Either form of enhancement can be performed on single-band (monochrome) images or on the individual components of multi-image composites. The resulting images may also be recorded or displayed in black and white or in color. Choosing the appropriate enhancement(s) for any particular application is an art and often a matter of personal preference.

Enhancement operations are normally applied to image data after the appropriate restoration procedures have been performed. Noise removal, in particular, is an important precursor to most enhancements. Without it, the image interpreter is left with the prospect of analyzing enhanced noise!

Below, we discuss the most commonly applied digital enhancement techniques. Three techniques can be categorized as *contrast manipulation*, *spatial feature manipulation*, or *multi-image manipulation*. Within these broad categories, we treat the following:

1. **Contrast manipulation.** Gray-level thresholding, level slicing, and contrast stretching.
2. **Spatial feature manipulation.** Spatial filtering, edge enhancement, and Fourier analysis.
3. **Multi-image manipulation.** Multispectral band ratioing and differencing, principal components, canonical components, vegetation components, intensity–hue–saturation (IHS) color space transformations, and decorrelation stretching.

7.4 CONTRAST MANIPULATION

Gray-Level Thresholding

Gray-level thresholding is used to *segment* an input image into two classes—one for those pixels having values below an analyst-defined gray level and one for those above this value. Below, we illustrate the use of thresholding to prepare a *binary mask* for an image. Such masks are used to segment an image into two classes so that additional processing can then be applied to each class independently.

Shown in Figure 7.11*a* is a TM1 image that displays a broad range of gray levels over both land and water. Let us assume that we wish to show the brightness variations in this band in the water areas only. Because many of the gray levels for land and water overlap in this band, it would be impossible to separate these two classes using a threshold set in this band. This is not the case in the TM4 band (Figure 7.11*b*). The histogram of DNs for the TM4 image (Figure 7.11*c*) shows that water strongly absorbs the incident energy in this near-infrared band (low DNs), while the land areas are highly reflective (high DNs). A threshold set at DN = 40 permits separation of these two classes in the TM4 data. This binary classification can then be applied to the

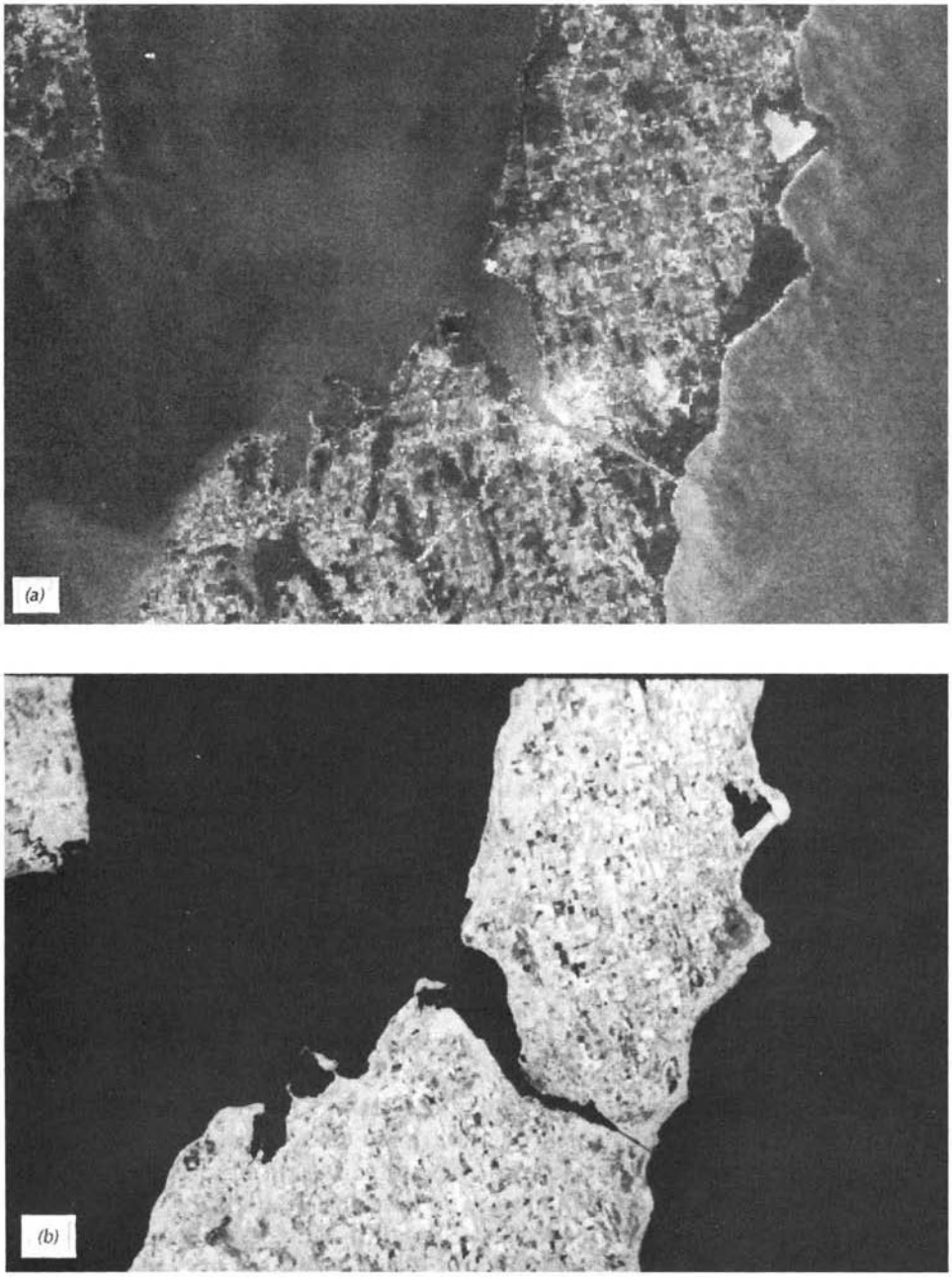
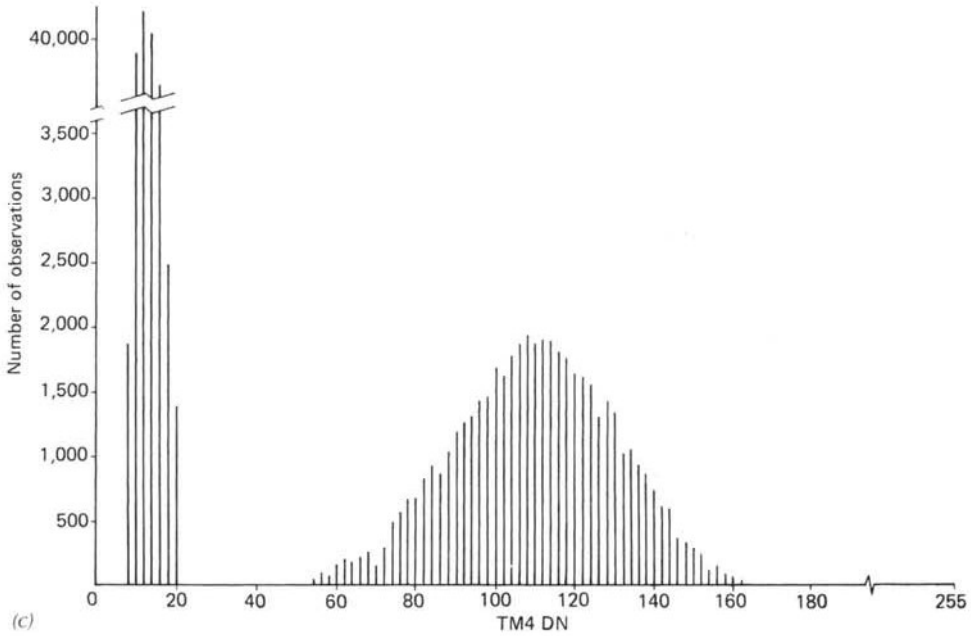


Figure 7.11 Gray-level thresholding for binary image segmentation: (*a*) original TM1 image containing continuous distribution of gray tones; (*b*) TM4 image; (*c*) TM4 histogram; (*d*) TM1 brightness variation in water areas only.



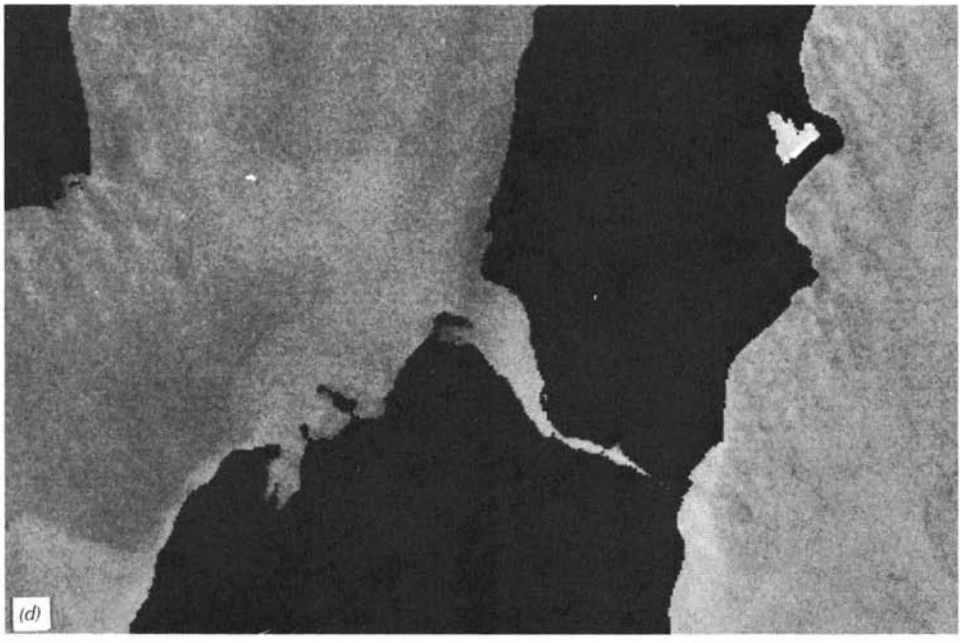


Figure 7.11 *(Continued)*

TM1 data to enable display of brightness variations in only the water areas. This is illustrated in Figure 7.11*d*. In this image, the TM1 land pixel values have all been set to 0 (black) based on their classification in the TM4 binary mask. The TM1 water pixel values have been preserved for display.

Level Slicing

Level slicing is an enhancement technique whereby the DNs distributed along the x axis of an image histogram are divided into a series of analyst-specified intervals or "slices." All of the DNs falling within a given interval in the input image are then displayed at a single DN in the output image. Consequently, if six different slices are established, the output image contains only six different gray levels. The result looks something like a contour map, except that the areas between boundaries are occupied by pixels displayed at the same DN. Each level can also be shown as a single color.

Figure 7.12 illustrates the application of level slicing to the "water" portion of the scene illustrated in Figure 7.11. Here, TM1 data have been level sliced into multiple levels in those areas previously determined to be water from the TM4 binary mask.

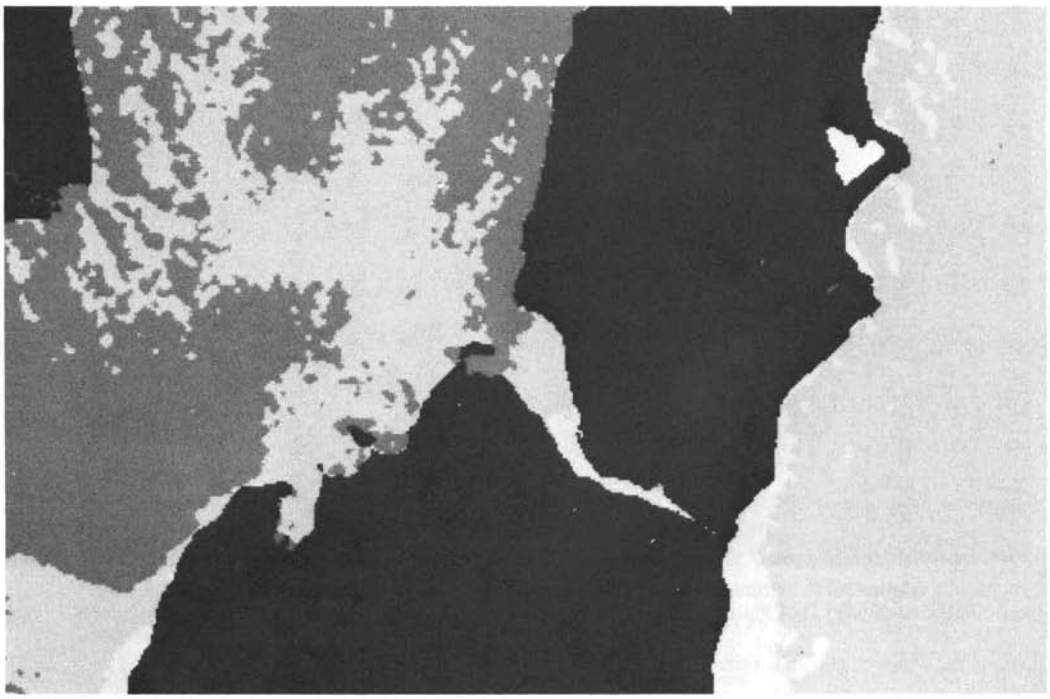


Figure 7.12 Level slicing operation applied to TM1 data in areas determined to be water in Figure 7.11.

Level slicing is used extensively in the display of thermal infrared images in order to show discrete temperature ranges coded by gray level or color. (See Plate 21 and Figure 6.18.)

Contrast Stretching

Image display and recording devices often operate over a range of 256 gray levels (the maximum number represented in 8-bit computer encoding). Sensor data in a single image rarely extend over this entire range. Hence, the intent of contrast stretching is to expand the narrow range of brightness values typically present in an input image over a wider range of gray values. The result is an output image that is designed to accentuate the contrast between features of interest to the image analyst.

To illustrate the contrast stretch process, consider a hypothetical sensing system whose image output levels can vary from 0 to 255. Figure 7.13*a* illustrates a histogram of brightness levels recorded in one spectral band over a scene. Assume that our hypothetical output device (e.g., computer monitor) is also capable of displaying 256 gray levels (0 to 255). Note that the histogram shows scene brightness values occurring only in the limited range of 60 to 158. If we were to use these image values directly in our display device (Figure 7.13*b*), we would be using only a small portion of the full range of possible display levels. Display levels 0 to 59 and 159 to 255 would not be utilized. Consequently, the tonal information in the scene would be compressed into a small range of display values, reducing the interpreter's ability to discriminate radiometric detail.

A more expressive display would result if we were to expand the range of image levels present in the scene (60 to 158) to fill the range of display values (0 to 255). In Figure 7.13*c*, the range of image values has been uniformly expanded to fill the total range of the output device. This uniform expansion is called a *linear stretch*. Subtle variations in input image data values would now be displayed in output tones that would be more readily distinguished by the interpreter. Light tonal areas would appear lighter and dark areas would appear darker.

In our example, the linear stretch would be applied to each pixel in the image using the algorithm

$$DN' = \left(\frac{DN - MIN}{MAX - MIN} \right) 255 \tag{7.6}$$

where

DN′ = digital number assigned to pixel in output image
DN = original digital number of pixel in input image
MIN = minimum value of input image, to be assigned a value of 0 in the output image (60 in our example)
MAX = maximum value of input image, to be assigned a value of 255 in the output image (158 in our example).

