6. Spectral Feature Extraction

1. Single image operations

When speaking of a single image, "spectral" feature extraction refers to limiting the number of gray values in the image. This is done by compressing the data in some manner (e.g., density slicing, thresholding, etc.) The result is a smaller feature space. The ultimate in simplicity is a binary image (1-bit). The binary image identifies only the presence or absence of a single feature (e.g., water, clouds).

2. Multiple-image feature extraction

When considering multiple image data, the 2-dimensional histogram is a basic tool. The 1-dimensional image histogram, $h(k)$, is a map of the frequency of occurrence of a gray value, $k$, in an image. For example, an image with the 1D histogram at right would have more pixels with a gray value $k=138$ than any other gray value.

Dividing the histogram by the number of pixels in the image yields a representation of the probability that a pixel picked at random from the image will have a particular gray value:

$$p(k) = \frac{h(k)}{N}$$

where $N$ is the number of pixels in the image and $p(k)$ is the called the probability density function (pdf).

A 2-dimensional histogram, $h(k_1, k_2)$, specifies the number of pixels that simultaneously have a gray value of $k_1$ in one image and a gray value of $k_2$ in a second, co-registered image. The 2-dimensional histogram is related to the joint probability density function, which describes the probability that a pixel selected at random will have a gray value of $k_1$ in image 1, and a value of $k_2$ in image 2:

$$p(k) = p(k_1, k_2) = \frac{h(k_1, k_2)}{N}$$

where $p(k_1, k_2)$ is the joint probability density function, and $k$ is a vector.

We may also define an n-dimensional histogram or pdf to describe the frequency of occurrence of a particular set of gray values in an n-channel image set, i.e.,

$$p(k_1, k_2, \ldots, k_n) = \frac{h(k_1, k_2, \ldots, k_n)}{N}$$

or, in vector notation,

$$p(k) = \frac{h(k)}{N}$$

where $k = k_1, k_2, \ldots, k_n$. 
Like the 1-dimensional histogram, a 2-dimensional histogram is devoid of any spatial information; it is strictly a statistical representation of the gray value distribution. The size of a 2–dimensional histogram is therefore independent of the size of an image. It is dependent only on the bit depth of the data.

**Feature extraction methods**

Consider an image having two target areas. The task is to discriminate between the two targets. Tonal variations within each target are of no concern.

a. **Select only those variables that contain useful information**

Suppose that two images of the same scene are available. In one, the targets of interest are separable by tone, in the other, the targets are only partially separable. In the illustration, this is apparent from the 1-d histograms. The first image is necessary and sufficient for discriminating between the two targets. Since the second image contributes nothing to significantly improve the discrimination, it may be discarded without loss for this application.

b. **Eliminate redundancy among the selected variables**

Again consider two images of the same scene. This time let the images contain essentially the same information, i.e., for the purpose of discriminating between $\omega_1$ and $\omega_2$ either image would be equally effective.
Algebraic Operations

Algebraic operations produce a single output image as a result of a pixel-by-pixel sum, difference, product or quotient of two or more input images.

Note: input images must be accurately registered.

Concepts: 2-Dimensional histogram

Image addition - tends to extract information that is positively correlated from image to image. Removes information that is uncorrelated from image to image. The output image gray value, \( k' \), is given by:

\[
k' = a_0 + a_1k_1 + a_2k_2 + \ldots + a_nk_n
\]

where:

\( k_1 = \) the gray value of a pixel in the \( i \)th image.

\( a_i = \) weighting coefficient for the \( i \)th image with \( a_i > 0 \)

Addition is often applied to data when the important information is correlated from band to band, as in the illustration. If the uncorrelated information is random noise, then the addition would tend to remove the random noise and thus improve the information content of the sum image. If the data are not noisy and image 1 and image 2 are different spectral bands, then the uncorrelated data is "color" information and the correlated data is intensity information. Image addition will then extract intensity information.

