitative vegetation remote sensing. They may be useful in the studies of environmental pollution in human settlement remote sensing.

4.4.3 Principal Component Analysis

The dimension of the multispectral space constructed by a remotely sensed image is the number of spectral bands (nb). For example, Landsat MSS image constructs a four dimensional multispectral space. For Landsat TM image, the multispectral space will have seven dimensions. Spectral bands of sensors are chosen based on their effectiveness on the monitoring of certain physical properties and differentiation of certain surface cover types. Because these spectral bands are always correlated, some of the spectral dimensions are redundant. It is desirable to reduce the redundancy by using a smaller number of dimensions than nb to represent the original image without losing much of the data variability. *Principal component analysis* (PCA) is a perfect tool for this. PCA rotates the original multispectral axes to a new set of principal component (PC) axes that are orthogonal to each other and are orientated along the eigenvector directions. The eigenvectors can be sorted according to their corresponding eigen values in a descending order so that the first PC axis is the eigenvector having the greatest eigen value and the last PC corresponds to the eigenvector with the lowest eigen value. Each eigen value is the amount of data variability recorded in the corresponding PC image after PCA is done.

For a given image with nb spectral bands. PCA is realized in four steps: construction of variance-covariance matrix, **V**, based on pixel values of image samples; find and sort the nb eigen values $(\lambda_1, \lambda_2, \ldots, \lambda_{nb})^T$; solve for the nb eigenvectors $G = (\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{nb})$; and transform (project) the original image into the nb eigenvector directions to produce nb PC images. The last step is also called *principal component transformation* (PCT).

Covariance Matrix and Correlation Matrix. The means are first calculated in vector form

$$\mathbf{m} = 1/N \sum_i \mathbf{x_i} \quad (4\text{-}28)$$

where N is the number of pixels sampled from the original image. **x** and **m** are vectors each with nb elements (e.g., $\mathbf{m} = (m_1, m_2, \ldots, m_{nb})^T$). For general purpose PCA, all pixels or a systematic sampling of the image is used. If the PCA is used to enhance certain phenomena such as urban areas, more samples from the urban areas or perhaps only pixels from the urban area should be selected.

For *variance-covariance matrix* **V**, since N is normally very large we can approximately write

$$\mathbf{V} = 1/N \sum_i \{(\mathbf{x_i} - \mathbf{m})(\mathbf{x_i} - \mathbf{m})^T\} \quad (4\text{-}29)$$

V is an nb x nb symmetric matrix. The *correlation matrix* **R** can be obtained from **V**.

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & \ldots\ldots & r_{1nb} \\ r_{21} & 1 & \ldots\ldots & \ldots\ldots \\ . & . & . & . \\ . & . & . & . \\ r_{nb1} & r_{nb2} & \ldots\ldots & 1 \end{pmatrix} \quad (4\text{-}30)$$

$$r_{ij} = \frac{V_{ij}}{\sqrt{V_{ii} - V_{jj}}} \quad , \quad r_{ij} = r_{ji}$$

The Principal Component Transformation. PCT has another name called Hoteling transformation.

Its purpose is to find a new coordinate system in which data can be represented without correlation. A vector system X is transformed into Y. A rotation matrix is used to complete the process. The rotation matrix is **G**,

$$\mathbf{Y} = \mathbf{G}^T\mathbf{x} \tag{4-31}$$

G can be found through deriving the eigenvalues and eigenvectors from the covariance matrix V_X. To find eigenvalues we need to solve

$$|\mathbf{V}_X - \lambda I| = 0 \tag{4-32}$$

where "I" is the identity matrix. λ is the eigenvalue vector $(\lambda_1, \lambda_2, ..., \lambda_{nb})^T$. For each non-zero eigenvalue, λ_i, we can find its corresponding eigenvector $\mathbf{g}_i = (g_{i1}, g_{i2}, ..., g_{inb})^T$. This can be obtained from

$$[\mathbf{V}_X - \lambda_i I] \bullet \mathbf{g}_i = 0 \tag{4-33}$$

Each eigenvector $\mathbf{g}_i$ is normalized to unit vectors. The rotation matrix **G** can then be determined by

$$\mathbf{G} = (\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_{nb})$$

Some of the properties of PCA are listed in the following:

(1) because the eigenvectors are unit vectors, rotation will not affect the total variance of the original data, i.e. the sum of λ_i is the same as the sum of variances in **V**

(2) transformation **G** simply redistributes the variances in **V**

(3) data variance along one axis in the new coordinate system is the same as its corresponding eigenvalue.