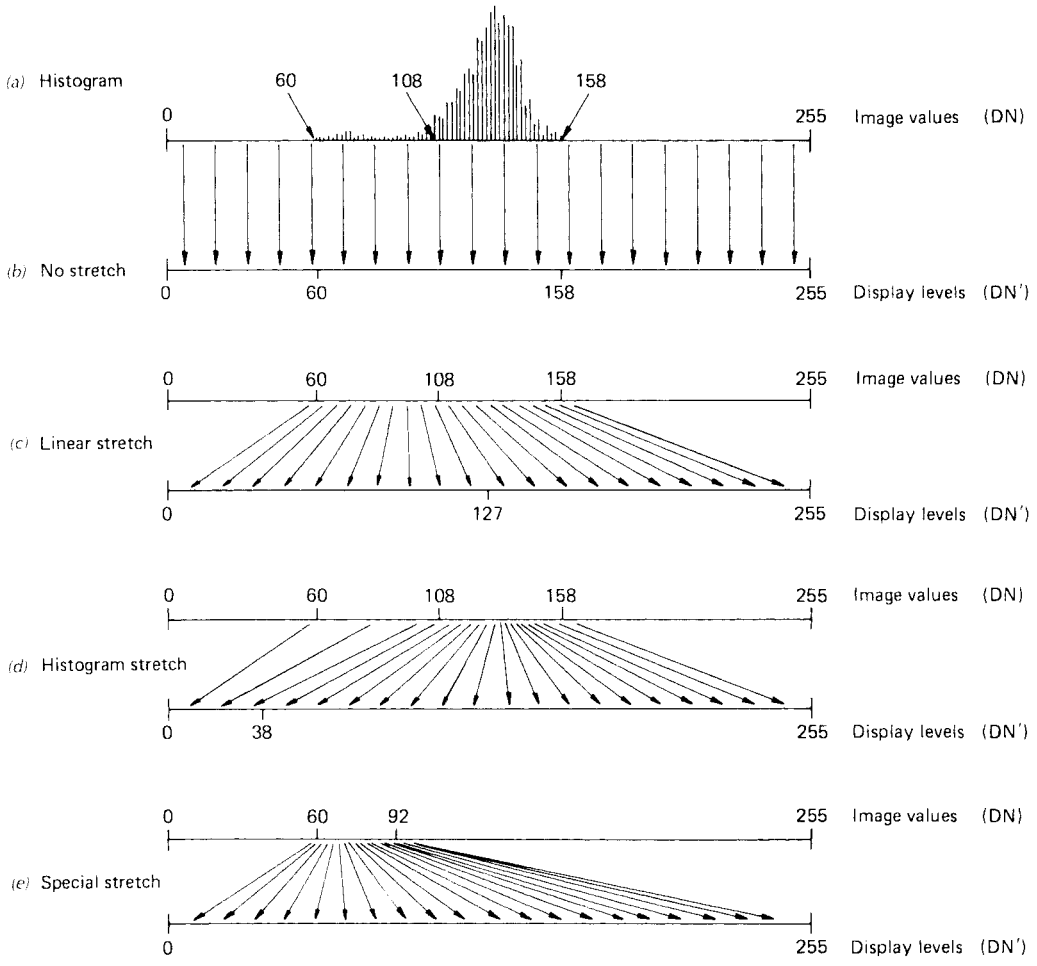


Figure 7.13 Principle of contrast stretch enhancement.

Figure 7.14 illustrates the above algorithm graphically. Note that the values for DN and DN' must be discrete whole integers. Since the same function is used for all pixels in the image, it is usually calculated for all possible values of DN *before* processing the image. The resulting values of DN' are then stored in a table (array). To process the image, no additional calculations are necessary. Each pixel's DN is simply used to index a location in the table to find the appropriate DN' to be displayed in the output image. This process is referred to as a *table lookup* procedure and the list of DN's associated with each DN is called a *lookup table* (*LUT*). The obvious advantage to the table lookup process is its computational efficiency. All possible values for DN' are computed only once (for a maximum of 256 times) and the indexing of a location in the table is then all that is required for each pixel in the image.

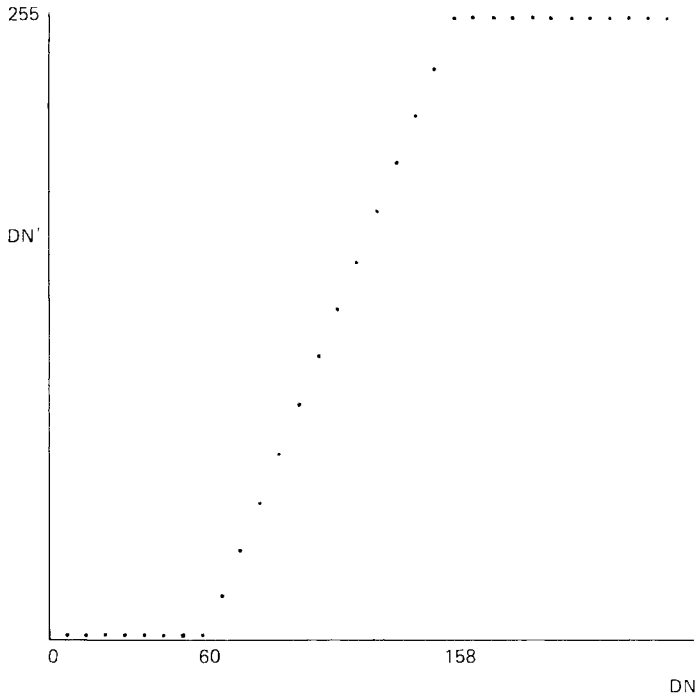


Figure 7.14 Linear stretch algorithm. Each point represents several discrete digital numbers.

One drawback of the linear stretch is that it assigns as many display levels to the rarely occurring image values as it does to the frequently occurring values. For example, as shown in Figure 7.13*c*, half of the dynamic range of the output device (0 to 127) would be reserved for the small number of pixels having image values in the range 60 to 108. The bulk of the image data (values 109 to 158) are confined to half the output display levels (128 to 255). Although better than the direct display in (*b*), the linear stretch would still not provide the most expressive display of the data.

To improve on the above situation, a *histogram-equalized stretch* can be applied. In this approach, image values are assigned to the display levels on the basis of their frequency of occurrence. As shown in Figure 7.13*d*, more display values (and hence more radiometric detail) are assigned to the frequently occurring portion of the histogram. The image value range of 109 to 158 is now stretched over a large portion of the display levels (39 to 255). A smaller portion (0 to 38) is reserved for the infrequently occurring image values of 60 to 108.

For special analyses, specific features may be analyzed in greater radiometric detail by assigning the display range exclusively to a particular range of image values. For example, if water features were represented by a narrow range of values in a scene, characteristics in the water features could be enhanced by stretching this small range to the full display range. As shown in Figure 7.13*e*,

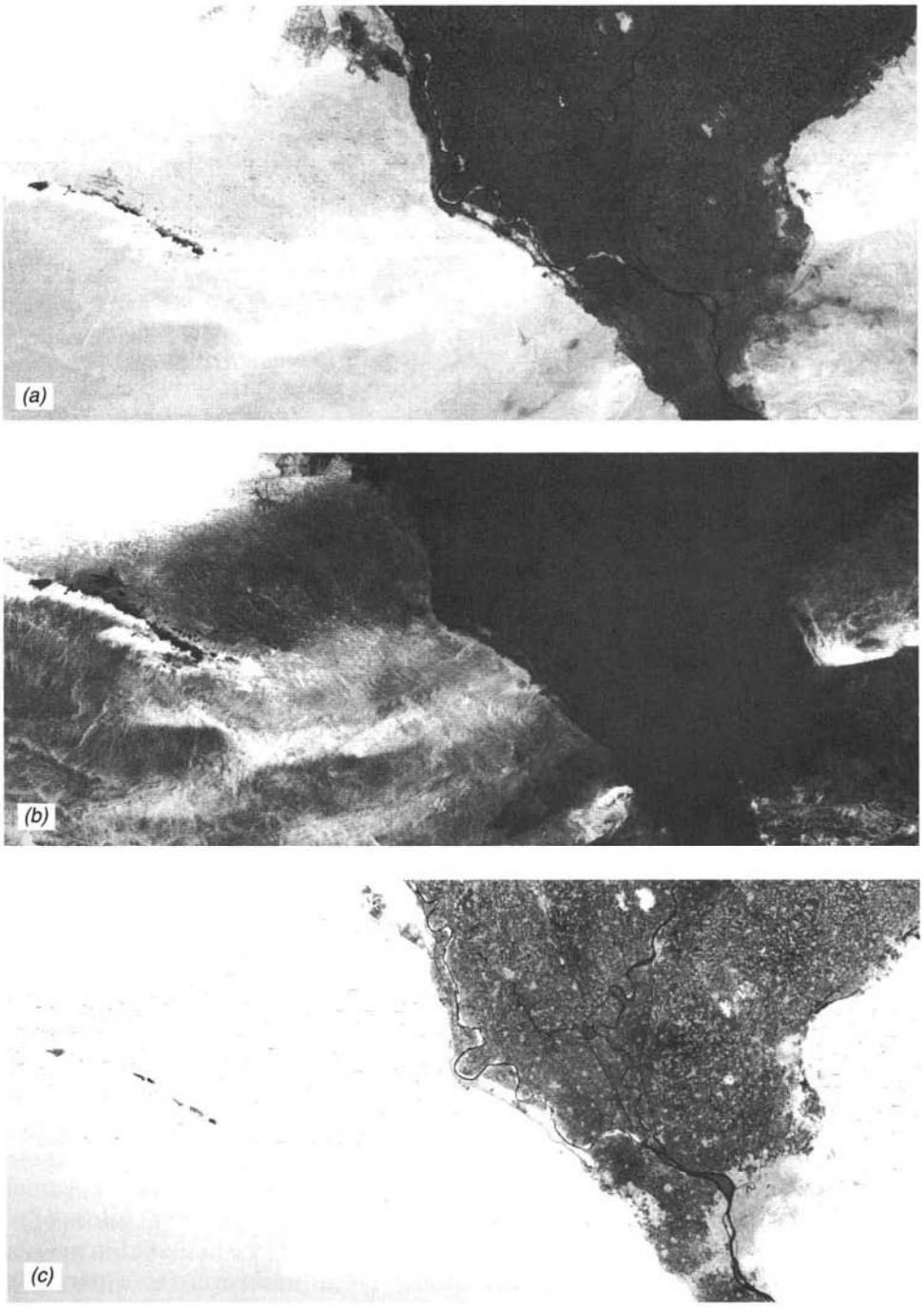


Figure 7.15 Effect of contrast stretching Landsat MSS data acquired over the Nile Delta; (*a*) original image; (*b*) stretch that enhances contrast in bright image areas; (*c*) stretch that enhances contrast in dark image areas. (Courtesy IBM Corp.)

the output range is devoted entirely to the small range of image values between 60 and 92. On the stretched display, minute tonal variations in the water range would be greatly exaggerated. The brighter land features, on the other hand, would be "washed out" by being displayed at a single, bright white level (255).

The visual effect of applying a contrast stretch algorithm is illustrated in Figure 7.15. An original Landsat MSS image covering the Nile Delta in Egypt is shown in (*a*). The city of Cairo lies close to the apex of the delta near the lower-right edge of the scene. Because of the wide range of image values present in this scene, the original image shows little radiometric detail. That is, features of similar brightness are virtually indistinguishable.

In Figure 7.15*b*, the brightness range of the desert area has been linearly stretched to fill the dynamic range of the output display. Patterns that were indistinguishable in the low contrast original are now readily apparent in this product. An interpreter wishing to analyze features in the desert region would be able to extract far more information from this display.

Because it reserves all display levels for the bright areas, the desert enhancement shows no radiometric detail in the darker irrigated delta region, which is displayed as black. If an interpreter were interested in analyzing a feature in this area, a different stretch could be applied, resulting in a display as shown in Figure 7.15*c*. Here, the display levels are devoted solely to the range of values present in the delta region. This rendering of the original image enhances brightness differences in the heavily populated and intensively cultivated delta, at the expense of all information in the bright desert area. Population centers stand out vividly in this display, and brightness differences between crop types are accentuated.

The contrast stretching examples we have illustrated represent only a small subset of the range of possible transformations that can be applied to image data. For example, nonlinear stretches such as sinusoidal transformations can be applied to image data to enhance subtle differences within "homogeneous" features such as forest stands or volcanic flows. Also, we have illustrated only monochromatic stretching procedures. Enhanced color images can be prepared by applying these procedures to separate bands of image data independently and then combining the results into a composite display.

7.5 SPATIAL FEATURE MANIPULATION

Spatial Filtering

In contrast to spectral filters, which serve to block or pass energy over various spectral ranges, spatial filters emphasize or deemphasize image data of various *spatial frequencies*. Spatial frequency refers to the "roughness" of the tonal variations occurring in an image. Image areas of high spatial frequency are tonally "rough." That is, the gray levels in these areas change abruptly over a relatively small number of pixels (e.g., across roads or field borders). "Smooth"

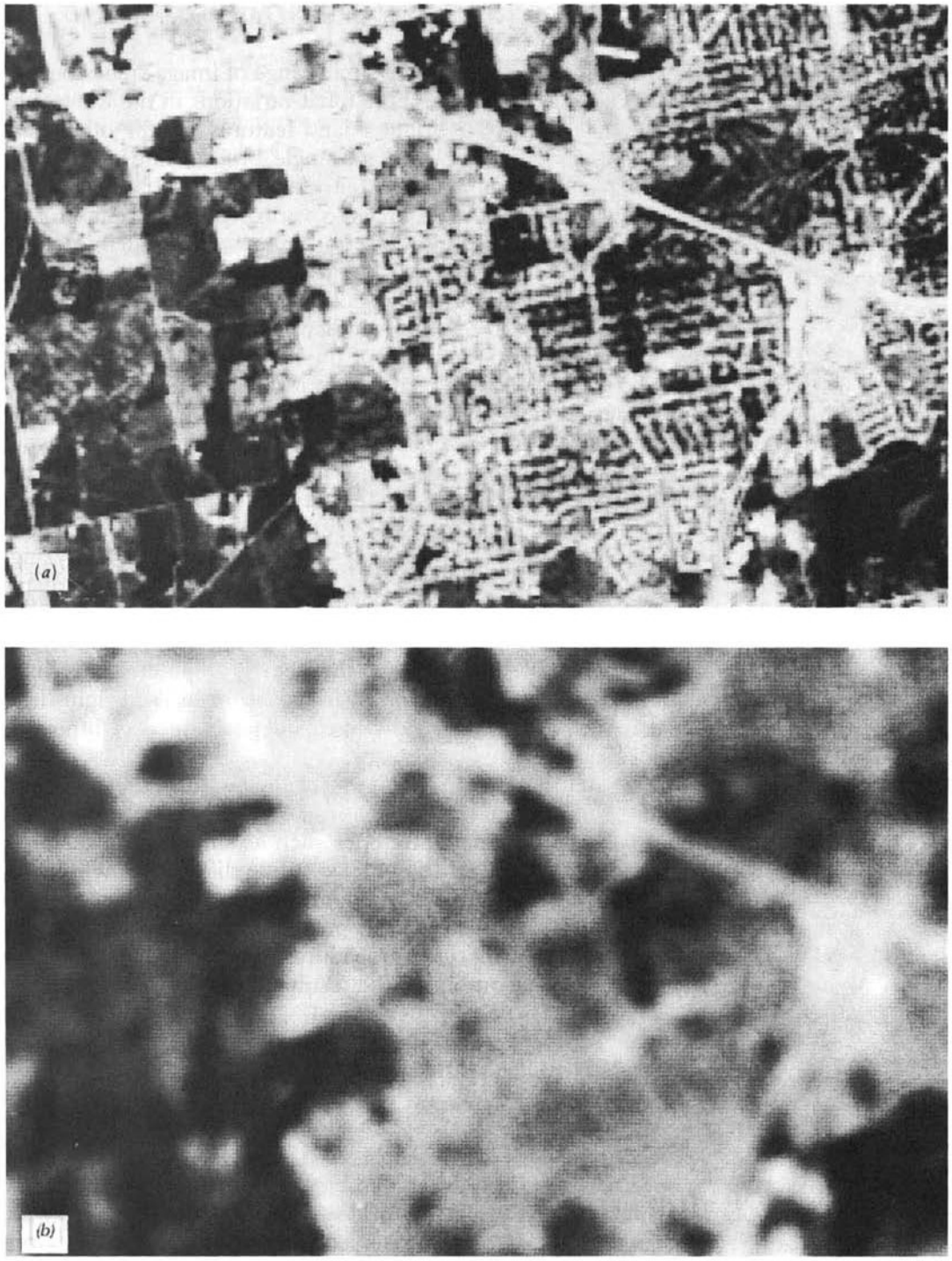


Figure 7.16 Effect of spatial filtering Landsat TM data: (*a*) original image; (*b*) low frequency component image; (*c*) high frequency component image.