Ex 1: Removal of random noise: addition may be used to remove random noise. If \( k_1 \) is the gray value of an image contaminated with random noise, then we might write:

\[
k_1 = k + n_1
\]

signal + noise

Note that \( k_0 \) will always be positive while \( n_0 \) may be positive or negative. Similarly, for a second image of the same scene, we have:

\[
k_2 = k + n_2
\]

Taking the average value of the two images gives:

\[
k_{avg} = (k_1 + k_2)/2
\]

\[= k + (n_1 + n_2)/2\]

Generalizing to the case when \( t \) images of the same scene are available:

\[
k_{avg} = k + \frac{1}{t}(n_1 + n_2 + n_3 + \cdots + n_t)
\]

Since \( n_i \) may be positive or negative with a randomly varying magnitude,

\[
\frac{1}{t}(n_1 + n_2 + n_3 + \cdots + n_t) \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty
\]

and

\[
k_{avg} \rightarrow k \quad \text{as} \quad t \rightarrow \infty
\]

The noise reduction achieved is proportional to \( \sqrt{t} \)
Ex 2. Spectral Addition: Adding gray values from different spectral bands.

2-band case: A new image is defined with gray values $k'$, given by a linear combination of the two original bands with gray values $k_1$ and $k_2$.

$$k' = a_0 + a_1 k_1 + a_2 k_2$$

The vector, $a = (a_1, a_2)$, defines the direction of changing gray value that is desired in the new image. The coefficients may be defined by choosing

$$(k_{11}, k_{12}) \Rightarrow k_{\text{min}} = \text{minimum gray value in the new image},$$

and

$$(k_{21}, k_{22}) \Rightarrow k_{\text{max}} = \text{maximum gray value in the new image}.$$ 

The coefficients are given by:

$$a_2 = m a_1$$

where: $m$ = slope of the best fit straight line.

$$m = \frac{k_{22} - k_{21}}{k_{12} - k_{11}} \Rightarrow a_0 = k_{\text{min}} - a_1 (k_{11} + m k_{21})$$

$$a_1 = \frac{k_{\text{max}} - k_{\text{min}}}{(k_{21} - k_{11}) + m (k_{22} - k_{12})}$$

$$a_2 = m a_1$$

Image addition can be regarded as a partitioning of the measurement space. The space is partitioned into plane-parallel regions which are perpendicular to the direction of the coefficient vector $a = (a_1, a_2)$. Intensity information extraction will be optimized if the coefficient vector is parallel to the direction of correlation. In the illustration, the "optimal" partitioning is obtained when the coefficients are related by:

$$a_2 = ma_1$$

where

$$m = \frac{(k_{22} - k_{21})}{(k_{12} - k_{11})}$$
Image subtraction - tends to extract information that is uncorrelated (or negatively correlated) from image to image. The output image gray value, k', is given by:

\[ k' = a_1k_1 + a_2k_2 \]

where:  
- \( k_1, k_2 \) = gray value for image 1 and image 2, respectively
- \( a_1, a_2 \) = coefficient, with \( a_1 > 0 \) and \( a_2 < 0 \) (or vice versa)

Subtraction is often applied to data when the important information is uncorrelated from band to band, as in the illustration. For two images that represent two different spectral bands, uncorrelated information is the spectral (color) difference between the two images. If the two images are the same spectral band and were imaged at two different times, then the uncorrelated information represents the temporal difference between the two images.

Ex. 1: Motion detection/change detection
A car travelling on a road will appear in a different position in two images taken only moments apart. As long as the illumination conditions and exposure (or gain) settings of the imaging device have not changed for the two images, the gray values of corresponding pixels will be identical for all background areas. Only pixels representing the moving vehicle will be different in the two images.

Ex. 2. Spectral differencing
As with addition, the coefficients are given by:

\[ a_2 = m a_1 \]

where:  
- \( m = \frac{k_{22} - k_{21}}{k_{12} - k_{11}} \)
- \( a_0 = k_{\text{min}} - a_1 (k_{11} + mk_{21}) \)
- \( a_1 = \frac{k_{\text{max}} - k_{\text{min}}}{(k_{21} - k_{11}) + m(k_{22} - k_{12})} \)
- \( a_2 = m a_1 \)

As with addition, image subtraction measurement space is partitioned into plane-parallel regions which are perpendicular to the direction of the coefficient vector \( \mathbf{a} = (a_1, a_2) \). Color information extraction will be optimized if the coefficient vector is parallel to the direction of corresponding to the important color differences.
**Image division** (Ratioing) - tends to extract information that is uncorrelated from image to image. Ideally, ratioing adjusts for differences in intensity while emphasizing color differences. The output image gray value, $k'$, is given by:

$$k_0 = \frac{a_1 k_1}{k_2}$$

where: $k_1, k_2 =$ gray value for image 1 and image 2, respectively  
$a_1, a_2 =$ coefficients for the two image values.

