RELATIVE ELEVATION DETERMINATION FROM LANDSAT IMAGERY

S. WANG[1], R.M. HARALICK[2] and J. CAMPBELL[3]

[1] *Department of Computer Science, George Mason University, Fairfax, VA 22030 (U.S.A.)*
[2] *Machine Vision International, Ann Arbor, MI 48104 (U.S.A.)*
[3] *Dept. of Geography, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061 (U.S.A.)*
(Received March 29, 1984)

ABSTRACT

Wang, S., Haralick, R.M. and Campbell, J., 1984. Relative elevation determination from LANDSAT imagery. *Photogrammetria*, 39: 193—215.

In LANDSAT imagery, spectral and spatial information can be used to estimate a relative digital terrain model in mountainous areas. To do this, the mixed information of direct and indirect illumination, material reflectance, and topographic modulation in the original LANDSAT imagery must be first separated. From the direct and indirect illumination information, ridges and valleys can be determined. From the material reflectance information, big visible rivers can be detected. Finally, a relative elevation model can be generated by elevation growing. In elevation growing valley pixels are assigned increasing elevations as they become more distant from the rivers or other valley pixels already assigned an elevation. It also proceeds in a direction perpendicular to valleys climbing up to the ridges assigning elevations to any unassigned pixel.

1. INTRODUCTION

It is a common task for a photointerpreter to examine the spatial pattern on an aerial image and by appropriate interpretation be able to tell the elevation of one area relative to another and be able to infer the stream network and the drainage network even though some of the streams may be below the resolution of the sensor. There is a wealth of information in spatial patterns on aerial imagery but most computer data processing of remotely sensed imagery, being limited to pixel spectral characteristics, does not make use of it.

In this paper, we describe a procedure by which the relative elevation model can be inferred from a LANDSAT scene of mountainous and hilly terrain. To a first-order effect, the cause of the intensity value at any pixel is due to whether it is only diffusely lit or directly lit. If it is directly illuminated there are additional effects due to the angle at which the sun illuminates the ground patch corresponding to the pixel and the reflectance of the surface material on the ground patch. To make sense of the spatial pattern first requires separation of these effects. For this purpose, we use a clustering

technique on ratio images to determine similar reflectance classes and then do a subclustering on these classes to determine directly lit from indirectly lit pixels. This subclustering creates a shadow image. Then we modify Eliason et al.'s (1981) technique to create two images from the one LANDSAT image (Haralick and Wang, 1983). The first image is a "reflectance" image; the second is a topographic modulation image portraying information related to surface slope and sun illumination. The details of this technique are given in Section 2.

As discussed in Section 3, the sun azimuth and the shadow image constitute sufficient information for the identification of the ridges and the valleys. With the valleys identified, each valley pixel may be assigned a relative elevation which increases as the valley path from the pixel to the river it empties in increases. Ridges must be assigned elevations higher than their neighboring valleys and each ridge pixel can be assigned a relative elevation which decreases on the ridge path from the pixel to the saddle point where the ridge crosses a valley. In order to do this, local slopes must be known. Some estimated local slopes are assigned initially to generate the first version of elevation model. Then the topographic modulation image is used to calculate more accurate local slopes to generate better elevation models. Finally, LANDSAT imagery is reconstructed to evaluate our illumination model and elevation reconstruction algorithm.

1.1. Study area

This research examines an area in southeastern West Virginia, shown in Fig. 1.1. This region is a portion of the Appalachian Plateau's physiographic province, within the "unglaciated Allegheny plateau" described by Thornbury (1967). In general, this region is a thoroughly dissected plateau-like surface. It receives about 1 m of precipitation each year and, as depicted on topographic maps, has a moderate drainage network density. Drainage is through tributaries of the New (Kanawha) River, which flows west into the Ohio River drainage system.

The overall drainage pattern within this region is that of a relatively large sinuous channel (the Gauley River) superimposed over the finer texture of a dentritic pattern formed by first, second, and third order streams. A number of the small first or second order streams flow directly into the large channel. Thus the overall pattern is composed of a mixture of many very small stream segments, many with very steep gradients, a prominent major channel with a relatively low gradient, and relatively few stream segments of intermediate length and gradient.

Throughout the area, flood plains (when present) are narrow and tend to closely follow the course of the stream channel. Valleys are narrow, with steep sides; the Gauley River, for example, follows a valley that is typically 150 m deep but only 100 m wide. Uplands often consist only of ridge crests, although plateau-like upland regions are present, they are not continuous or