Figure 7.16 *(Continued)*

image areas are those of low spatial frequency, where gray levels vary only gradually over a relatively large number of pixels (e.g., large agricultural fields or water bodies). *Low pass filters* are designed to emphasize low frequency features (large-area changes in brightness) and deemphasize the high frequency components of an image (local detail). *High pass filters* do just the reverse. They emphasize the detailed high frequency components of an image and deemphasize the more general low frequency information.

Spatial filtering is a "local" operation in that pixel values in an original image are modified on the basis of the gray levels of neighboring pixels. For example, a simple low pass filter may be implemented by passing a moving window throughout an original image and creating a second image whose DN at each pixel corresponds to the local average within the moving window at each of its positions in the original image. Assuming a 3×3-pixel window is used, the center pixel's DN in the new (filtered) image would be the average value of the 9 pixels in the original image contained in the window at that point. This process is very similar to that described previously under the topic of noise suppression. (In fact, low pass filters are very useful for reducing random noise.)

A simple high pass filter may be implemented by subtracting a low pass filtered image (pixel by pixel) from the original, unprocessed image. Figure 7.16

illustrates the visual effect of applying this process to an image. The original image is shown in Figure 7.16*a*. Figure 7.16*b* shows the low frequency component image and Figure 7.16*c* illustrates the high frequency component image. Note that the low frequency component image (*b*) reduces deviations from the local average, which smooths or blurs the detail in the original image, reduces the gray-level range, but emphasizes the large-area brightness regimes of the original image. The high frequency component image (*c*) enhances the spatial detail in the image at the expense of the large-area brightness information. Both images have been contrast stretched. (Such stretching is typically required because spatial filtering reduces the gray-level range present in an image.)

Convolution

Spatial filtering is but one special application of the generic image processing operation called *convolution*. Convolving an image involves the following procedures:

1. A moving window is established that contains an array of coefficients or weighting factors. Such arrays are referred to as *operators* or *kernels*, and they are normally an odd number of pixels in size (e.g., 3×3, 5×5, 7×7).
2. The kernel is moved throughout the original image, and the DN at the center of the kernel in a second (convoluted) output image is obtained by multiplying each coefficient in the kernel by the corresponding DN in the original image and adding all the resulting products. This operation is performed for each pixel in the original image.

Figure 7.17 illustrates a 3×3-pixel kernel with all of its coefficients equal to 1/9. Convolving an image with this kernel would result in simply averaging the values in the moving window. This is the procedure that was used to prepare the low frequency enhancement shown in Figure 7.16*b*. However, images emphasizing other spatial frequencies may be prepared by simply altering the kernel coefficients used to perform the convolution. Figure 7.18 shows three successively lower frequency enhancements (*b*, *c*, and *d*) that have been derived from the same original data set (*a*).

The influence convolution may have on an image depends directly upon the size of the kernel used and the values of the coefficients contained within the kernel. The range of kernel sizes and weighting schemes is limitless. For example, by selecting the appropriate coefficients, one can center-weight kernels, make them of uniform weight, or shape them in accordance with a particular statistical model (such as a Gaussian distribution). In short, convolution is a generic image processing operation that has numerous appli-

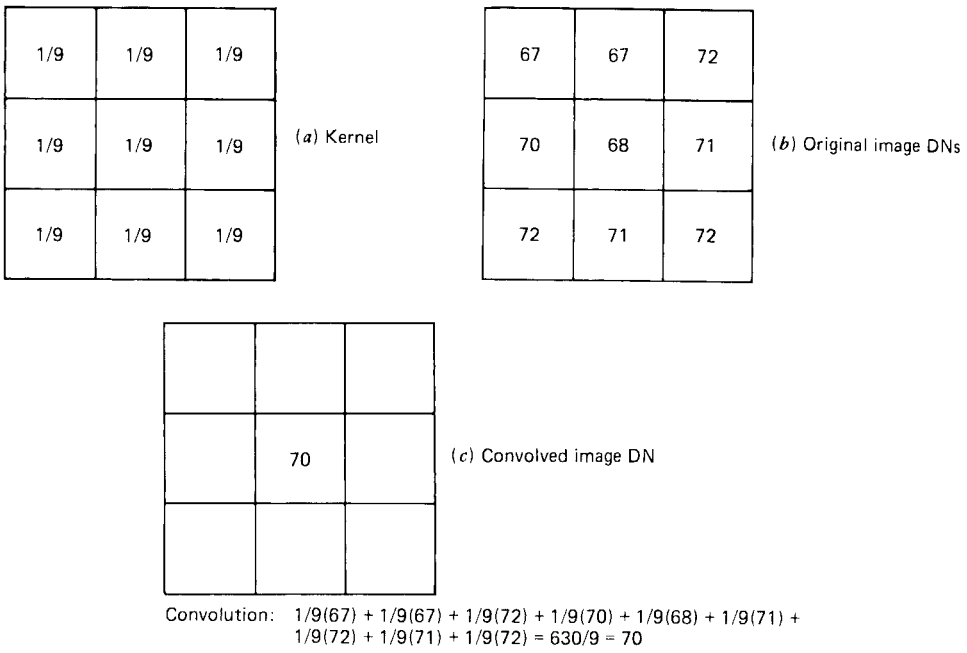


Figure 7.17 Concept of convolution. Shown is a 3 × 3-pixel kernel with all coefficients equal to $\frac{1}{9}$. The central pixel in the convolved image (in this case) contains the average of the DNs within the kernel.

cations in addition to spatial filtering. (Recall the use of "cubic convolution" as a resampling procedure.)

Edge Enhancement

We have seen that high frequency component images emphasize the spatial detail in digital images. That is, these images exaggerate local contrast and are superior to unenhanced original images for portraying linear features or edges in the image data. However, high frequency component images do not preserve the low frequency brightness information contained in original images. Edge-enhanced images attempt to preserve both local contrast and low frequency brightness information. They are produced by "adding back" all or a portion of the gray values in an original image to a high frequency component image of the same scene. Thus, edge enhancement is typically implemented in three steps:

1. A high frequency component image is produced containing the edge information. The kernel size used to produce this image is chosen based on the roughness of the image. "Rough" images suggest small

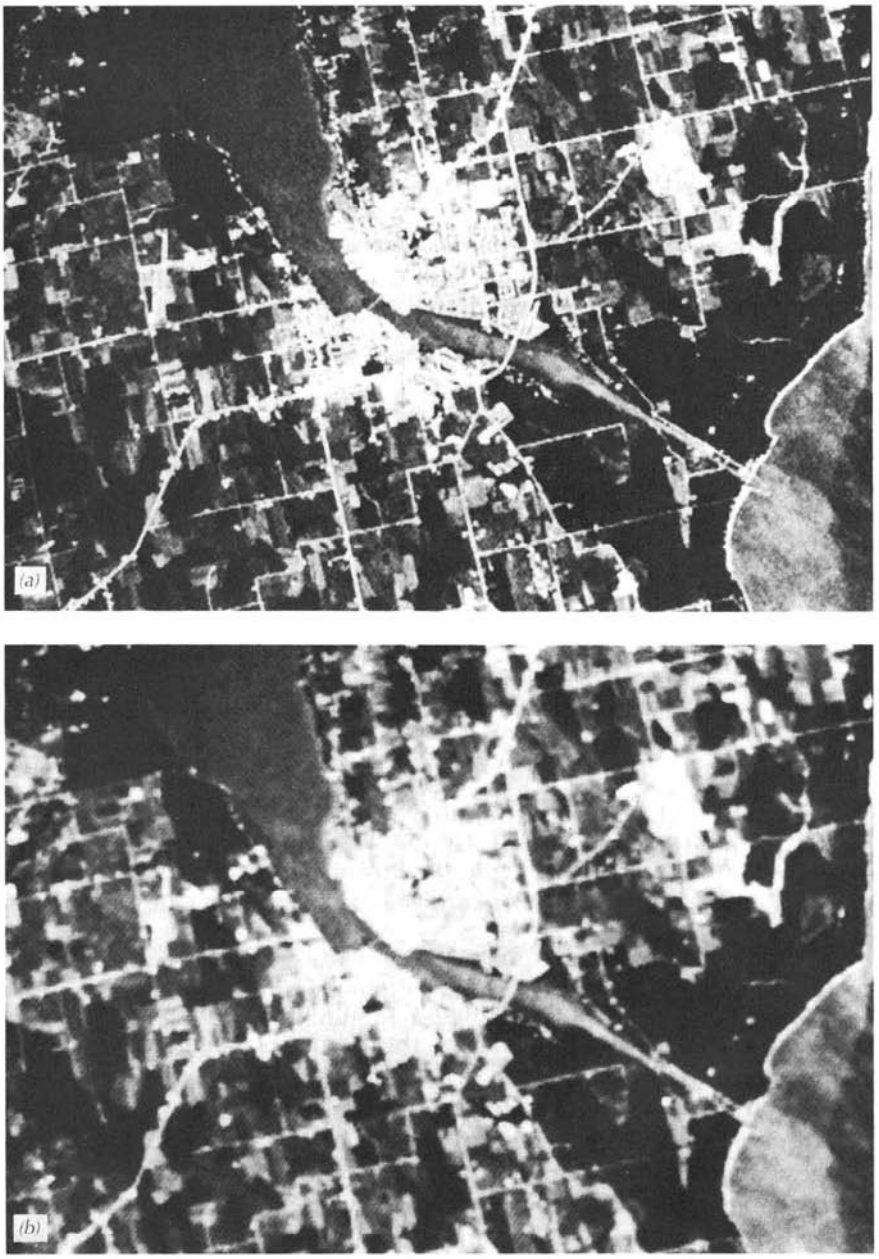


Figure 7.18 Frequency components of an image resulting from varying the kernel used for convolution: (*a*) original image; (*b–d*) successively lower frequency enhancements.

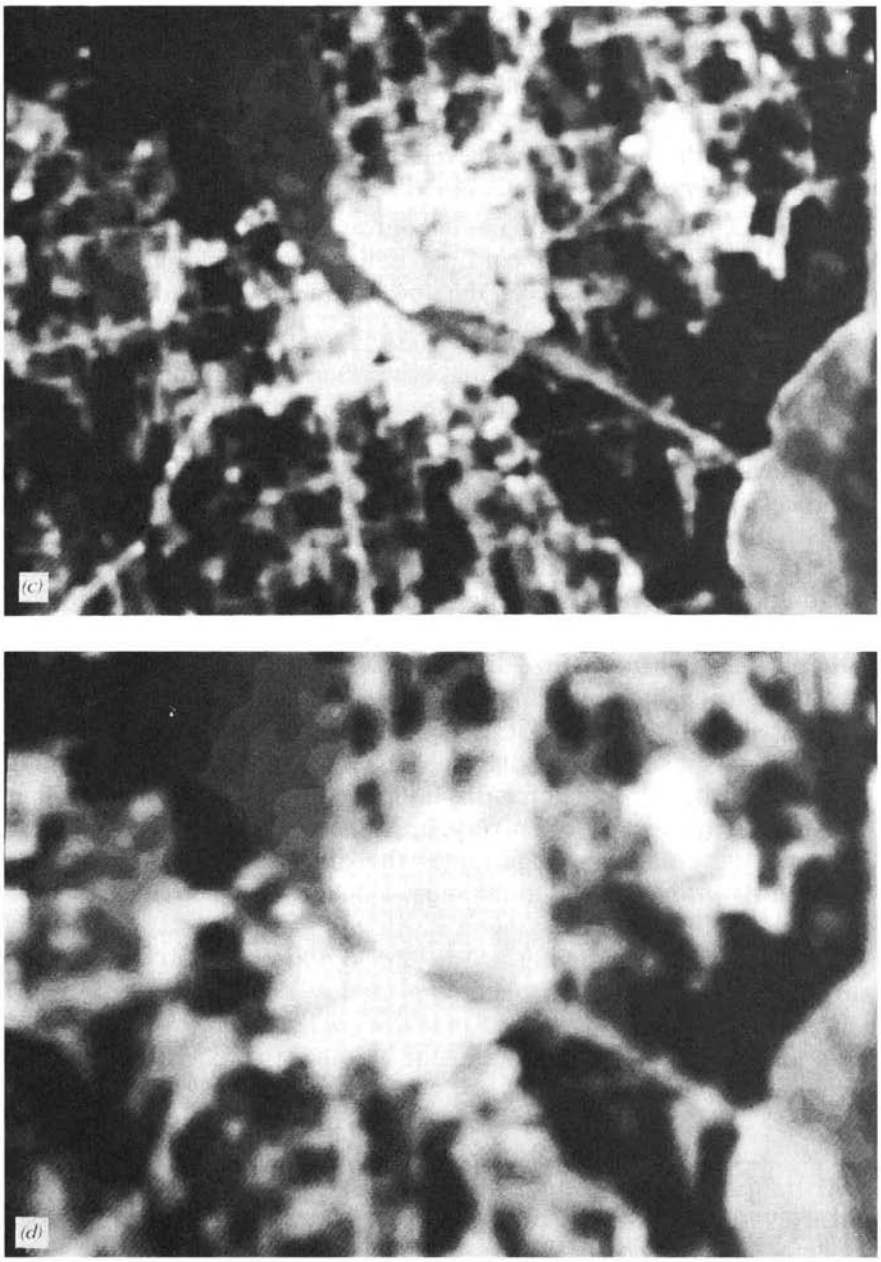


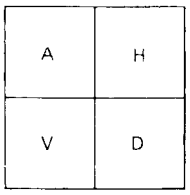
Figure 7.18 *(Continued)*

filter sizes (e.g., 3 × 3 pixels), whereas large sizes (e.g., 9 × 9 pixels) are used with "smooth" images.

2. All or a fraction of the gray level in each pixel of the original scene is added back to the high frequency component image. (The proportion of the original gray levels to be added back may be chosen by the image analyst.)
3. The composite image is contrast stretched. This results in an image containing local contrast enhancement of high frequency features that also preserves the low frequency brightness information contained in the scene.

Directional first differencing is another enhancement technique aimed at emphasizing edges in image data. It is a procedure that systematically compares each pixel in an image to one of its immediately adjacent neighbors and displays the difference in terms of the gray levels of an output image. This process is mathematically akin to determining the first derivative of gray levels with respect to a given direction. The direction used can be horizontal, vertical, or diagonal. In Figure 7.19, a horizontal first difference at pixel A would result from subtracting the DN in pixel H from that in pixel A. A vertical first difference would result from subtracting the DN at pixel V from that in pixel A; a diagonal first difference would result from subtracting the DN at pixel D from that in pixel A.

It should be noted that first differences can be either positive or negative, so a constant such as the display value median (127 for 8-bit data) is normally added to the difference for display purposes. Furthermore, because pixel-to-pixel differences are often very small, the data in the enhanced image often span a very narrow range about the display value median and a contrast stretch must be applied to the output image.



Horizontal first difference = $DN_A - DN_H$

Vertical first difference = $DN_A - DN_V$

Diagonal first difference = $DN_A - DN_D$

Figure 7.19 Primary pixel (A) and reference pixels (H, V, and D) used in horizontal, vertical, and diagonal first differencing, respectively.