Ratioing partitions measurement space into radial segments. The optimal partitioning will be dependent both on the position of the origin and on the scaling of the output gray value. A simple ratio will partition the space into radial segments that emanate from an origin at $(0,0)$. The implicit assumption is that $(0,0)$ represents black (zero intensity) and that distance from $(0,0)$ corresponds to an increase in intensity. In remote sensing imagery this is frequently not the case, largely due to atmospheric effects on the observed radiation. The true zero intensity point is usually offset in measurement space.

Scaling is also an issue. A simple ratio will result in a non-linear partitioning of the measurement space as illustrated in figure (a). Intuitively, one would like to partition the space each output gray value occupied an equal slice of measurement space.

Optimal scaling of a ratio image is not always easily accomplished, however, if the appropriate scaling values can be determined, a reasonable scaling is given by:

$$k = \frac{n_f - n_0 + 1}{\theta_{\max} - \theta_{\min}} \left[ \tan^{-1} \left( \frac{k_2 - b_2}{k_1 - b_1} \right) - \theta_{\min} \right]$$

where:  
$\theta_{\min} =$ minimum angle  
$\theta_{\max} =$ maximum angle  
$n_f =$ maximum gray value in output image (usually 255)  
$n_0 =$ minimum gray value in output image (usually 0)  
b$_1, b_0 =$ offsets of the zero-intensity point in measurement space.
**Principal Component Analysis** (Richards: p. 127 -138)
- characteristic vector analysis
- eigenvector analysis
- Karhunen-Loeve transformation

Multispectral images often exhibit high correlations between spectral bands, i.e., there is considerable redundancy in the information content from band to band. Principal component analysis is a procedure that determines a linear transformation of the spectral data into a new set of variables.

**Vegetation Indices**

Vegetation Indices are feature extraction operations designed to yield estimates of vegetative cover in an image. These indices are based on the fact that vegetation absorbs well in the red (visible) and reflects very efficiently in the near infrared. We will consider four vegetation indices:

a. Environmental Vegetation Index \[ \text{EVI} = \frac{\text{IR}}{\text{Red}} \]
b. Normalized Difference Vegetation Index \[ \text{NDVI} = \frac{\text{IR} - \text{R}}{\text{IR} + \text{R}} \]
c. Path-corrected Normalized Vegetation Index \[ \text{NVI}^* \]
d. Bounded Vegetation Index \[ \text{BVI} \]
e. Soil-Adjusted Vegetation Index \[ \text{SAVI} \]
f. Kauth-Thomas transformation \[ \text{Tasseled cap, K-T} \]

The EVI, NDVI, NVI, BVI and SAVI belong to a group of indices that partition spectral space by some form of ratio, wherein the vegetation index of a pixel is defined by the slope of the line passing through its value in spectral space. Any of these indices can be expressed in terms the ratio of two new variables, \( x \) and \( y \), obtained by a rotation of the spectral axes by some angle \( \theta \) with or without a change in the location of the origin:

\[
x = (R - R_0) \cos(\theta) + (IR - IR_0) \sin(\theta)
\]

\[
y = -(R - R_0) \sin(\theta) + (IR - IR_0) \cos(\theta)
\]

where \( R_0 \) and \( IR_0 \) are the offsets to the new coordinate system origin. The simplest vegetation index is the Environmental Vegetation Index (EVI) which is simply the ratio \( IR/\text{Red} \) with \( \theta = 0^\circ \) and \( IR_0 = 0^\circ \):
In this case, the index represents the slope of lines passing through (0,0), and ranges from zero for a pixel lying on the red axis, to infinity for a pixel on the IR axis.