(4) λ_i devided by the total variance is called a *loading factor* of new axis (PC) i.

For multispectral space with nb dimensions, after PCT, some of the λ_i's are so low that their corresponding loading factors can be ignored in subsequent analysis. We can therefore make use of only those principal components with high loading factors in data analysis. By so doing, we will not lose much of the original data variability. This serves as a purpose for the reduction of data dimensionality.

Figure 4-17 shows an example of PCT applied to a TM image of Berkeley, California. The image was taken in July 1990 during the dry seasons of this area. The first image is a color composite of the original imagery. The remaining are the first five principal component (PC) images. Table 4-3, lists the variance-covariance matrix used to calculate the eigenvalues and their corresponding eigenvectors (arranged in each row). The percentage is the loading factor for the corresponding eigenvalue in the same row. It can be seen that the first PC image explains over 70% of the total variance of the original image. The first two PC images account for over 95% of the total variance. It can be seen that the first PC image agrees generally well with the color composite image. From the eigenvector values of PC 1 image, we can see that it is a weighted average of the original six bands of images with Bands 5, 4 and 7 having a stronger contribution.

We can regard this image as depicting surface brightness. On PC 1 image, the paved surface, dry soil and dry vegetation (grass and shrubs) appear brighter. The second PC image contrast paved surface with vegetation. The eigenvector values for PC 2 show all but Bands 4 and 5 have negative weights. Because the weights for Band 5 and 7 are low, this image can be regarded as the Band 4 image subtracted by the Bands 1, 2, and 3. The PC 3 image shows the difference between the two middle infrared bands and the visible and near infrared bands. Although PC 4 and PC 5 images look rather noisy and indeed they do not carry much of the total variance of the original images, the eigenvector values indicates that they are dominated by the difference between the band 7 and band 5, and between band 3 and band 1, respectively. An analysis of the eigenvector structure, we can see that the dominant variability in a multispectral image is the brightness, followed by the difference between near infrared and the other spectral bands. PCA decomposes the differences in an image and allows us to search for the causes of those differences. However, the axes as defined by the eigenvectors varies with image statistics. Images covering different types of landscapes may have considerably different eigen structures. Table 4-6 shows the different eigenstructure of two types of landscape in Figure 4-15a. The red boxed area is primarily an urban area that includes the Berkeley downtown, University of California campus and the surrounding residential areas. The green boxed area is a natural area covering part of Strawberry Canyon and Tilden Park with conifers, eucalyptus trees, and shrubs and grasslands distributed at various portions of the hills and valleys.