First-difference images emphasize those edges normal to the direction of differencing and deemphasize those parallel to the direction of differencing. For example, in a horizontal first-difference image, vertical edges will result in large pixel-to pixel changes in gray level. On the other hand, the vertical first differences for these same edges would be relatively small (perhaps zero). This effect is illustrated in Figure 7.20 where vertical features in the original image (*a*) are emphasized in the horizontal first-difference image (*b*). Horizontal features in the original image are highlighted in the vertical first-difference image (*c*). Features emphasized by the diagonal first difference are shown in (*d*).

Figure 7.21 illustrates yet another form of edge enhancement involving diagonal first differencing. This image was produced by adding the absolute value of the upper-left-to-lower-right diagonal first difference to that of the upper-right-to-lower-left diagonal. This enhancement tends to highlight all edges in the scene.

Fourier Analysis

The spatial feature manipulations we have discussed thus far are implemented in the *spatial domain*—the (x, y) coordinate space of images. An alternative coordinate space that can be used for image analysis is the *frequency domain*. In this approach, an image is separated into its various spatial frequency components through application of a mathematical operation known as the *Fourier transform*. A quantitative description of how Fourier transforms are computed is beyond the scope of this discussion. Conceptually, this operation amounts to fitting a continuous function through the discrete DN values if they were plotted along each row and column in an image. The "peaks and valleys" along any given row or column can be described mathematically by a combination of sine and cosine waves with various amplitudes, frequencies, and phases. A Fourier transform results from the calculation of the amplitude and phase for each possible spatial frequency in an image.

After an image is separated into its component spatial frequencies, it is possible to display these values in a two-dimensional scatter plot known as a *Fourier spectrum*. Figure 7.22 illustrates a digital image in (*a*) and its Fourier spectrum in (*b*). The lower frequencies in the scene are plotted at the center of the spectrum and progressively higher frequencies are plotted outward. Features trending horizontally in the original image result in vertical components in the Fourier spectrum; features aligned vertically in the original image result in horizontal components in the Fourier spectrum.

If the Fourier spectrum of an image is known, it is possible to regenerate the original image through the application of an *inverse Fourier transform*. This operation is simply the mathematical reversal of the Fourier transform. Hence, the Fourier spectrum of an image can be used to assist in a number of

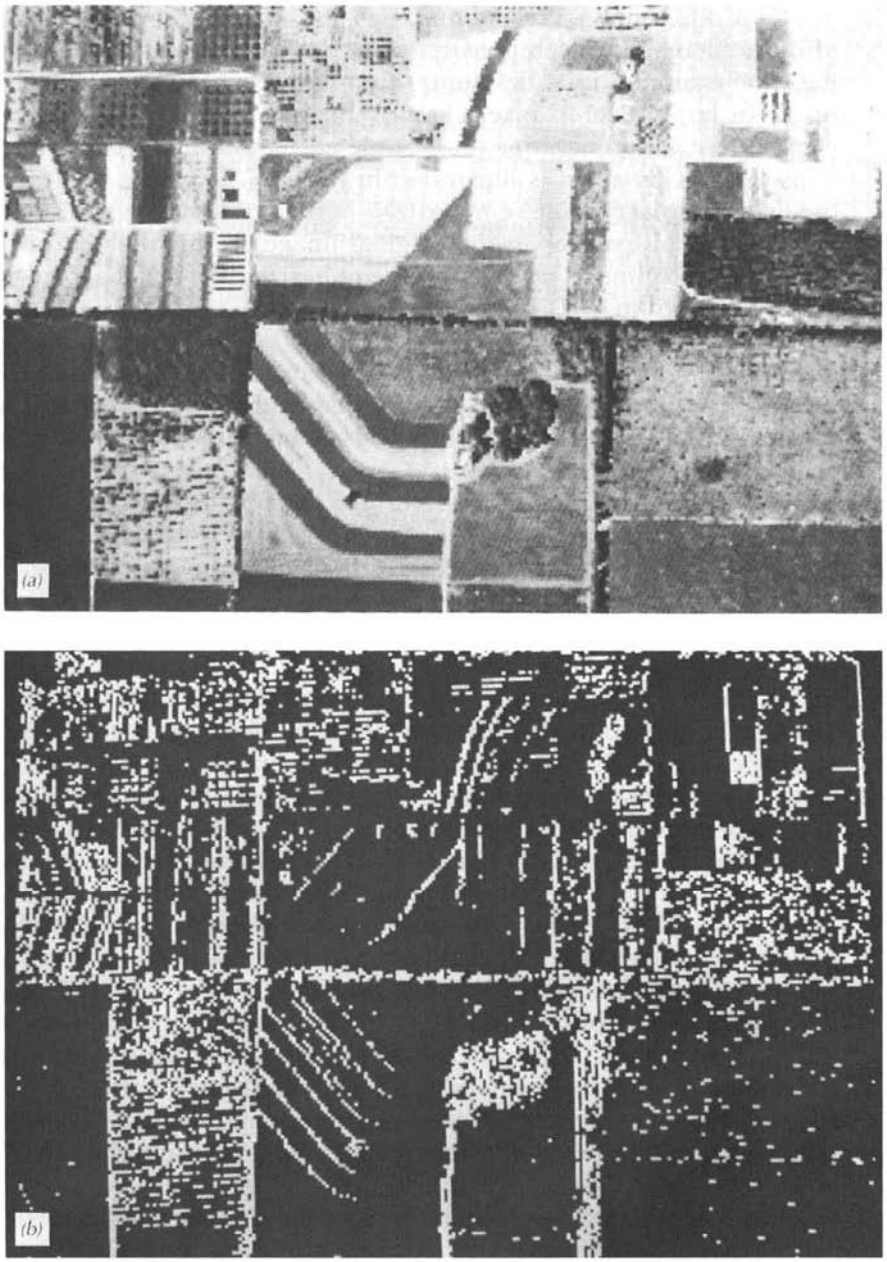


Figure 7.20 Effect of directional first differencing: (*a*) original image; (*b*) horizontal first difference; (*c*) vertical first difference; (*d*) diagonal first difference.

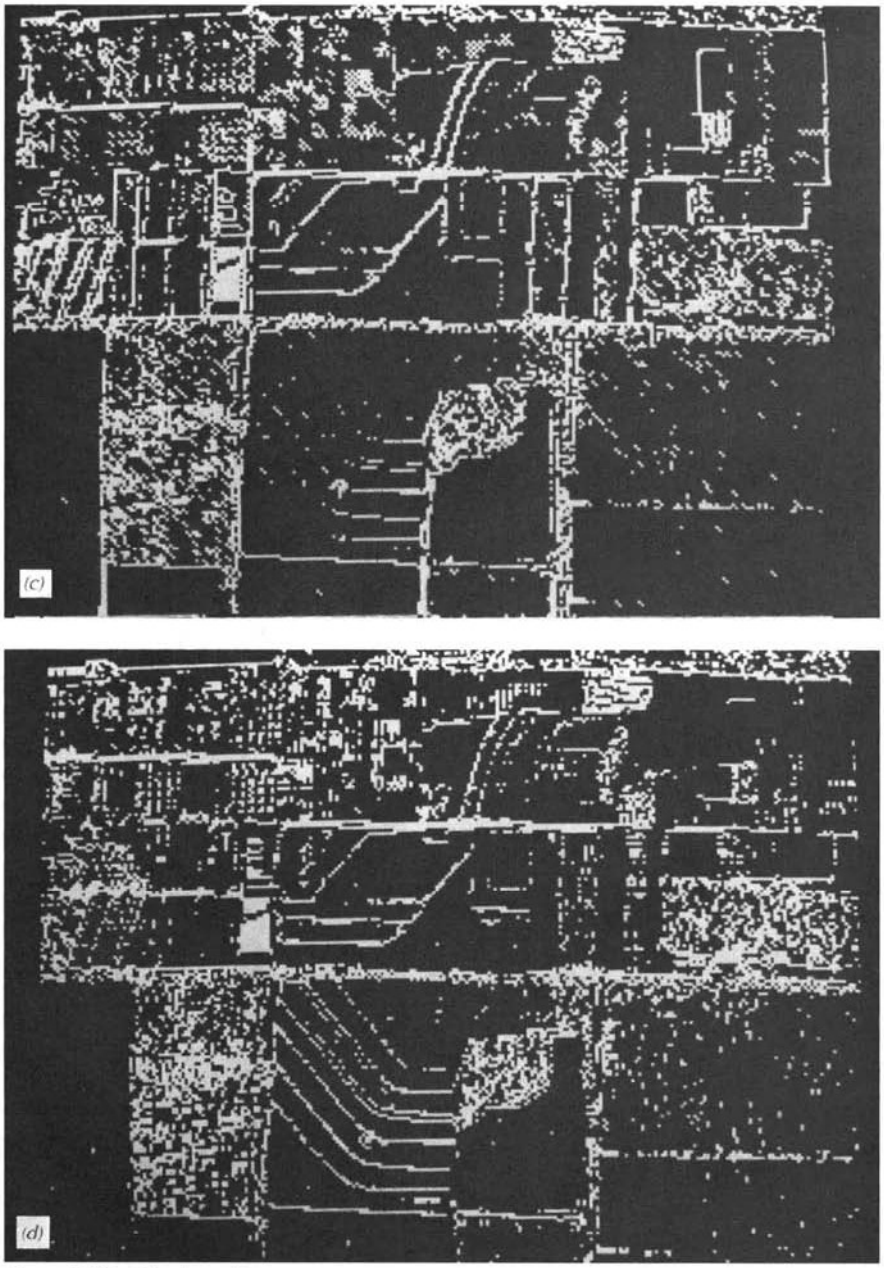


Figure 7.20 *(Continued)*

Figure 7.21 Edge enhancement through cross-diagonal first differencing.

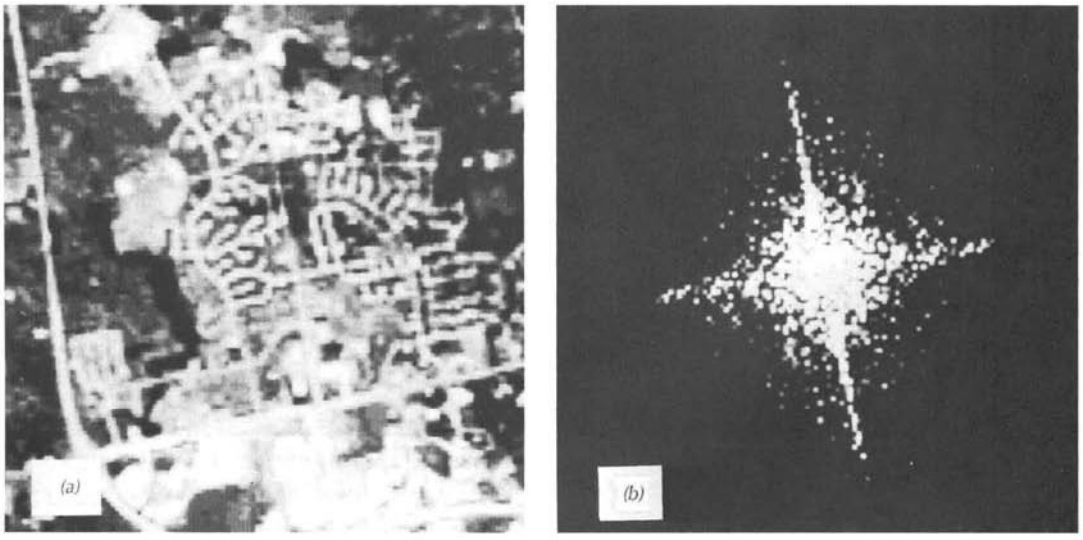


Figure 7.22 Applicaiton of Fourier transform: (*a*) original scene; (*b*) Fourier spectrum of (*a*).

image processing operations. For example, spatial filtering can be accomplished by applying a filter directly on the Fourier spectrum and then performing an inverse transform. This is illustrated in Figure 7.23. In Figure 7.23*a*, a circular high frequency blocking filter has been applied to the Fourier spectrum shown previously in Figure 7.22*b*. Note that this image is a low pass

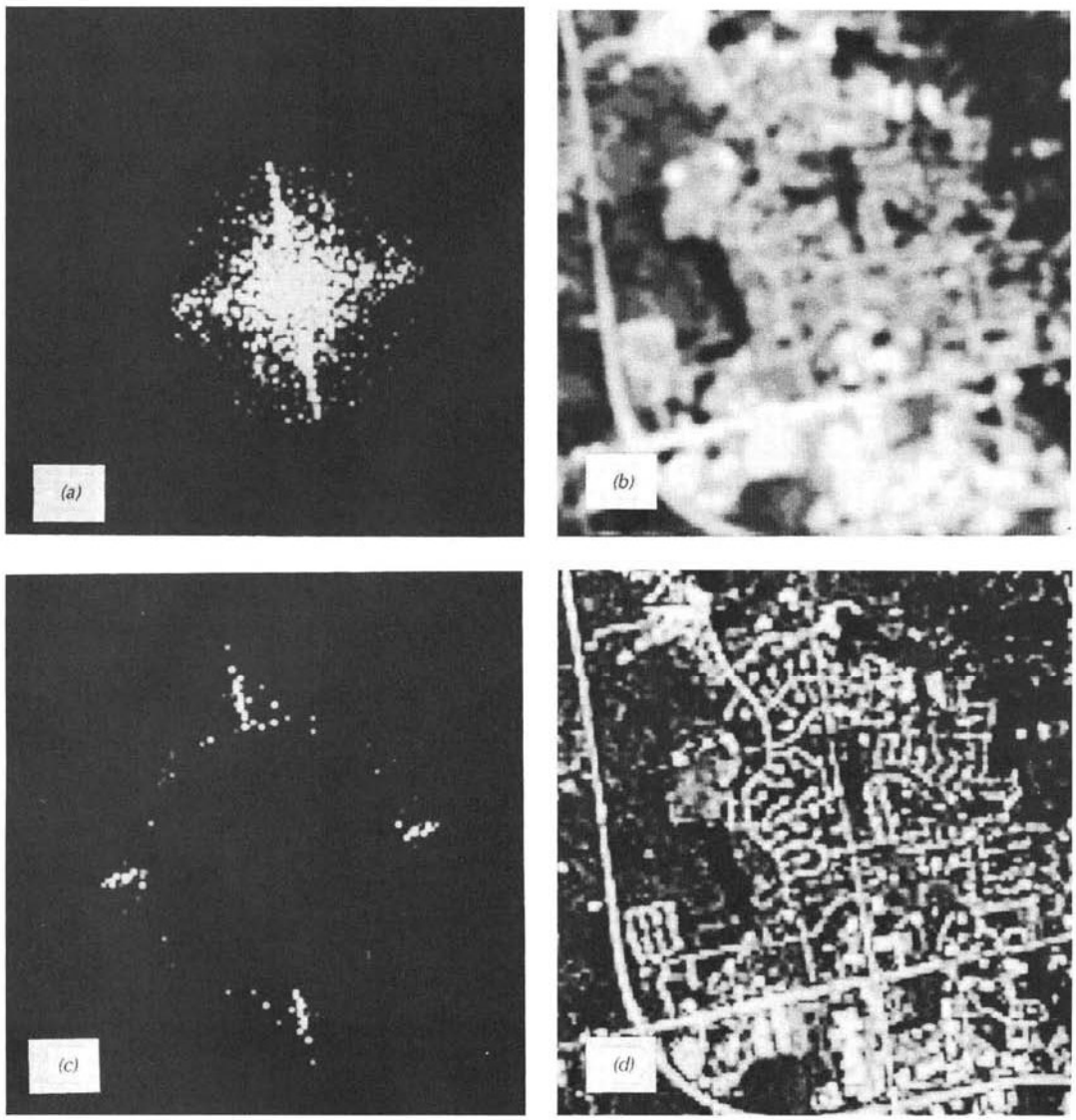


Figure 7.23 Spatial filtering in the frequency domain: (*a*) high frequency blocking filter; (*b*) inverse transform of (*a*); (*c*) low frequency blocking filter; (*d*) inverse transform of (*c*).

filtered version of the original scene. Figures 7.23*c* and *d* illustrate the application of a circular low frequency blocking filter (*c*) to produce a high pass filtered enhancement (*d*).