The soil line usually lies near the diagonal in ir-red spectral space. Since vegetation lies to the ir side of the soil line, a better scaling of the index is provided by a rotation of the axes by 45°. This rotation, which requires no prior knowledge of atmospheric conditions and is mathematically convenient, yields the Normalized Vegetation Index (NVI):

\[
\text{NVI} = \frac{y}{x} = \frac{-(R - 0)\sin(45^\circ) + (IR - 0)\cos(45^\circ)}{(R - 0)\cos(45^\circ) + (IR - 0)\sin(45^\circ)} = \frac{IR - R}{IR + R}
\]

In the NVI coordinate system, the x-axis lies at an angle of 45° to the red axis, and points on the IR axis have a value of 1. Thus, the range of the NVI index is from 0 to 1.

Path radiance is an additive factor to the radiance observed at the remote sensor. The effect of path radiance is to shift the entire data structure away from the origin in spectral space. To accommodate this shift, the index coordinate system can be translated by an amount equivalent to the path radiance in each band. A modified normalized vegetation index, NVI*, is defined by setting the red and ir offsets, \( R_0 \) and \( IR_0 \), equal to the band path radiances, \( R^* \) and \( IR^* \), respectively:

\[
\text{NVI}^* = \frac{y}{x} = \frac{-(R - R^*)\sin(45^\circ) + (IR - IR^*)\cos(45^\circ)}{(R - R^*)\cos(45^\circ) + (IR - IR^*)\sin(45^\circ)} = \frac{(IR - IR^*) - (R - R^*)}{(IR - IR^*) + (R - R^*)}
\]
The origin of the NVI* coordinate system represents the location of a black (zero reflectance) target. Thus, angular change in position relative to the new origin represents a change in color but not in intensity. If a change in color in the ir-red space can be equated to a change in the amount of vegetation, then the angular change represented by the NVI* should be an effective measure of vegetative cover. (This is an implicit premise of the NVI and the simple ratio index. But without accounting for path radiance, the relationship between angle and color change is invalid, and the prediction of vegetation change is weakened.)

With the NVI and NVI*, vegetation is estimated by the angular distance in ir-red spectral space from the 45° line, which is presumed to be coincident with the soil line. However, the soil line can be skewed due to multiplicative atmospheric effects. The BVI is designed to account for these effects by using the true soil line as a reference. If the angle of the true soil line relative to 45° is $\theta_s$, and the angle of a vegetation pixel relative to 45° is $\theta_v$, then the angle of the vegetation pixel from the soil line is simply $\theta_v - \theta_s$. As the NVI* is also expressed relative to 45°, $\theta_v$ is equal to $\tan^{-1}(\text{NVI}*)$. The angle of the vegetation pixel relative to the soil line is, therefore:

$$\theta_v = \tan^{-1}(\text{NVI}*) - \theta_s.$$

The maximum possible vegetation angle, $\theta_v$, is the angle of a line parallel to the ir axis, and the distance of this line from the ir axis is determined by the red path radiance, $R_o$. No pixel has a lower red value than this. In the NVI* coordinate system this line corresponds to an index value of 1, and hence an angle of 45° relative to the diagonal. Thus, the angular distance of a vegetated pixel from the soil line, $\theta_v = \tan^{-1}(\text{NVI}*) - \theta_s$, can be expressed as a proportion of the maximum possible angular distance of a vegetated pixel, 45°-$\theta_s$, giving the BVI:

$$\text{BVI} = \frac{\tan^{-1}(\text{NVI}*) - \theta_s}{45^\circ - \theta_s}$$

In summary, with the BVI, the soil-vegetation structure is bounded by the soil line and the maximum vegetation line. The locations of these lines are defined in deriving the BVI, and the calculated index of any pixel is expressed as a fraction of the maximum possible vegetation. This means that the BVI's from different scenes can be compared reliably. In contrast, with the EVI and NVI, the maximum vegetation value in a particular scene is unknown; variation in this value is an unpredictable and important source of error when comparing calculated indices.

The SAVI is another variation on this theme (Reference). In contrast, the Kauth-Thomas transformation is a significant departure from the other vegetation indices. (See Richards, p. 148-152.)