extensive. The area is forested with a dense cover of deciduous trees (Kuchler's "mixed mesophytic forest", 1964). Cleared areas for agriculture (chiefly pasture) tend to follow the valleys of intermediate-sized streams. Settlements are small and dispersed, usually positioned in valleys.
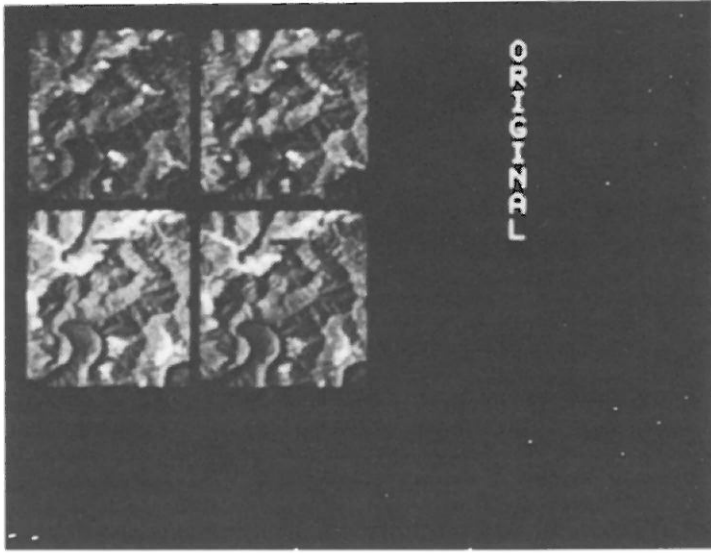

Fig. 1.1. Original LANDSAT subscene of 80 pixels by 80 pixels.

This region appears on the Charleston, West Virginia/Ohio USGS 1:250,000 quadrangle (NJ 17-5). Our investigations include areas in Nicholas County, W. VA and neighboring counties. This area was imaged by the LANDSAT-1 MSS on April 13, 1976 (scene id: 5360-14502; path 18, row 34). This date reflects important qualities of the scene. First, at this date the atmosphere was unusually clear — there is no evidence of atmospheric (Mie) scattering or degradation of the data. Also, at this spring date most of the forested areas are without leaves, especially at higher elevations. Lower elevations have a cover of newly emerged leaves and grasses. Within a few weeks leaves will have emerged in vegetation throughout the entire region, but at this time in April, there is a sharp spectral contrast between the vegetation cover of the higher elevations and that of some of the valleys.

2. THE PROBLEM OF MIXED INFORMATION

Four kinds of information are mixed in LANDSAT imagery: surface reflectance, topography, diffuse light and haze. Assuming the ground surface is flat, vegetated areas have high reflectance for some spectral regions and appear as bright areas to the LANDSAT sensor. On the contrary, water areas have low reflectance and appear as dark areas to the LANDSAT sensor. If

topography is then considered, there is a pattern of directly illuminated and shadowed slopes due to varied heights and orientations of the slopes. However, graytones for image pixels corresponding to shadowed locations are not zero because of diffuse light coming indirectly from the sun. Finally, when light is reflected from the ground back to the sensor, there is additive haze due to atmospheric scattering. The difficulty of interpreting LANDSAT scenes of mountainous areas is due to the mixing of topographic data with reflectance data. To begin to separate these individual components we need to begin with an illumination model.

2.1. Separating the information

The basic data model for a Lambertian surface illuminated by a point source is:

$$G(x, y) = r(x, y) I \cos \theta (x, y) \qquad (2.1)$$

where G is brightness value of a pixel within the image,

  x, y are pixel coordinates,
  r is surface reflectance,
  I is the illumination flux from the sun, and
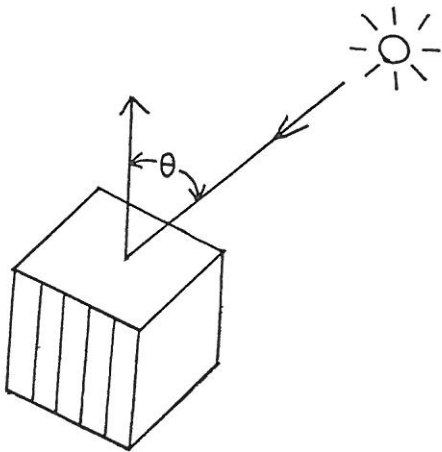 θ is the angle between sun incidence direction and surface normal (Fig. 2.1).

Fig. 2.1. θ is the angle between sun incidence direction and surface normal.

Adding band number, diffuse light, and haze into this model, one has the general model for LANDSAT data as:
  (I) For directly illuminated pixels:

$$G(x, y, b) = r(x, y, b) I(b) \cos \theta (x, y) + r(x, y, b)D(b) + H(b)$$

(II) For shadowed pixels:

$$G(x, y, b) = r(x, y, b) D(b) + H(b) \qquad (2.2)$$

where b is the spectral band number,

D is diffuse light, and

H is the haze due to atmospheric scattering.

Because haze is an additive constant independent of pixel locations, we use Switzer et al.'s (1981) technique for haze removal. The haze corrected image $G - H$ is defined as G'.