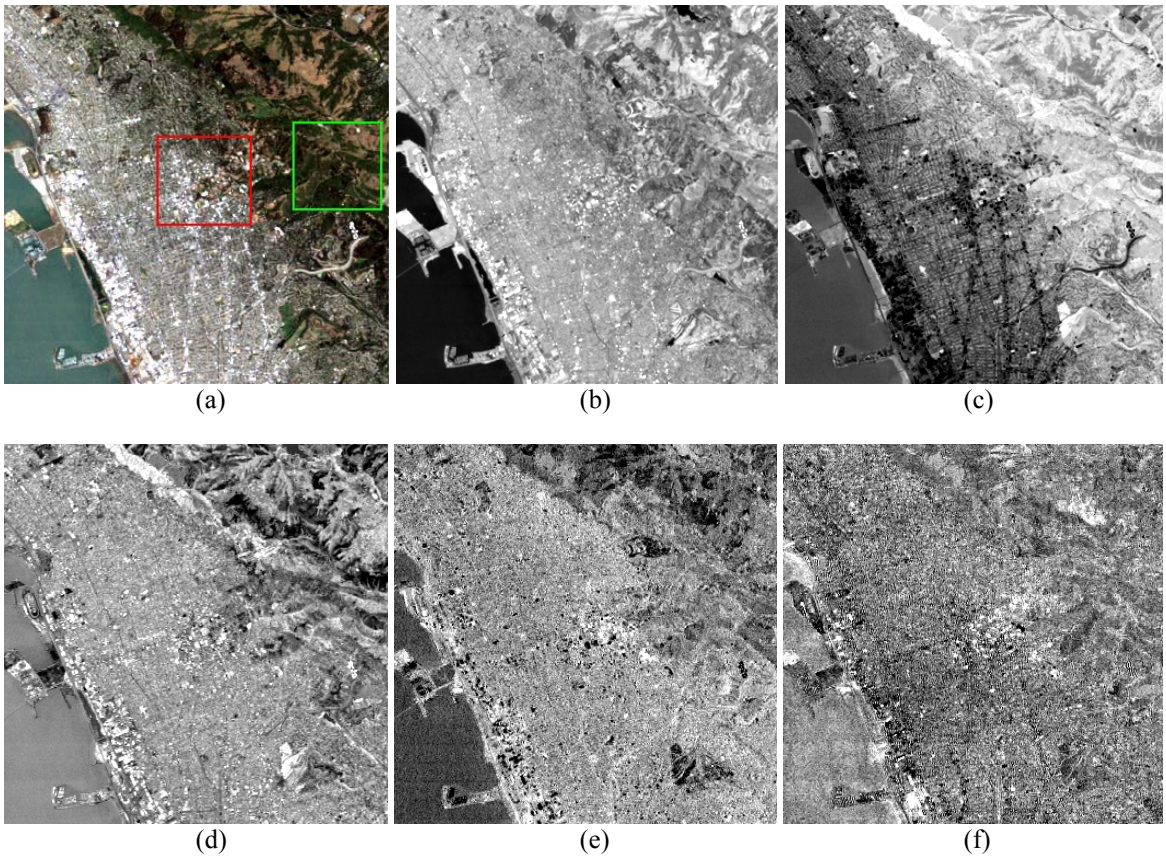


Figure 4-17. Principal component transformation of a TM image acquired over Berkeley, California, July 1990. a. color composite of TM Bands 1, 2, and 3 displayed in blue, green, and red. b-f, the principal component images corresponding to the first eigenvalues. [B/W]

From PC1 of the urban area, the dominant bands are Bands 1, 3, and 5. For PC2, this changes to Bands 4 and 5 with negative loadings. These are clearly different than the eigen structure of the entire image. For

the natural image, its PC1 is dominated by the two middle infrared bands while PC 2 is almost completely dominated by a positive loading of the near infrared band.

Table 4-3. Covariance Matrix, Eigenvalues and Their Corresponding Eigenvectors and Factor Loadings

		TM 1	TM 2	TM 3	TM 4	TM 5	TM 7	
TM 1		350.92						
TM 2		185.38	102.65					
TM 3		289.87	163.15	281.89				
TM 4		-47.62	2.71	61.94	681.26			
TM 5		176.83	135.28	333.84	709.11	1229.49		
TM 7		193.28	120.03	248.41	273.50	604.88	348.10	
	Eigenvalues	Eigenvectors						Loading factor
Eigen 1	2121.35	0.15	0.11	0.24	0.45	0.75	0.38	70.85%
Eigen 2	729.45	-0.62	-0.31	-0.46	0.51	0.10	-0.18	24.36%
Eigen 3	126.37	0.40	0.20	0.12	0.72	-0.45	-0.25	4.22%
Eigen 4	9.32	-0.09	-0.20	-0.05	0.13	-0.45	0.86	0.31%
Eigen 5	6.29	-0.61	0.16	0.76	0.06	-0.14	-0.07	0.21%
Eigen 6	1.54	0.24	-0.89	0.37	0.02	0.02	-0.15	0.05%