Figure 7.24 illustrates another common application of Fourier analysis—the elimination of image noise. Shown in Figure 7.24*a* is an airborne multispectral scanner image containing substantial noise. The Fourier spectrum

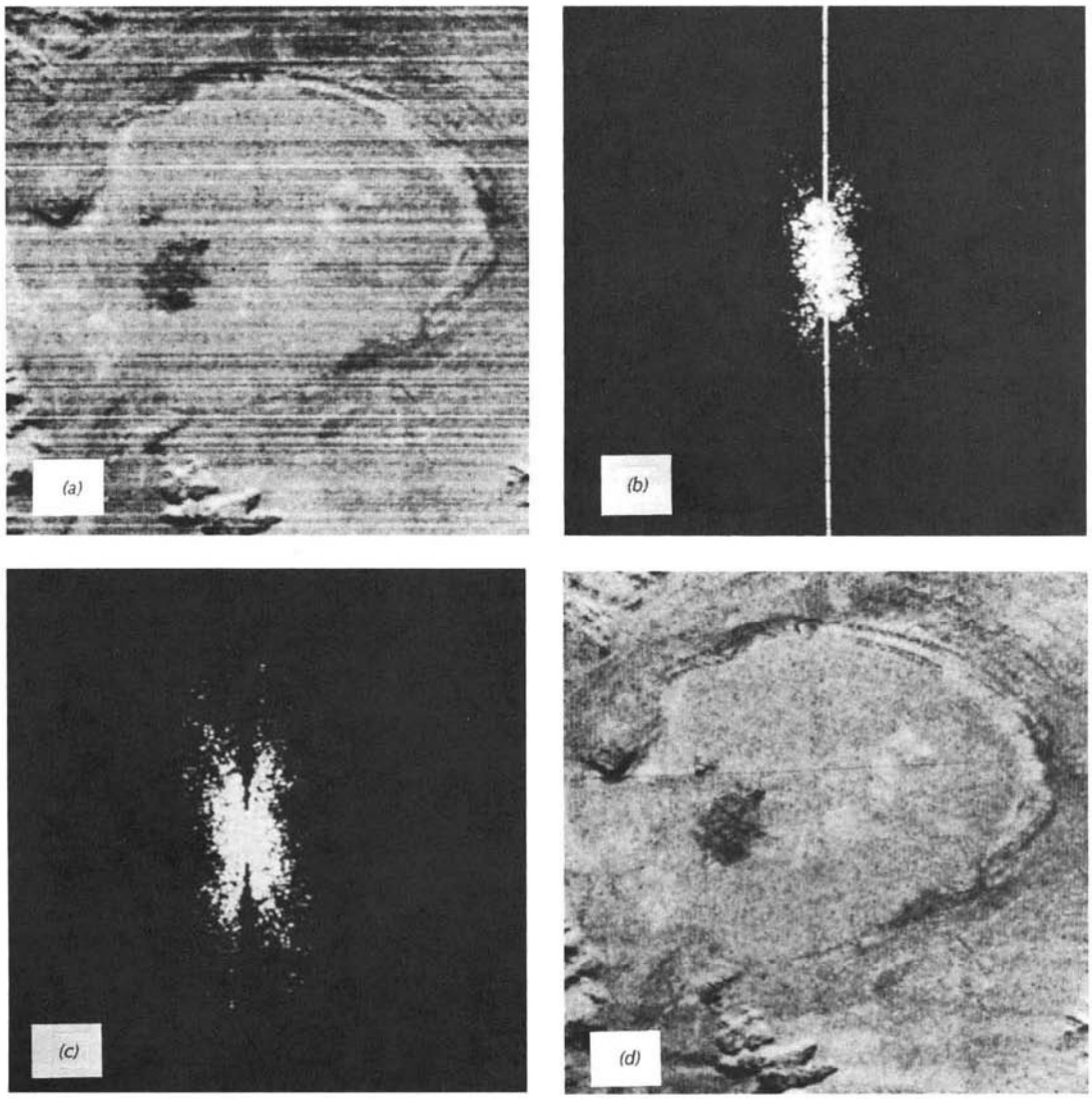


Figure 7.24 Noise elimination in the frequency domain. (*a*) Airborne multispectral scanner image containing noise. (Courtesy NASA.) (*b*) Fourier spectrum of (*a*). (*c*) Wedge block filter. (*d*) Inverse transform of (*c*).

for the image is shown in Figure 7.24*b*. Note that the noise pattern, which occurs in a horizontal direction in the original scene, appears as a band of frequencies trending in the vertical direction in the Fourier spectrum. In Figure 7.24*c* a vertical *wedge block filter* has been applied to the spectrum. This filter passes the lower frequency components of the image but blocks the high frequency components of the original image trending in the horizontal direction. Figure 7.24*d* shows the inverse transform of (*c*). Note how effectively this operation eliminates the noise inherent in the original image.

Fourier analysis is useful in a host of image processing operations in addition to the spatial filtering and image restoration applications we have illustrated in this discussion. However, most image processing is currently implemented in the spatial domain because of the number and complexity of computations required in the frequency domain. (This situation is likely to change with improvements in computer hardware and advances in research on the spatial attributes of digital image data.)

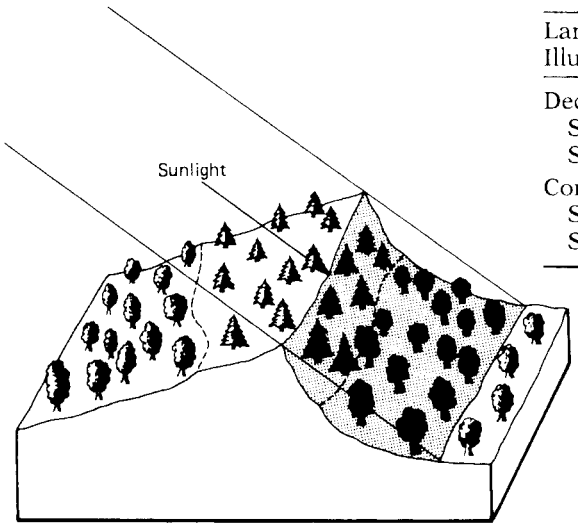
Before leaving the topic of spatial feature manipulation, it should be reemphasized that we have illustrated only a representative subset of the range of possible processing techniques available. Several of the references included at the end of this chapter describe and illustrate numerous other procedures that may be of interest to the reader.

7.6 MULTI-IMAGE MANIPULATION

Spectral Ratioing

Ratio images are enhancements resulting from the division of DN values in one spectral band by the corresponding values in another band. A major advantage of ratio images is that they convey the spectral or color characteristics of image features, regardless of variations in scene illumination conditions. This concept is illustrated in Figure 7.25, which depicts two different land cover types (deciduous and coniferous trees) occurring on both the sunlit and shadowed sides of a ridge line. The DNs observed for each cover type are substantially lower in the shadowed area than in the sunlit area. However, the ratio values for each cover type are nearly identical, irrespective of the illumination condition. Hence, a ratioed image of the scene effectively compensates for the brightness variation caused by the varying topography and emphasizes the color content of the data.

Ratioed images are often useful for discriminating subtle *spectral* variations in a scene that are masked by the *brightness* variations in images from individual spectral bands or in standard color composites. This enhanced discrimination is due to the fact that ratioed images clearly portray the variations in the *slopes* of the spectral reflectance curves between the two bands involved, regardless of the absolute reflectance values observed in the bands.



Land Cover/ Illumination	Digital Number Band A	Band B	Ratio (Band A/Band B)
Deciduous			
Sunlit	48	50	0.96
Shadow	18	19	0.95
Coniferous			
Sunlit	31	45	0.69
Shadow	11	16	0.69

Figure 7.25 Reduction of scene illumination effects thorugh spectral ratioing. (Adapted from Sabins, 1997.)

These slopes are typically quite different for various material types in certain bands of sensing. For example, the near-infrared-to-red ratio for healthy vegetation is normally very high. That for stressed vegetation is typically lower (as near-infrared reflectance decreases and the red reflectance increases). Thus a near-infrared-to-red (or red-to-near-infrared) ratioed image might be very useful for differentiating between areas of the stressed and nonstressed vegetation. This type of ratio has also been employed extensively in vegetation indices aimed at quantifying relative vegetation greenness and biomass.

Obviously, the utility of any given spectral ratio depends upon the particular reflectance characteristics of the features involved and the application at hand. The form and number of ratio combinations available to the image analyst also varies depending upon the source of the digital data. The number of possible ratios that can be developed from n bands of data is $n(n - 1)$. Thus, for Landsat MSS data, $4(4 - 1)$, or 12, different ratio combinations are possible (six original and six reciprocal). For the six nonthermal bands of Landsat TM or ETM+ data there are $6(6 - 1)$, or 30, possible combinations.

Figure 7.26 illustrates four representative ratio images generated from TM data. These images depict higher ratio values in brighter tones. Shown in (*a*) is the ratio TM1/TM2. Because these two bands are highly correlated for this scene, the ratio image has low contrast. In (*b*) the ratio TM3/TM4 is depicted so that features such as water and roads, which reflect highly in the red band (TM3) and little in the near-infrared band (TM4), are shown in lighter tones. Features such as vegetation appear in darker tones because of their rel-

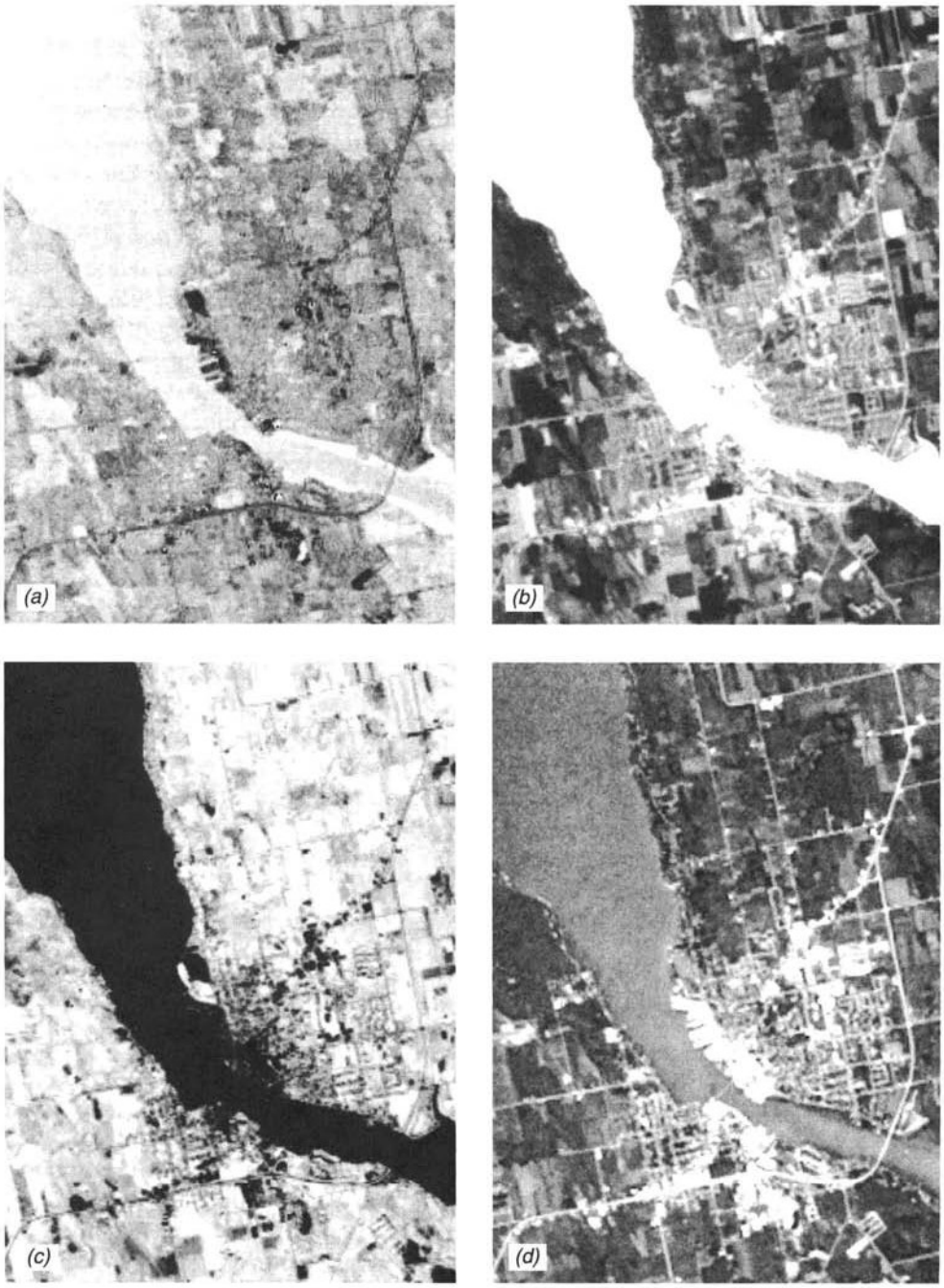


Figure 7.26 Ratioed images derived from midsummer Landsat TM data, near Sturgeon Bay, WI (higher ratio values are displayed in brighter image tones): (*a*) TM1/TM2; (*b*) TM3/TM4; (*c*) TM5/TM2; (*d*) TM3/TM7.

atively low reflectance in the red band (TM3) and high reflectance in the near infrared (TM4). In (*c*) the ratio TM5/TM2 is shown. Here, vegetation generally appears in light tones because of its relatively high reflectance in the mid-infrared band (TM5) and its comparatively lower reflectance in the green band (TM2). However, note that certain vegetation types do not follow this trend due to their particular reflectance characteristics. They are depicted in very dark tones in this particular ratio image and can therefore be discriminated from the other vegetation types in the scene. Part (*d*) shows the ratio TM3/TM7. Roads and other cultural features appear in lighter tone in this image due to their relatively high reflectance in the red band (TM3) and low reflectance in the mid-infrared band (TM7). Similarly, differences in water turbidity are readily observable in this ratio image.

Ratio images can also be used to generate false color composites by combining three monochromatic ratio data sets. Such composites have the twofold advantage of combining data from more than two bands and presenting the data in color, which further facilitates the interpretation of subtle spectral reflectance differences. Choosing which ratios to include in a color composite and selecting colors in which to portray them can sometimes be difficult. For example, excluding reciprocals, 20 color combinations are possible when the 6 original ratios of Landsat MSS data are displayed 3 at a time. The 15 original ratios of nonthermal TM data result in 455 different possible combinations.

Various quantitative criteria have been developed to assist in selecting which ratio combinations to include in color composites. The *Optimum Index Factor* (*OIF*) is one such criterion (Chavez, Berlin, and Sowers, 1982). It ranks all possible three-ratio combinations based on the total variance present in each ratio and the degree of correlation between ratios. That combination containing the most variance and least correlation is assumed to convey the greatest amount of information throughout a scene. A limitation of this procedure is that the best combination for conveying the *overall* information in a scene may not be the best combination for conveying the *specific* information desired by the image analyst. Hence, some trial and error is often necessary in selecting ratio combinations.

Certain caution should be taken when generating and interpreting ratio images. First, it should be noted that such images are "intensity blind." That is, dissimilar materials with different absolute radiances but having similar slopes of their spectral reflectance curves may appear identical. This problem is particularly troublesome when these materials are contiguous and of similar image texture. One way of minimizing this problem is by using a *hybrid color ratio composite*. This product is prepared by displaying two ratio images in two of the primary colors but using the third primary color to display an individual band of data. This restores a portion of the lost absolute radiance information and some of the topographic detail that may be needed to dis-

criminate between certain features. (As we illustrate later, IHS color space transformations can also be used for this purpose.)