After haze is removed, it can be seen that resolution of the remaining components amounts to extracting diffuse light Df containing the information of $r(x, y, b)D(b)$, reflectance data R which contains the information of $r(x, y, b)I(b)$, and topographic modulation data Tp which contains the information of $\cos\theta(x, y)$:

(I) For directly illuminated pixels:

$$G'(x, y, b) = R(x, y, b) Tp(x, y) + Df(x, y, b)$$

(II) For shadowed pixels:

$$G'(x, y, b) = Df(x, y, b) \qquad (2.3)$$

The first problem to be solved to accomplish this unmixing is the determination of which pixels are directly lit from which pixels are in shadow. Once this is accomplished the unraveling can begin. For example, for the diffuse light image, pixels which are in shadow take their value as the dehazed data value. Pixels which are directly lit take their value as the average dehazed data value taken over all shadowed pixels which are likely to be from the same material as they are.

To separate the shadow pixels from the directly lit pixels, we seek to transform the images in a way in which the only effect is reflectance. Then within groups of pixels with similar reflectance, we can separate the bright appearing ones from the dark appearing ones. This two step technique is more accurate than a simple thresholding technique (Campbell et al., 1981; Wang et al., 1983).

One way to transform the data so that the only remaining effect is reflectance is to take ratios of one band to another. The ratio image has been widely used by remote sensing researchers to subdue surface topographic effects (Vincent, 1973; Raines et al., 1978). An alternative rationing procedure is to calculate a ratio of each pixel value in band to the total brightness for that pixel, summed over all bands, as suggested by Mulder (1982). For our procedure either approach is computationally feasible, provided the denominator in the ratio is composed of a linear combination of values. For this study, we prefer to use the ratios of individual pairs of bands, as ratios of band pairs are known to be effective in distinguishing reflectance of surface materials (Eliason et al., 1981). From eq. 2.2, the ratio image of two bands with band number b_1 and b_2 for directly illuminated pixels after haze is removed is:

$$\frac{G'(x, y, b_1)}{G'(x, y, b_2)} = \frac{r(x, y, b_1) \ [I(b_1) \cos\theta (x, y) + D(b_1)]}{r(x, y, b_2) \ [I(b_2) \cos\theta (x, y) + D(b_2)]}$$

If one assumes illumination and diffuse light, bands b_1 and b_2 are related by:

$$I(b_1) = a \ I(b_2),$$

$$D(b_1) = a \ D(b_2), \text{ then:}$$

$$\frac{G'(x, y, b_1)}{G'(x, y, b_2)} = \frac{r(x, y, b_1) \ a \ [I(b_2) \cos\theta (x, y) + D(b_2)]}{r(x, y, b_2) \ [I(b_2) \cos\theta (x, y) + D(b_2)]} = a \frac{r(x, y, b_1)}{r(x, y, b_2)} \qquad (2.4.1)$$

Similarly, the ratio image for shadowed pixels is:

$$\frac{G'(x, y, b_1)}{G'(x, y, b_2)} = a \frac{r(x, y, b_1)}{r(x, y, b_2)} \qquad (2.4.2)$$

Thus, whether shadowed or directly illuminated, the ratio image is independent of $\cos\theta$. Three independent ratio images taken from the 4-band imagery in Fig. 1.1 are shown in Fig. 2.2. It can be clearly seen that the effects of shadows have been removed.
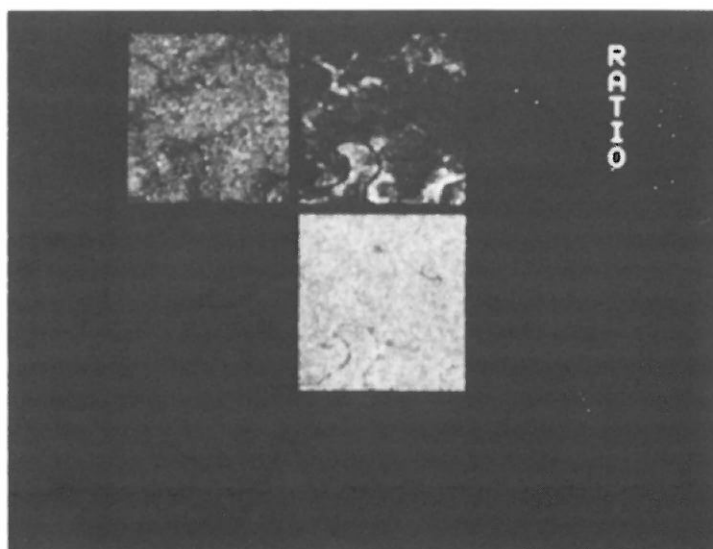

Fig. 2.2. Ratio images 5/4, 6/5, 7/6.

2.2. Clustering

Because the three ratio images depend upon material reflectance only, regions of the same material reflectance can be identified by grouping together pixels of similar spectral characteristics. Because we desire to conduct

the analysis using a minimum of prior information, unsupervised classification is favored over the supervised approach, which would require detail knowledge of the number, identify, and characteristics of groups. Unsupervised classification permits identification of the natural structure of the image with a minimum of prior information.