Table 4-4. Eigen Structure of Two Different Types of Landscapes

		TM 1	TM 2	TM 3	TM 4	TM 5	TM 7	
Urban	Eigenvalues	Eigenvectors						Loading factor
Eigen 1	1040.32	0.58	0.32	0.49	0.10	0.44	0.34	75.84%
Eigen 2	217.57	0.33	0.11	0.18	-0.78	-0.49	-0.06	15.86%
Eigen 3	90.90	-0.38	-0.19	-0.09	-0.57	0.52	0.46	6.63%
Eigen 4	13.68	-0.60	0.19	0.71	0.09	-0.27	0.14	1.00%
Eigen 5	7.51	0.12	-0.13	-0.23	0.22	-0.48	0.80	0.55%
Eigen 6	1.69	-0.18	0.89	-0.40	-0.06	0.00	0.07	0.12%
Natural								
Eigen 1	712.36	0.25	0.17	0.33	0.09	0.77	0.44	74.91%
Eigen 2	204.43	-0.15	-0.05	-0.22	0.94	0.14	-0.17	21.50%
Eigen 3	28.19	0.59	0.33	0.49	0.29	-0.46	-0.07	2.96%
Eigen 4	2.90	-0.31	-0.05	0.05	0.17	-0.41	0.84	0.30%
Eigen 5	2.24	0.64	-0.05	-0.72	-0.01	-0.03	0.27	0.24%
Eigen 6	0.78	-0.25	0.92	-0.28	-0.06	0.01	0.00	0.08%

Urban – the left box in Figure 4-17; Natural – the right box in Figure 4-17.

The above loading factors indicate that there are only 2-3 significant axes in the original six bands of TM imagery of the example area. The eigenvectors, i.e., the rotation axes, vary from one landscape to another. Interestingly, the first eigenvector all reflect surface brightness while the second one point to the direction

of vegetation difference. This has some relevance to the indices developed from KT transform which will be introduced shortly. The application of PCA in classification (keeping the maximum variance) serves the purpose of reducing data complexity and improving efficiency. A lot of efforts have been devoted to the use of PCA in change detection. When it is applied to multi-date images, change information may be preserved in those PC images with low λ_i's (e.g., Richard, 1984; Fung and LeDrew, 1987;88). Gong (1993) applied PCA to multi-date difference images and obtained change information in the first two principal components. Sometimes the PC images having low loading factors may contain important subtle information for a particular application. For examples, striping can be highlighted in some of the PCs with low loading factors revealing the problem of sensor detectors.

4.4.4 Gram-Schmidt Orthogonalization

Gram-Schmidt orthogonalization is a mathematical tool that operates rotational transformation in n-dimensional space, resulting in a series of orthogonal axes, each of which is a linear combination of the original axes. Therefore, it is a linear transform, and is sometimes called Gram-Schmidt transform (GST).

It requires m+1 pixel gray level vectors ($X_1, X_2, \ldots, X_{m+1}$) to derive m indices. The first index is defined by a difference vector between two of the gray-level vectors, $X_2 - X_1$. Normalizing this vector to unit length we have the coefficients for the first index, $V_1 = (X_2 - X_1)/| X_2 - X_1|$. The coefficients for the second index are obtained by finding the unit vector along the perpendicular line of V_1 passing through X_3. This can be obtained by first projecting vector $X_3 - X_1$ to the direction of V_1, $(X_3 - X_1)\bullet V_1V_1$, then subtracting the projected vector from vector $X_3 - X_1$ to obtain the perpendicular vector to V_1, $U_2 = (X_3 - X_1) - (X_3 - X_1)\bullet V_1V_1$. Normalizing U_2 we have the coefficients of the second index: $V_2 = U_2/|U_2|$. To derive the coefficients for the third index, we first find the line that is perpendicular to both V_1 and V_2 and passing through X_4. $U_3 = (X_4 - X_1) - (X_4 - X_1)\bullet V_1V_1 - (X_4 - X_1)\bullet V_2V_2$. Similarly we get V_3 by normalizing U_3, $V_3 = U_3/|U_3|$. This procedure guarantees that all indices derived in this manner is perpendicular to each other implying the newly generated indices having no correlation among them. This procedure can be generalized to produce m indices and for the mth index, $U_m = (X_{m+1} - X_1) - (X_{m+1} - X_1)\bullet V_1V_1 - (X_{m+1} - X_1)\bullet V_2V_2 - \ldots - (X_{m+1} - X_1)\bullet V_{m-1}V_{m-1}$ followed by a normalization of U_m.