Noise removal is an important prelude to the preparation of ratio images since ratioing enhances noise patterns that are uncorrelated in the component images. Furthermore, ratios only compensate for multiplicative illumination effects. That is, division of DNs or radiances for two bands cancels only those factors that are operative equally in the bands and not those that are additive. For example, atmospheric haze is an additive factor that might have to be removed prior to ratioing to yield acceptable results. Alternatively, ratios of between-band differences and/or sums may be used to improve image interpretability in some applications.

The manner in which ratios are computed and displayed will also greatly influence the information content of a ratio image. For example, the ratio between two raw DNs for a pixel will normally be quite different from that between two radiance values computed for the same pixel. The reason for this is that the detector response curves for the two channels will normally have different offsets, which are additive effects on the data. (This situation is akin to the differences one would obtain by ratioing two temperatures using the Fahrenheit scale versus the Celsius scale.) Some trial and error may be necessary before the analyst can determine which form of ratio works best for a particular application.

It should also be noted that ratios can "blow up" mathematically (become equal to infinity) in that divisions by zero are possible. At the same time, ratios less than 1 are common and rounding to integer values will compress much of the ratio data into gray level 0 or 1. Hence, it is important to scale the results of ratio computations somehow and relate them to the display device used. One means of doing this is to employ an algorithm of the form

$$DN' = R \arctan\left(\frac{DN_X}{DN_Y}\right) \tag{7.7}$$

where

DN' = digital number in ratio image
R = scaling factor to place ratio data in appropriate integer range
$\arctan (DN_X/DN_Y)$ = angle (in radians) whose tangent is the ratio of the digital numbers in bands X and Y; if DN_Y equals 0, this angle is set to 90°

In the above equation the angle whose tangent is equal to the ratio of the two bands can range from 0° to 90°, or from 0 to approximately 1.571 rad. Therefore, DN′ can range from 0 to approximately $1.571R$. If an 8-bit display is used, R is typically chosen to be 162.3, and DN′ can then range from 0 to 255.

Principal and Canonical Components

Extensive interband correlation is a problem frequently encountered in the analysis of multispectral image data (later illustrated in Figure 7.49). That is, images generated by digital data from various wavelength bands often appear similar and convey essentially the same information. Principal and canonical component transformations are two techniques designed to reduce such redundancy in multispectral data. These transformations may be applied either as an enhancement operation prior to visual interpretation of the data or as a preprocessing procedure prior to automated classification of the data. If employed in the latter context, the transformations generally increase the computational efficiency of the classification process because both principal and canonical component analyses may result in a reduction in the dimensionality of the original data set. Stated differently, the purpose of these procedures is to compress all of the information contained in an original n-band data set into fewer than n "new bands." The new bands are then used in lieu of the original data.

A detailed description of the statistical procedures used to derive principal and canonical component transformations is beyond the scope of this discussion. However, the concepts involved may be expressed graphically by considering a two-band image data set such as that shown in Figure 7.27. In (*a*), a random sample of pixels has been plotted on a scatter diagram according to their gray levels as originally recorded in bands A and B. Superimposed on the band A–band B axis system are two new axes (axes I and II) that are rotated with respect to the original measurement axes and that have their origin at the mean of the data distribution. Axis I defines the direction of the *first*

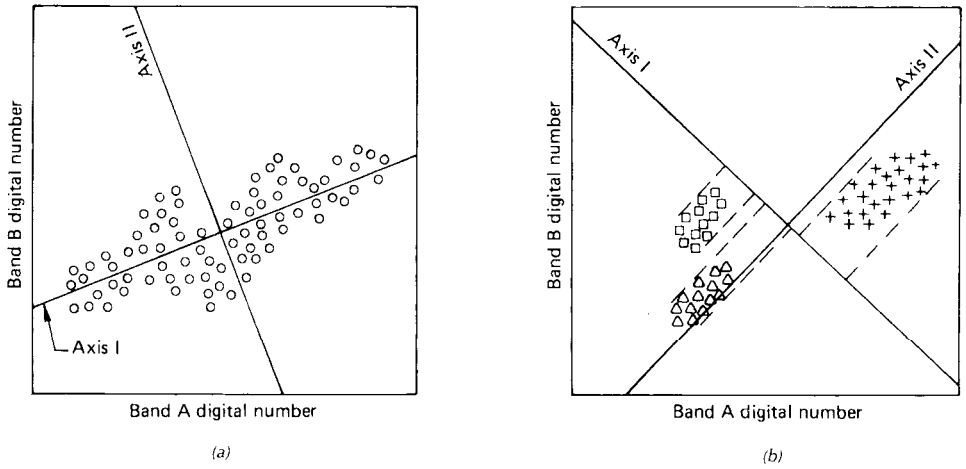


Figure 7.27 Rotated coordinate axes used in (*a*) principal component and (*b*) canonical component transformations.

principal component and axis II defines the direction of the *second principal component*. The form of the relationship necessary to transform a data value in the original band A–band B coordinate system into its value in the new axis I–axis II system is

$$DN_I = a_{11}DN_A + a_{12}DN_B \qquad DN_{II} = a_{21}DN_A + a_{22}DN_B \tag{7.8}$$

where

DN_I, DN_{II} = digital numbers in new (principal component image) coordinate system
DN_A, DN_B = digital numbers in old (original) coordinate system
$a_{11}, a_{12}, a_{21}, a_{22}$ = coefficients for the transformation

In short, the principal component image data values are simply linear combinations of the original data values multiplied by the appropriate transformation coefficients. These coefficients are statistical quantities known as *eigenvectors* or *principal components*. They are derived from the variance/covariance matrix for the original image data.

Hence, a principal component image results from the linear combination of the original data and the eigenvectors on a pixel-by-pixel basis throughout the image. Often, the resulting principal component *image* is loosely referred to as simply a *principal component*. This is theoretically incorrect in that the eigenvalues themselves are the principal components, but we will sometimes not make this distinction elsewhere in this book.

It should be noted in Figure 7.27*a* that the data along the direction of the first principal component (axis I) have a greater variance or dynamic range than the data plotted against either of the original axes (bands A and B). The data along the second principal component direction have far less variance. This is characteristic of all principal component images. In general, the first principal component image (PC1) includes the largest percentage of the total scene variance and succeeding component images (PC2, PC3, . . . , PC*n*) each contain a decreasing percentage of the scene variance. Furthermore, because successive components are chosen to be orthogonal to all previous ones, the data they contain are uncorrelated.

Principal component enhancements are generated by displaying contrast-stretched images of the transformed pixel values. We illustrate the nature of these displays by considering the Landsat MSS images shown in Figure 7.28. This figure depicts the four MSS bands of a scene covering the Sahl al Matran area, Saudi Arabia. Figure 7.29 shows the principal component images for this scene. Some areas of geologic interest labeled in Figure 7.28 are (A) alluvial material in a dry stream valley, (B) flat-lying quaternary and tertiary basalts, and (C) granite and granodiorite intrusion.

Note that in Figure 7.29, PC1 expresses the majority (97.6 percent) of the variance in the original data set. Furthermore, PC1 and PC2 explain virtually all of the variance in the scene (99.4 percent). This compression of

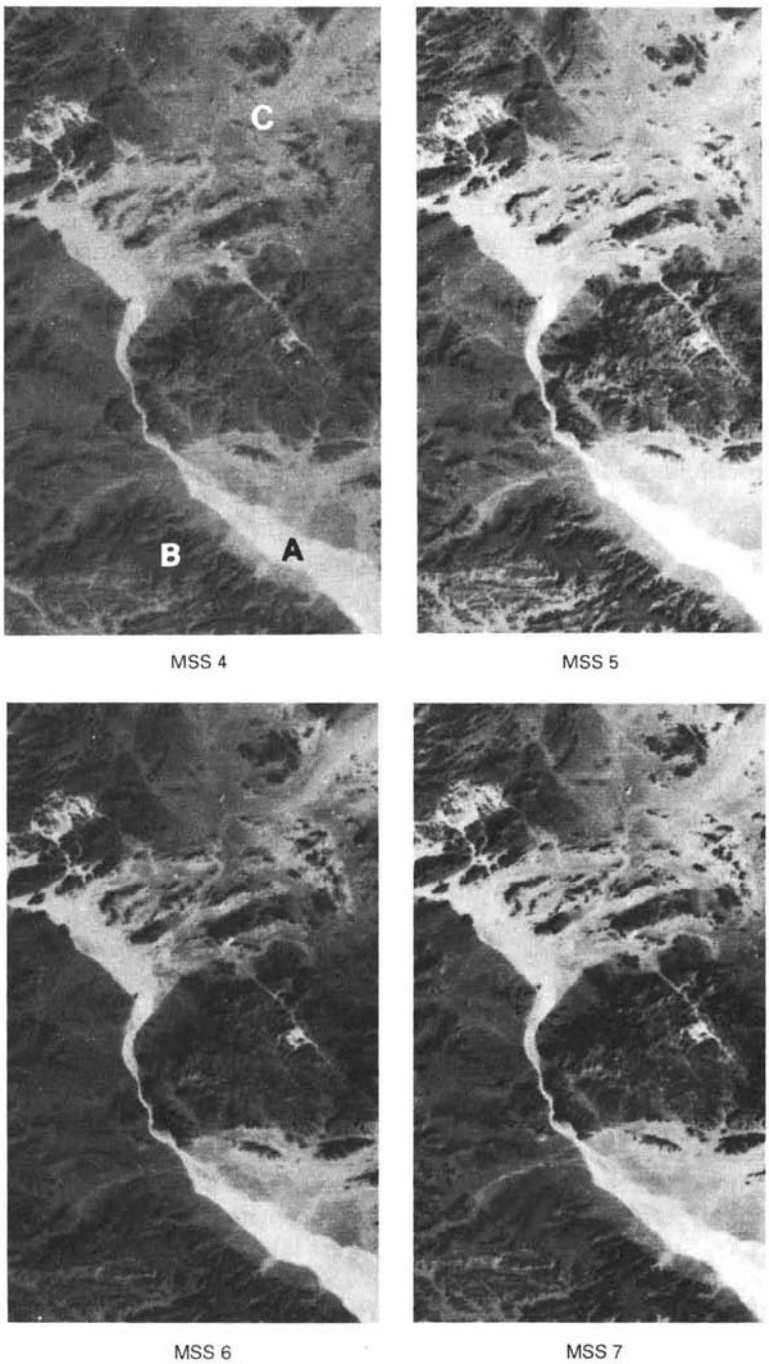


Figure 7.28 Four MSS bands covering the Sahl al Matran area of Saudi Arabia. Note the redundancy of information in these original image displays. (Courtesy NASA.)

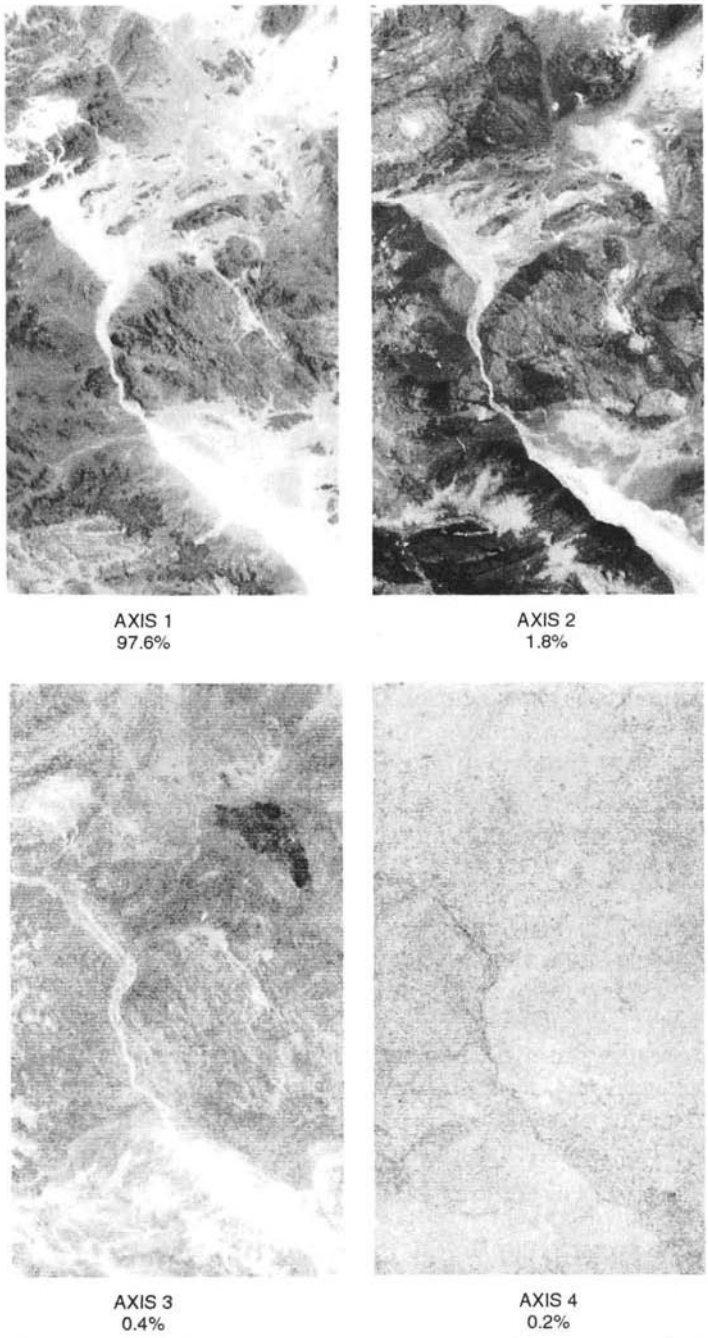


Figure 7.29 Transformed data resulting from principal component analysis of the MSS data shown in Figure 7.28. The percentage of scene variance contained in each axis is indicated. (Courtesy NASA.)

image information in the first two principal component images of Landsat MSS data is typical. Because of this, we refer to the *intrinsic dimensionality* of Landsat MSS data as being effectively 2. Also frequently encountered with Landsat MSS data, the PC4 image for this scene contains virtually no information and tends to depict little more than system noise. However, note that both PC2 and PC3 illustrate certain features that were obscured by the more dominant patterns shown in PC1. For example, a semicircular feature (labeled C in Figure 7.28) is clearly defined in the upper right portion of the PC2 and PC3 images (appearing bright and dark, respectively). This feature was masked by more dominant patterns both in the PC1 image and in all bands of the original data. Also, its tonal reversal in PC2 and PC3 illustrates the lack of correlation between these images.

As in the case of ratio images, principal component images can be analyzed as separate black and white images (as illustrated here), or any three component images may be combined to form a color composite. If used in an image classification process, principal component data are normally treated in the classification algorithm simply as if they were original data. However, the number of components used is normally reduced to the intrinsic dimensionality of the data, thereby making the image classification process much more efficient by reducing the amount of computation required. (For example, Landsat TM or ETM+ data can often be reduced to just three principal component images for classification purposes.)

Principal component enhancement techniques are particularly appropriate where little prior information concerning a scene is available. *Canonical component analysis*, also referred to as multiple discriminant analysis, may be more appropriate when information about particular features of interest is known. Recall that the principal component axes shown in Figure 7.27*a* were located on the basis of a random, undifferentiated sample of image pixel values. In Figure 7.27*b*, the pixel values shown are derived from image areas containing three different analyst-defined feature types (the feature types are represented by the symbols △, □, and +). The canonical component axes in this figure (axes I and II) have been located to maximize the *separability* of these classes while minimizing the variance within each class. For example, the axes have been positioned in this figure such that the three feature types can be discriminated solely on the basis of the first canonical component images (CC1) values located along axis I.