In the noisy ratio images of Fig. 2.2, there are three major clusters: water area, vegetated area, and non-vegetated area. The size of water area is much smaller than that of the other two. For this reason, the mode approach (Goldberg and Shlien, 1978) which uses a fixed threshold to get cluster centers does not work. AMOEBA (Bryant, 1979) works better, but fails to obtain unbroken river segments. Despite its simplicity, it was found that ISODATA modified in such a way that class sizes are also taken into consideration works best.

The algorithm of basic ISODATA is given in the textbook by Duda and Hart (1973):

(1) Choose some initial values for the means u_1, u_2, ..., u_c, where c is the total number of classes.

(2) Classify the samples by assigning them to the class of the closest mean.

(3) Recompute the means as the average of the samples in their class.

(4) If any mean changed value, go to 2; otherwise, stop.

The modified ISODATA algorithm is 1. Choose some initial values for the means u_1, ..., u_c, and sizes z_1, ..., z_c, where c is the total number of classes.

Set ACTIVE (i) = 1, $1 <= i <= c$. 2. Classify the samples by assigning them to the class of closest mean. 3. Recompute the means as the average of the samples in their class. Also compute the size of each class. 4. Set ACTIVE (i) = 0 if size of class i is greater than zi. If no mean changed value or ACTIVE (i) = 0 for all i, stop. 5. For samples which are in class i with ACTIVE (i) = 1, classify them by assigning them to the class j of the closest mean, ACTIVE (j) = 1.

Go to 3.

Now the crucial point is how to select the initial mean and size for each class. One simple yet powerful method is developed by the inspiration of looking at the pseudocolor image (Moik, 1980) which is created by assigning a color to a pixel (x, y) according to the quantized value of $G(x, y, b_1)$, the quantized value of $G(x, y, b_2)$, and the quantized value of $G(x, y, b_3)$ for a three-band graytone imagery. The result even can be considered to be a reasonable cluster image. Thus, the following recursive algorithm was developed to compute the total number of classes as well as initial class means and sizes.

(1) For an n-band imagery, find the minimum graytone value MIN and maximum graytone value MAX for each band i and quantisize the graytones $G(x, y, b_i)$ into $g + 1$ values:

$$Q(1, b_i) = \text{MIN}, ..., Q(g + 1, b_i) = \text{MAX}$$

Calculate the total number of pixels in sets:

$$S(a_1, b_1) = \{(x, y) \mid Q(a_1, b_1) <= G(x, y, b_1) <= Q(a_1 + 1, b_1), \text{ and}$$
$$Q(a_2, b_2) <= G(x, y, b_2) <= Q(a_2 + 1, b_2)\}$$

Also calculate the graytone means for each set.

Mark all the sets as active.

(2) Find the set with highest number of pixels among the active sets. This is a cluster center. Its mean can be used as initial class mean in ISODATA. An estimated size of this class is the sum of its number of pixels and the numbers of pixels of its neighbors.

(3) If no peak can be found, stop; otherwise mark this peak set and its neighbors as inactive and go to 2. Using these initial class means and sizes in the modified ISODATA, one gets material cluster image mc of Fig. 2.3.

Each material cluster c_1 can be defined as a set of pixels (x, y) in which $Mc(x, y) = c_1$:
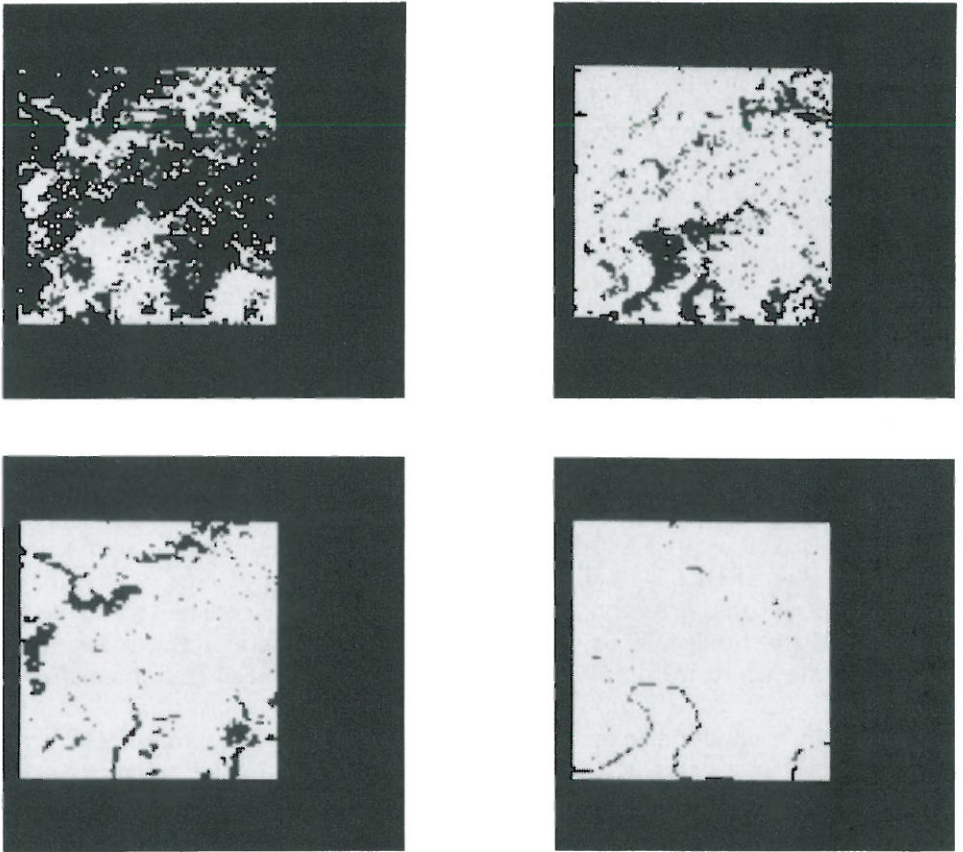
$$C(c_1) = \{(x, y) \mid Mc(x, y) = c_1\}$$