Because gray-level vectors can be selected by the image analyst, this technique gives us flexibility in designing physically meaningful transforms for image enhancement. This is a distinct property of GST as compared to PCA which is totally based on the statistical properties of selected sample pixels.

A successful use of GST in remote sensing is the development of the Kauth-Thomas transform (*KT Transform*), also known as *tasseled cap transform*. Applying GST to a multispectral image, Kauth and Thomas (1976) developed a set of new indices within the multispectral space constructed by the four bands of Landsat multispectral scanner (MSS). The new indices are *brightness*, *greenness*, *yellowness* and nonesuch. The brightness was developed based on a soil line linked between wet and dry soils in the multispectral space while the greenness was developed along a line drown from a closed canopy vegetation point perpendicular to the soil line. These two indices can explain the majority of the image variation and are found to be related to surface cover conditions while the yellowness and nonesuch are sensitive to conditions of haze and water vapor in the atmosphere, respectively (Jackson, 1983). Since the last two indices were related to exogenous factors that are considered as noise in terrestrial remote sensing, only brightness and greenness have been widely used for Landsat MSS data. The linear coefficients for transforming the MSS data into the four indices are listed in Table 4-5. For Landsat TM data the number of useful indices has been increased to three – brightness, greenness and *wetness* (Table 4-6) (Crist and Cicone, 1984).

Table 4-5. KT Transform coefficients for Landsat MSS data (Kauth and Thomas, 1976)

	Brightness	Greenness	Yellowness	Nonesuch (noise)
MSS Green Band	0.433	-0.290	-0.829	0.223
MSS Red Band	0.632	-0.562	0.522	0.012
MSS NIR Band 1	0.586	0.600	-0.039	-0.543
MSS NIR Band 2	0.264	0.491	0.194	0.810

Table 4-6 shows the KT coefficients for TM imagery. Because GST enables us to create physically meaningful indices based on knowledge extracted from the images under investigation, it can be applied to images acquired by any type of sensor to produce either image specific or sensor specific transforms. For different applications, a different set of indices can be produced using GST.

Table 4-6. Two sets of KT transform coefficients for Landsat TM data (Crist and Cicone, 1984)

	Brightness	Greenness	Wetness
TM Band 1 – blue	0.304	-0.285	0.151
TM Band 2 – green	0.279	-0.244	0.197
TM Band 3 – red	0.474	-0.543	0.328
TM Band 4 – NIR	0.559	0.724	0.341
TM Band 5 – mid IR 1	0.508	0.084	-0.711
TM Band 7 – mid IR 2	0.186	-0.180	-0.457

* the two sets of coefficients are taken from the two references

With the vectors listed in Tables 4-5 and 4-6, we can arrange them into a matrix K. K can be applied to the gray level vector ($X=(x_1, x_2, \ldots, x_{nb})^T$) of each pixel in an image to produce a series of transformed images. For example, any set of the indices in Table 4-6 can be used to any TM image with the following equation:

$$\begin{pmatrix} B \\ G \\ W \end{pmatrix} = K^T X \tag{4-34}$$

where B G and W respectively stands for the brightness, greenness, and wetness. T denotes the transpose of matrix K.