In Figure 7.30, canonical component images are shown for the Landsat MSS scene that was shown in Figure 7.28. Once again, CC1 expresses the highest percentage variation in the data with subsequent component images representing lesser amounts of uncorrelated additional information. These displays, as in cases such as feature C, may enhance subtle features not evident in the original image data. Like principal component data, canonical component data can also be used in image classification. Canonical component images not only improve classification efficiency but also can improve

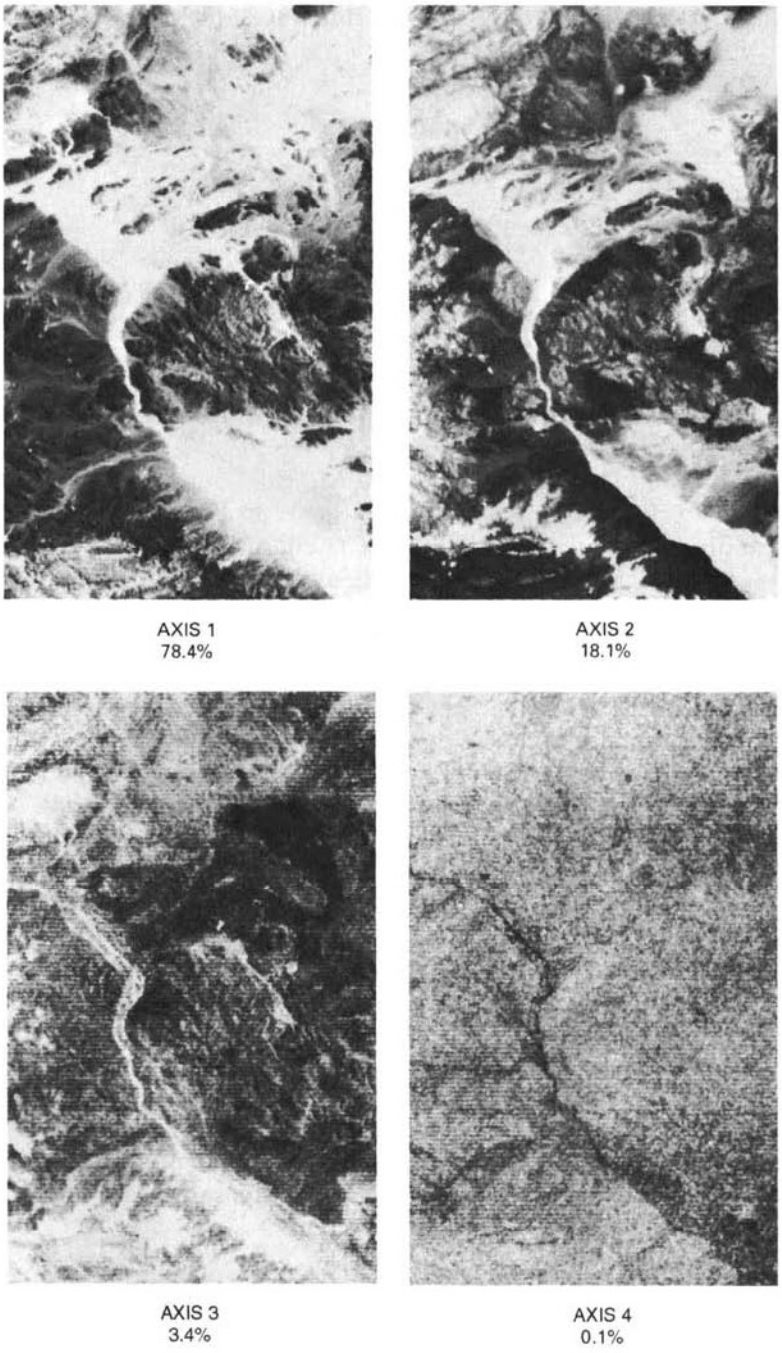


Figure 7.30 Transformed data resulting from canonical component analysis of the MSS data shown in Figure 7.28. The percentage of scene variance contained in each axis is indicated. (Courtesy NASA.)

classification accuracy for the identified features due to the increased spectral separability of classes.

Vegetation Components

Previously (Section 6.16), we introduced the concept of vegetation indices and the use of between-band differences and ratios to produce vegetation index images from AVHRR data. Here we wish to point out that numerous other forms of linear data transformations have been developed for vegetation monitoring, with differing sensors and vegetation conditions dictating different transformations. For example, Kauth and Thomas (1976) derived a linear transformation of the four Landsat MSS bands that established four new axes in the spectral data that can be interpreted as *vegetation components* useful for agricultural crop monitoring. This "tasseled cap" transformation rotates the MSS data such that the majority of information is contained in two components or features that are directly related to physical scene characteristics. Brightness, the first feature, is a weighted sum of all bands and is defined in the direction of the principal variation in soil reflectance. The second feature, greenness, is approximately orthogonal to brightness and is a contrast between the near-infrared and visible bands. Greenness is strongly related to the amount of green vegetation present in the scene. Brightness and greenness together typically express 95 percent or more of the total variability in MSS data and have the characteristic of being readily interpretable features generally applicable from scene to scene (once illumination and atmospheric effects are normalized).

Crist and Cicone (1984) extended the tasseled cap concept to Landsat TM data and found that the six bands of reflected data effectively occupy three dimensions, defining planes of soils, vegetation, and a transition zone between them. The third feature, called wetness, relates to canopy and soil moisture.

Figure 7.31 illustrates the application of the tasseled cap transformation to TM data acquired over north central Nebraska. The northern half of the area is dominated by circular corn fields that have been watered throughout the summer by center-pivot irrigators. The southern half of the area is at the edge of the Nebraska Sand Hills, which are covered by grasslands. Lighter tones in these images correspond to larger values of each component.

Figure 7.32 illustrates another vegetation transformation, namely, the *transformed vegetation index* (*TVI*), applied to the same data set shown in Figure 5.8 and Plate 14. The TVI is computed as

$$\text{TVI} = \left[\frac{\text{DN (near IR)} - \text{DN (red)}}{\text{DN (near IR)} + \text{DN (red)}} + 0.5\right]^{1/2} \times 100 \qquad (7.9)$$

where

DN (near IR) = digital number in the near-IR band
DN (red) = digital number in the red band

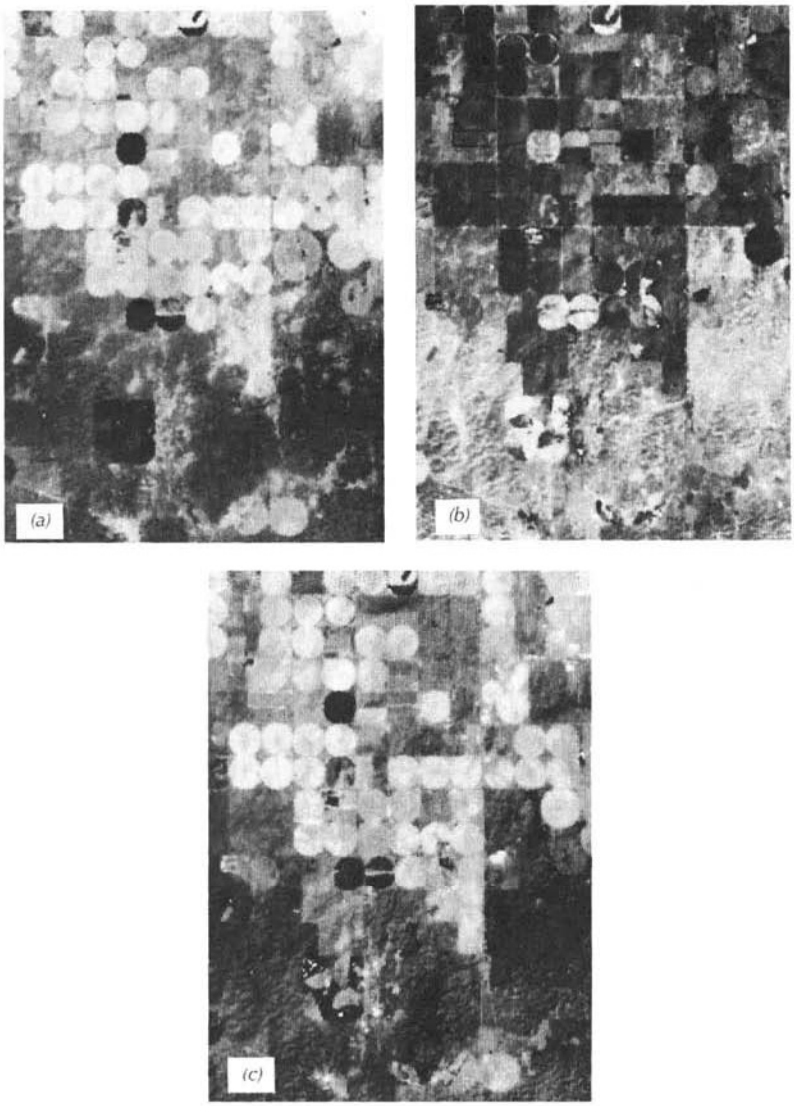


Figure 7.31 Tasseled cap transformation for a late-summer TM image of north central Nebraska. The three components shown here illustrate (*a*) relative greenness, (*b*) brightness, and (*c*) wetness. (Courtesy Institute of Agriculture and Natural Resources, University of Nebraska.)

Using ground reference data, it is frequently possible to "calibrate" TVI values to the green biomass present on a pixel-by-pixel basis. Usually, separate calibration relationships must be established for each cover type present in an image. These relationships may then be used in such applications as "precision crop management" or precision farming to guide the application of

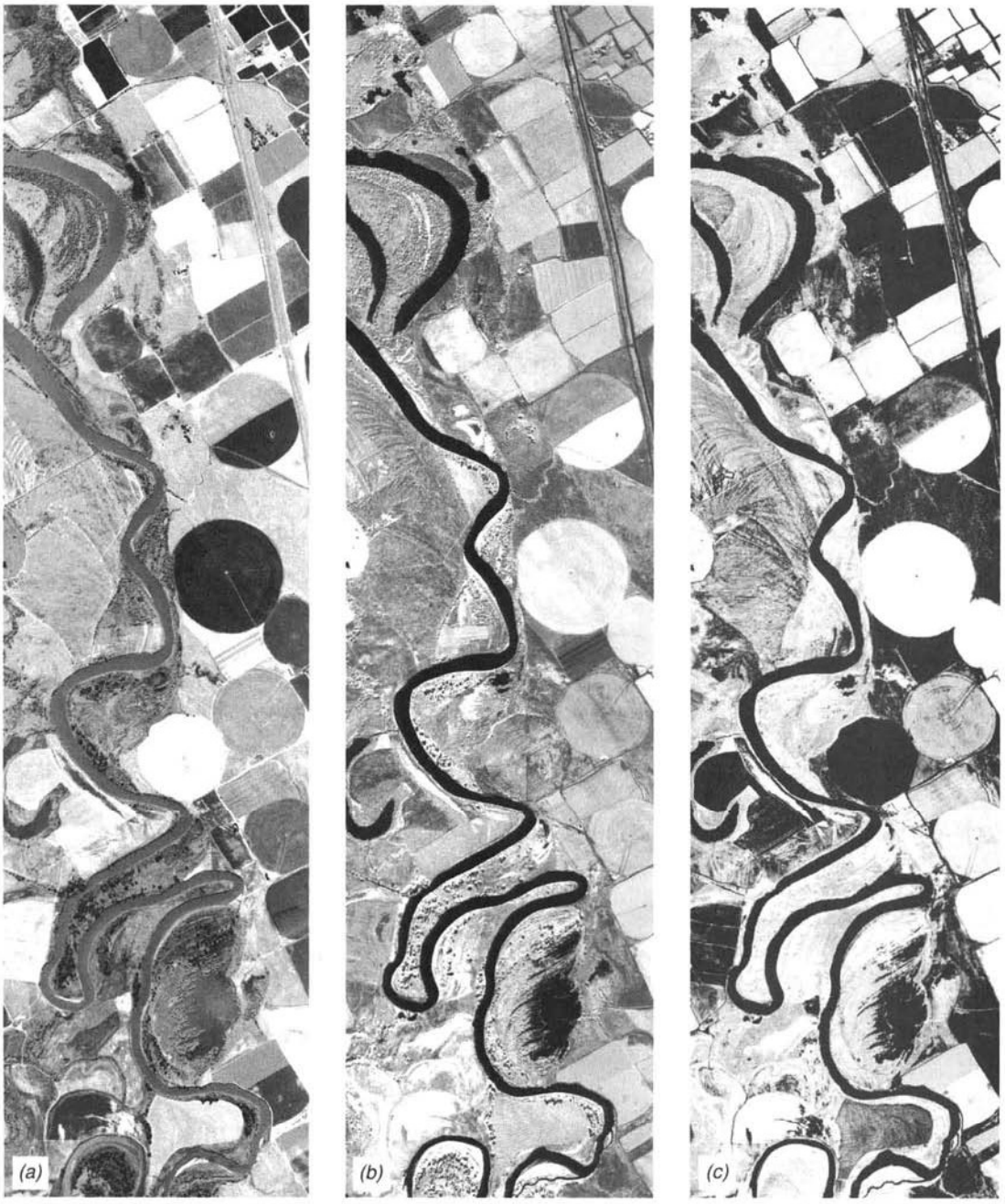


Figure 7.32 Transformed vegetation index (TVI) image derived from mid-August multispectral scanner data, Yakima River valley, WA: (*a*) red band; (*b*) near-IR band; (*c*) TVI image. Same image area as shown in Figure 5.8 and Plate 14. (Courtesy Sensys Technologies, Inc.)

irrigation water, fertilizers, herbicides, and so on (Section 4.6). Similarly, TVI values have been used to aid in making ranch management decision when the TVI data correlate with the estimated level of forage present in pastures contained in an image (Miller et al., 1985).

Indices such as the TVI and NDVI, which are based on the near-infrared and red spectral bands, have been shown to be well correlated not only with crop biomass accumulation, but also with leaf chlorophyll levels, leaf area index values, and the photosynthetically active radiation absorbed by a crop canopy. However, when such biophysical parameters reach moderate to high levels, the *green normalized difference vegetation index* (*GNDVI*) may be a more reliable indicator of crop conditions. The GNDVI is identical in form to the NDVI except that the green band is substituted for the red band.

With the availability of MODIS data on a global basis, several new vegetation indices have been proposed. For example, the *enhanced vegetation index* (*EVI*) has been developed as a modified NDVI with an adjustment factor to minimize soil background influences and a blue band correction of red band data to lessen atmospheric scattering. Additional treatment of this and the scores of other vegetation indices in use is beyond the scope of this discussion. Readers interested in additional information on this subject are encouraged to consult any of the numerous summaries of vegetation indices available in the literature (e.g., Richardson and Everitt, 1992; Running et al., 1994; Lyon et al., 1998; and Jensen, 2000).

Intensity–Hue–Saturation Color Space Transformation

Digital images are typically displayed as additive color composites using the three primary colors: red, green, and blue (RGB). Figure 7.33 illustrates the interrelation among the RGB components of a typical color display device (such as a color monitor). Shown in this figure is the RGB *color cube*, which is defined by the brightness levels of each of the three primary colors. For a display with 8-bit-per-pixel data encoding, the range of possible DNs for each color component is 0 to 255. Hence, there are 256^3 (or 16,777,216) possible combinations of red, green, and blue DNs that can be displayed by such a device. Every pixel in a composited display may be represented by a three-dimensional coordinate position somewhere within the color cube. The line from the origin of the cube to the opposite corner is known as the *gray line* since DNs that lie on this line have equal components of red, green, and blue.