Fig. 2.3. Material cluster image. For each cluster, a binary image is shown.

Once the material clusters are defined on the basis of the ratio images, one can find directly illuminated and shadowed pixels and define a binary shadow image. To do this, we collect together all dehazed 4-band pixel values belonging to a single material cluster and subcluster these 4-tuples into dark and bright subcluster classes. The next few paragraphs describe this in detail.

If one overlays the material cluster image Mc over any band of the dehazed image, one can see, within each material cluster, some pixels are bright and the others are dark. These differences are due to topographic variations; the bright pixels are directly illuminated pixels, and the dark pixels are in shadow. To separate the shadow pixels from the directly lit pixel, for each material cluster c_1, one performs a subclustering on the dehazed pixel values in the set:

$$\{G'(x, y, b) \mid Mc(x, y) = c_1\}$$

which is the set of all dehazed values for pixels whose material cluster index is c_1. This subclustering on cluster c_1 separates the directly illuminated pixels $C_0(c_1)$ from the indirectly illuminated pixels $C_1(c_1)$:

$$C_0(c_1) = \{(x, y) \mid (x, y) \text{ is directly illuminated on the basis of the subclustering}\}$$

$$C_1(c_1) = \{(x, y) \mid (x, y) \text{ is indirectly illuminated on the basis of the subclustering}\}$$

The subclustering of getting C_0, C_1 uses only the basic ISODATA program. In this case, the initial class mean for C_1 includes all the minimum graytones
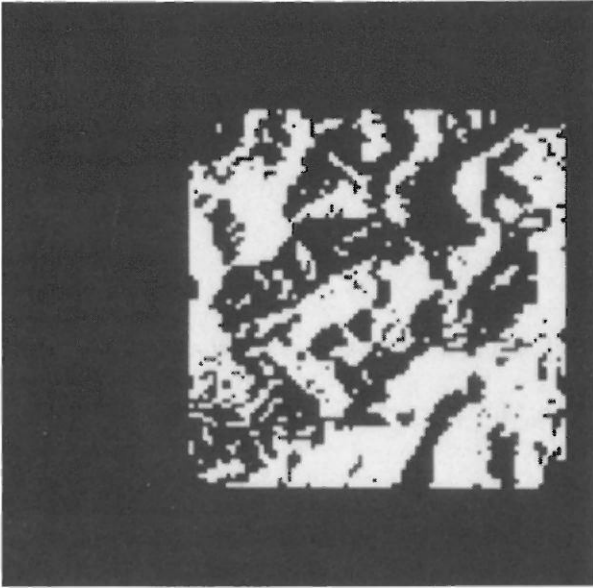

Fig. 2.4. Binary shadow image.

for four bands, and the initial mean for C_0 includes all the maximum gray-tones for four bands.

A shadow image Sw can be defined as:

$$Sw : X \times Y \to \{0, 1\}$$

$$Sw : (x, y) = 0 \text{ if } (x, y) \in C_0(Mc(x, y))$$

$$\qquad\qquad 1 \text{ if } (x, y) \in C_1(Mc(x, y)) \tag{2.5}$$

The shadow image for Fig. 1.1 is shown in Fig. 2.4. The correspondence between this and a topographic map is quite good. Now that directly lit and shadowed pixels have been identified, it is possible to use the dehazed image to get the diffuse light image Df, the reflectance image R, and the topographic modulation image Tp.

2.3. Diffuse light data, reflectance data, and topographic modulation data

From eq. 2.3, shadowed pixels contain only the information of diffuse light, but directly illuminated pixels contain the information of both diffuse light and direct sun illumination. By the method of clustering, each pixel in the image belongs to a material cluster. Each material cluster has a bright and a dark subcluster. For pixels in the dark subcluster, one can simply use their dehazed values as their values in the diffuse light image Df. On the other hand, for pixels in the bright subcluster, one can define their value in the diffuse light image to be the average value of all the pixels from the dark subcluster associated with the material cluster to which these pixels belong.