The RGB displays are used extensively in digital processing to display normal color, false color infrared, and arbitrary color composites. For example, a normal color composite may be displayed by assigning TM or

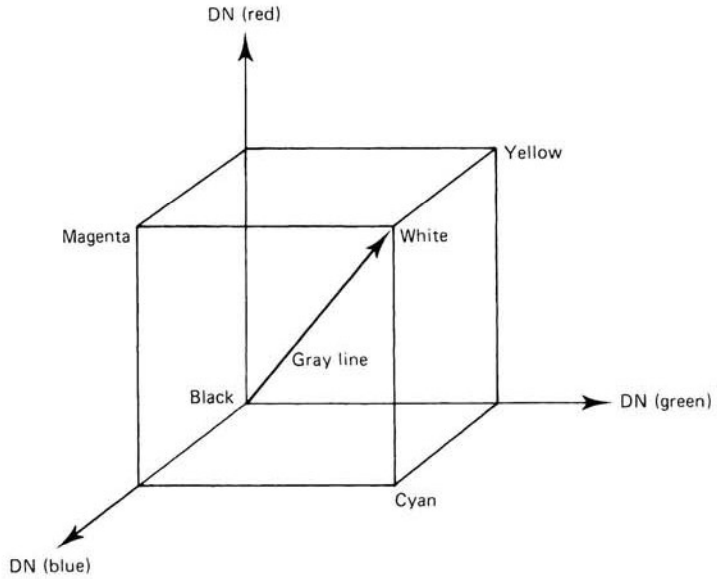


Figure 7.33 The RGB color cube. (Adapted from Schowengerdt, 1997.)

ETM+ bands 1, 2, and 3 to the blue, green, and red components, respectively. A false color infrared composite results when bands 2, 3, and 4 are assigned to these respective components. Arbitrary color composites result when other bands or color assignments are used. Color composites may be contrast stretched on a RGB display by manipulating the contrast in each of the three display channels (using a separate lookup table for each of the three color components).

An alternative to describing colors by their RGB components is the use of the *intensity–hue–saturation* (*IHS*) system. "Intensity" relates to the total brightness of a color. "Hue" refers to the dominant or average wavelength of light contributing to a color. "Saturation" specifies the purity of color relative to gray. For example, pastel colors such as pink have low saturation compared to such high saturation colors as crimson. Transforming RGB components into IHS components *before* processing may provide more control over color enhancements.

Figure 7.34 shows one (of several) means of transforming RGB components into IHS components. This particular approach is called the *hexcone model*, and it involves the *projection* of the RGB color cube onto a plane that is perpendicular to the gray line and tangent to the cube at the corner farthest from the origin. The resulting projection is a hexagon. If the plane of projection is moved from white to black along the gray line, successively smaller color *subcubes* are projected and a series of hexagons of decreasing size re-

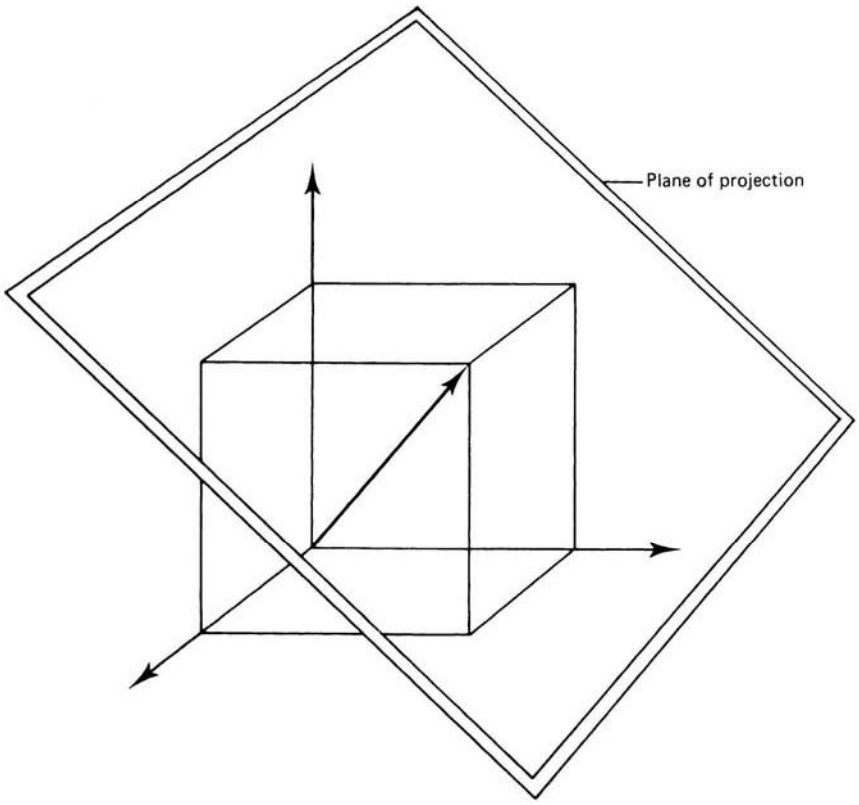


Figure 7.34 Planar projection of the RGB color cube. A series of such projections results when progressively smaller subcubes are considered between white and black.

sult. The hexagon at white is the largest and the hexagon at black degenerates to a point. The series of hexagons developed in this manner define a solid called the *hexcone* (Figure 7.35*a*).

In the hexcone model *intensity* is defined by the distance along the gray line from black to any given hexagonal projection. Hue and saturation are defined at a given intensity, within the appropriate hexagon (Figure 7.35*b*). *Hue* is expressed by the angle around the hexagon, and *saturation* is defined by the distance from the gray point at the center of the hexagon. The farther a point lies away from the gray point, the more saturated the color. (In Figure 7.35*b*, linear distances are used to define hue and saturation, thereby avoiding computations involving trigonometric functions.)

At this point we have established the basis upon which any pixel in the RGB color space can be transformed into its IHS counterpart. Such transformations are often useful as an intermediate step in image enhancement. This

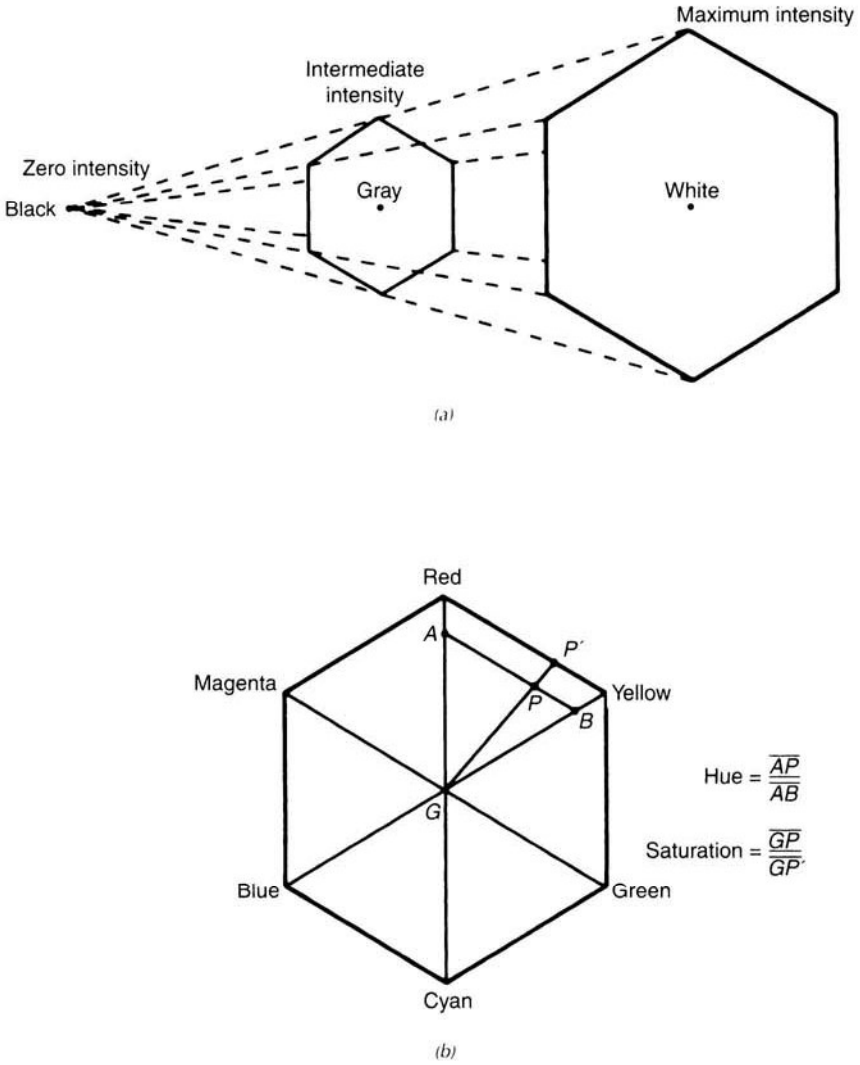


Figure 7.35 Hexcone color model. (*a*) Generation of the hexcone. The size of any given hexagon is determined by pixel intensity. (*b*) Definition of hue and saturation components for a pixel value, *P*, having a typical, nonzero intensity. (Adapted from Schowengerdt, 1997.)

is illustrated in Figure 7.36. In this figure the original RGB components are shown transformed first into their corresponding IHS components. The IHS components are then manipulated to enhance the desired characteristics of the image. Finally, these modified IHS components are transformed back to the RGB system for final display.

Among the advantages of IHS enhancement operations is the ability to vary each IHS component independently, without affecting the others. For example,

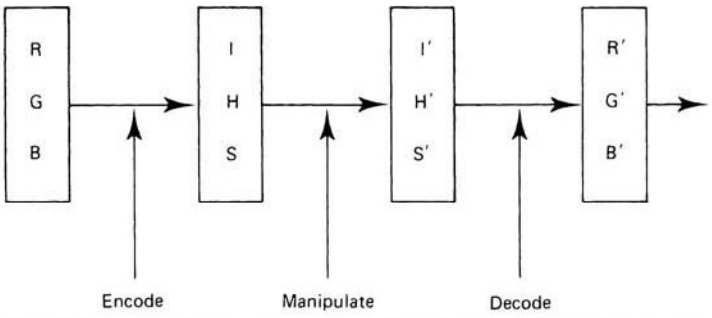


Figure 7.36 IHS/RGB encoding and decoding for interactive image manipulation. (Adapted from Schowengerdt, 1997.)

a contrast stretch can be applied to the intensity component of an image, and the hue and saturation of the pixels in the enhanced image will not be changed (as they typically are in RGB contrast stretches). The IHS approach may also be used to display spatially registered data of varying spatial resolution. For example, high resolution data from one source may be displayed as the intensity component, and low resolution data from another source may be displayed as the hue and saturation components. Such an approach was used to produce Plate 19 (Section 6.15). This plate is a merger of IKONOS 1-m-resolution panchromatic data (used in the intensity component) and 4-m-resolution multispectral data (hue and saturation components). The result is a composite image having the spatial resolution of the 1-m panchromatic data and the color characteristics of the original 4-m multispectral data. Similar procedures are used to merge other sources of same-sensor, multiresolution data sets (e.g., 10-m panchromatic and 20-m multispectral SPOT data, 15-m panchromatic and 30-m ETM+ data). Likewise, IHS techniques are often used to merge data from different sensing systems (e.g., digital orthophotos with satellite data).

One caution to be noted in using IHS transformations to merge multiresolution data is that direct substitution of the panchromatic data for the intensity component may not always produce the best final product in terms of color balance. In such situations, weighted combinations of the panchromatic and multispectral data might be used. The approach used in the production of Plate 19 was to employ a histogram matching operation to match the histogram of the new panchromatic data to that of the intensity component derived from the multispectral RGB data. The modified panchromatic data resulting from this operation were then used in the intensity component of the IHS image prior to transforming the data back to RGB format.

The development and application of various IHS encoding and enhancement schemes are the subject of continuing research. An interesting project in this regard has been the use of IHS transformations to display two bands (bands 1 and 2) of raw AVHRR data in a three-color composite. In such

composites, the sum of bands 1 and 2 is used to represent intensity. The band 2/1 ratio is used to define hue, and the difference between bands 1 and 2 is used to define saturation. The resulting image looks very similar to a standard color infrared composite.

Decorrelation Stretching

Decorrelation stretching is a form of multi-image manipulation that is particularly useful when displaying multispectral data that are highly correlated. Data from the NASA Thermal Infrared Multispectral Scanner (TIMS) and other hyperspectral data collected in the same region of the spectrum often fall into this category. Traditional contrast stretching of highly correlated data as R, G, and B displays normally only expands the range of intensities; it does little to expand the range of colors displayed, and the stretched image still contains only pastel hues. For example, no areas in a highly correlated image are likely to have high DNs in the red display channel but low values in the green and blue (which would produce a pure red). Instead, the reddest areas are merely a reddish-gray. To circumvent this problem, decorrelation stretching involves exaggeration of the least correlated information in an image primarily in terms of saturation, with minimal change in image intensity and hue.

As with IHS transformations, decorrelation stretching is applied in a transformed image space, and the results are then transformed back to the RGB system for final display. The major difference in decorrelation stretching is that the transformed image space used is that of the original image's principal components. The successive principal components of the original image are stretched independently along the respective principal component axes (Figure 7.27*a*). By definition, these axes are statistically independent of one another so the net effect of the stretch is to emphasize the poorly correlated components of the original data. When the stretched data are then transformed back to the RGB system, a display having increased color saturation results. There is usually little difference in the perceived hues and intensities due to enhancement. This makes interpretation of the enhanced image straightforward, with the decorrelated information exaggerated primarily in terms of saturation. Previously pastel hues become much more saturated.

Because decorrelation stretching is based on principal component analysis, it is readily extended to any number of image channels. Recall that the IHS procedure is applied to only three channels at a time.

7.7 IMAGE CLASSIFICATION

The overall objective of image classification procedures is to automatically categorize all pixels in an image into land cover classes or themes. Normally, mul-

tispectral data are used to perform the classification and, indeed, the spectral pattern present within the data for each pixel is used as the numerical basis for categorization. That is, different feature types manifest different combinations of DNs based on their inherent spectral reflectance and emittance properties. In this light, a spectral "pattern" is not at all geometric in character. Rather, the term *pattern* refers to the set of radiance measurements obtained in the various wavelength bands for each pixel. *Spectral pattern recognition* refers to the family of classification procedures that utilizes this pixel-by-pixel spectral information as the basis for automated land cover classification.

Spatial pattern recognition involves the categorization of image pixels on the basis of their spatial relationship with pixels surrounding them. Spatial classifiers might consider such aspects as image texture, pixel proximity, feature size, shape, directionality, repetition, and context. These types of classifiers attempt to replicate the kind of spatial synthesis done by the human analyst during the visual interpretation process. Accordingly, they tend to be much more complex and computationally intensive than spectral pattern recognition procedures.

Temporal pattern recognition uses time as an aid in feature identification. In agricultural crop surveys, for example, distinct spectral and spatial changes during a growing season can permit discrimination on multidate imagery that would be impossible given any single date. For example, a field of winter wheat might be indistinguishable from bare soil when freshly seeded in the fall and spectrally similar to an alfalfa field in the spring. An interpretation of imagery from either date alone would be unsuccessful, regardless of the number of spectral bands. If data were analyzed from both dates, however, the winter wheat fields could be readily identified, since no other field cover would be bare in late fall and green in late spring.

As with the image restoration and enhancement techniques we have described, image classifiers may be used in combination in a hybrid mode. Also, there is no single "right" manner in which to approach an image classification problem. The particular approach one might take depends upon the nature of the data being analyzed, the computational resources available, and the intended application of the classified data.

In the remaining discussion we emphasize spectrally oriented classification procedures for land cover mapping. (As stated earlier, this emphasis is based on the relative state of the art of these procedures. They currently form the backbone of most multispectral classification activities.) First, we describe *supervised classification*. In this type of classification the image analyst "supervises" the pixel categorization process by specifying, to the computer algorithm, numerical descriptors of the various land cover types present in a scene. To do this, representative sample sites of known cover type, called *training areas*, are used to compile a numerical "interpretation key" that describes the spectral attributes for each feature type of interest. Each pixel in the data set is then compared numerically to each category in the interpretation key