With this definition, the diffuse light image Df is:
(I) For directly illuminated pixels:

$$Df(x, y, b) = \sum_{\substack{(u, v) \in C_1(c_1) \\ c_1 = Mc(x, y)}} \frac{G'(u, v, b)}{\#C_1(c_1)}$$

where, for a set S, $\#S$ means the size of this set.
(II) For shadowed pixels:

$$Df(x, y, b) = G'(x, y, b)$$

If there were no variations in reflectance for pixels from the same material, we have:

Assumption 1: $r(x, y, b)$ is a constant $r'(c_1, b)$ for all pixels (x, y) in $C(c_1)$, where $c_1 = Mc(x, y)$.

Under this assumption, for directly illuminated pixels:

$$Df(x, y, b) = r'(c_1, b)D(b) \sum_{(u, v) \in C_1(c_1)} \frac{1}{\#C_1(c_1)}$$

$$\qquad\qquad = r'(c_1, b)\, D(b) \tag{2.6}$$

From eqs. 2.2 and 2.6 and Assumption 1, $G' - Df$ is:

(I) For directly illuminated pixels:

$$G'(x, y, b) - Df(x, y, b) = r(x, y, b) I(b) \cos\theta (x, y) + r(x, y, b) D(b)$$

$$- r'(c_1, b) D(b)$$

$$= r(x, y, b) I(b) \cos\theta (x, y) \qquad\qquad [c_1 = Mc(x, y)]$$

(II) For shadowed pixels:

$$G'(x, y, b) - Df(x, y, b) = 0 \qquad\qquad (2.7)$$

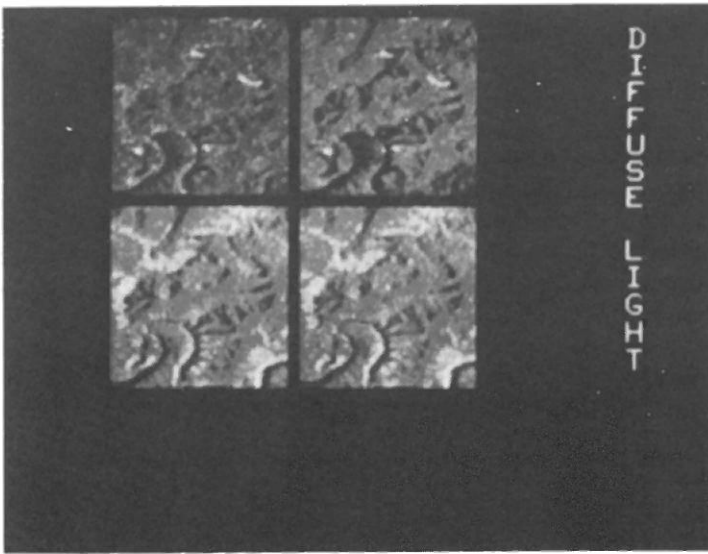The Df image for Fig. 1.1 is shown in Fig. 2.5.

Fig. 2.5. Diffuse light image.

An initial or raw estimated reflectance image R' can be calculated by assigning each pixel's value to be the average $G' - Df$ value of all the pixels from the bright subcluster associated with the material cluster to which the pixel belongs.

$$R'(x, y, b) = \sum_{\substack{(u, v) \in C_0(c_1) \\ c_1 = Mc(x, y)}} \frac{G'(u, v, b) - Df(u, v, b)}{\#C_0(c_1)}$$

$$= \sum_{(u, v) \in C_0(c_1)} \frac{r(u, v, b) I(b) \cos\theta (u, v)}{\#C_0(c_1)}$$

$$= r'(c_1, b) I(b) \sum_{(u, v) \in C_0(c_1)} \frac{\cos\theta (u, v)}{\#C_0(c_1)}$$

Define: $Xc(c_1) = \sum\limits_{(u,\,v) \in C_o(c_1)} \dfrac{\cos\theta\,(u,\,v)}{\#C_o(c_1)}$

$Xc(c_1)$ is the spatial average of $\cos\theta$ for pixels in the bright subcluster. Then:

$$R'(x,\,y,\,b) = r'(c_1,\,b)\,I(b)\,Xc(c_1) \tag{2.8}$$

We use the raw estimated reflectance images to produce raw estimates of the topographic modulation image. By eqs. 2.3, 2.7 and 2.8, the raw estimated topographic modulation image Tp' for band b is:

(I) For directly illuminated pixels:

$$Tp'(x,\,y,\,b) = \frac{G'(x,\,y,\,b) - Df(x,\,y,\,b)}{R'(x,\,y,\,b)} = \frac{r(x,\,y,\,b)\,I(b)\,\cos\theta\,(x,\,y)}{r'(c_1,\,b)\,I(b)\,Xc(c_1)} =$$

$$= \frac{\cos\theta\,(x,\,y)}{Xc(c_1)} \qquad\qquad [c_1 = Mc(x,\,y)]$$

(II) For shadowed pixels:

$$Tp'(x,\,y,\,b) = 0 \tag{2.9}$$

We use the following assumption to convert the raw topographic images to our final estimated topographic modulation image.

Assumption 2: $Xc(c_1)$ is the same for all material clusters, $1 <= c_1 <= Nc$.

With this assumption, the principal component image generated from Tp' corresponding to the largest eigenvalue becomes our final single-band estimated Tp image.

At this point, we have a four band image Tp each band of which is estimate of $\cos\theta\,(x,\,y)$. The first principal component of this four band image provides an estimate which is close to the minimum variance estimate of Tp.

(I) For directly illuminated pixels:

$$Tp(x,\,y) = k\,\cos\theta\,(x,\,y) \text{ for a constant } k \tag{2.10}$$

(II) For shadowed pixels:

$$Tp(x,\,y) = 0$$

Having the topographic modulation image, the final estimated reflectance image is easily computed using eq. 2.3.

(I) For directly illuminated pixels:

$$R(x,\,y,\,b) = \frac{G'(x,\,y,\,b) - Df(x,\,y,\,b)}{Tp(x,\,y)} = \frac{r(x,\,y,\,b)\,I(b)\,\cos\theta\,(x,\,y)}{k\,\cos\theta\,(x,\,y)} =$$

$$= \frac{r(x,\,y,\,b)I(b)}{k} \tag{2.11}$$

(II) For shadowed pixels:

$$R(x, y, b) = \sum_{\substack{(u, v) \in C_0(c_1) \\ c_1 = Mc(x, y)}} \frac{R(u, v, b)}{\#C_0(c_1)} = \frac{I(b)}{k} \sum_{(u, v) \in C_0(c_1)} \frac{r(u, v, b)}{\#C_0(c_1)}$$

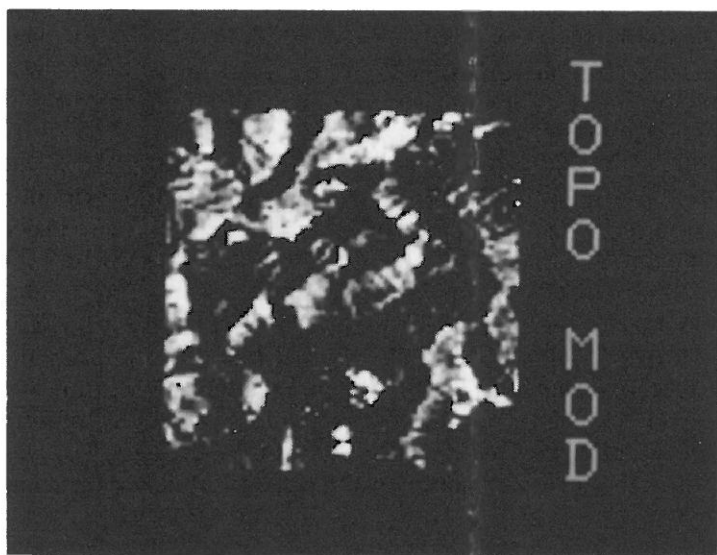The Tp image is shown in Fig. 2.6, and R image is shown in Fig. 2.7.
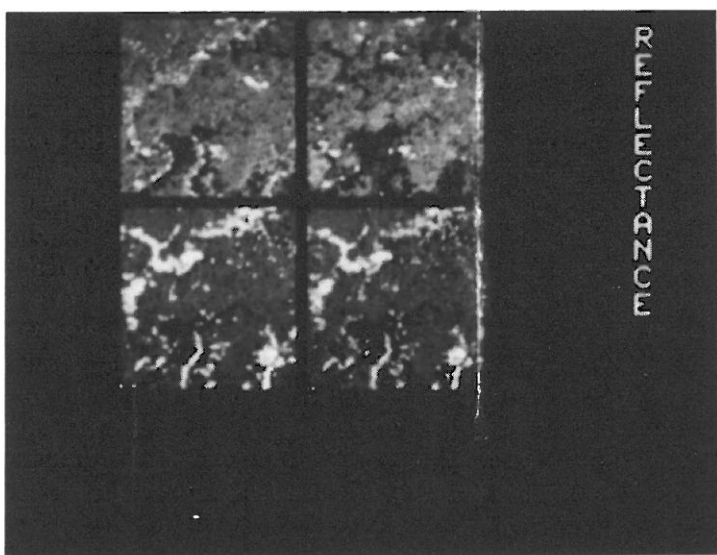

Fig. 2.6. Topographic modulation image.

Fig. 2.7. Reflectance image.

3. ELEVATION ESTIMATION

3.1. Identification of ridges and valleys

In the last section, the problem of confounded data is handled in such a way that the material information is contained in the reflectance image and the diffuse light image. The topographic information is contained in the shadow image and the topographic modulation image. In this section, we will show how to detect ridge segments, valley segments, and peak junctions from the shadow image. In the next section, we will perform an elevation growing to obtain initial raw estimates of elevations for all pixels on the basis of these ridge and valley segments. We then obtain a refined estimate by making the elevation model have slopes consistent with the information on the topographic modulation image.

Sides of hillsides facing the sun must be directly lit. Sides of hillsides facing away from the sun must be indirectly lit. A directly lit to indirectly lit transition in a direction moving away from the sun is a ridge. An indirectly lit to directly lit transition in a direction moving away from the sun is a valley. Thus, valleys and ridges exist on the borders between shadowed and directly lit areas. To find these areas we use the binary shadow image. First, a connected components operation determines regions on the shadow image. Then small, noisy regions are eliminated.

Next, the perimeters of these bright and shadowed regions are segmented into border segments according to their left regions, right regions, and orientations. A border segment is a maximally long sequence of connected pixels which are on the border between two given regions. Because the detection of ridges and valleys is highly orientation-dependent and the sun illumination comes from east in Fig. 1.1, each border segment is further broken into several pieces according to orientation; all the east—west parts are separated from the north—south parts. The binary result is shown in Fig. 3.1.

As the sun illumination is from the east in LANDSAT imagery, those border segments which are valley segments or ridge segments can be identified according to the brightness of the regions adjacent on the left and on the right. Because most of the trees in this area in April are unfoliated, the strongest region boundaries are shadow boundaries rather than reflectance boundaries, and the strongest boundaries are those at the extremes of steep slopes oriented normal to the sun direction. Because the sun illumination is predominantly east—west, a boundary that is dark on the left and bright on the right will correspond to a ridge, and the reverse will correspond to a valley.

East—west region boundaries are classified according to the labeling of neighboring north—south boundaries as well as their orientation relative to the east—west boundaries. As shown in Fig. 3.2, each east—west boundary B_1 has a left intersecting north—south boundary B_2 and a right intersecting
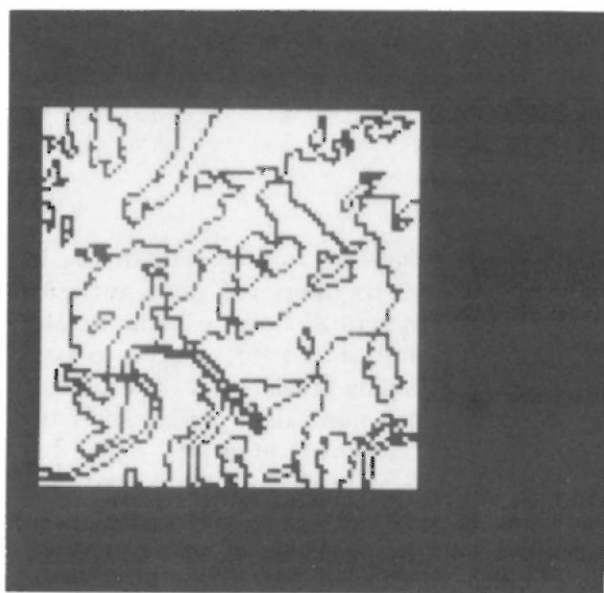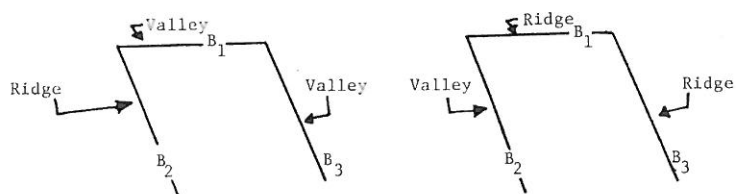
Fig. 3.1. Border segments.

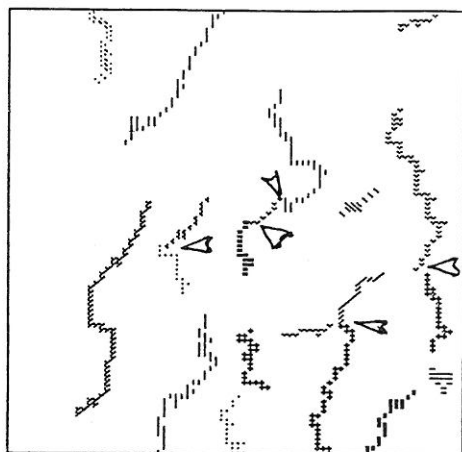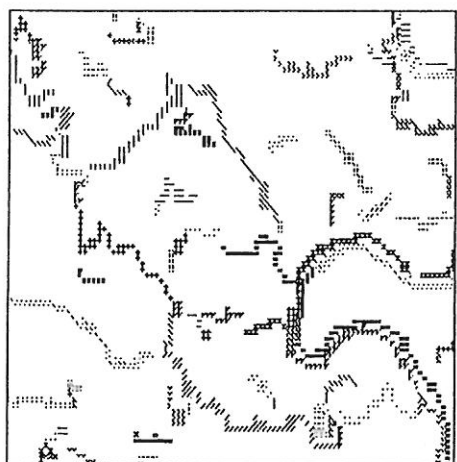Fig. 3.2. Classifying east—west border segments.

Fig. 3.3a. Valley map consisting of the border segments which are identified as valleys.
Fig. 3.3b. Ridge map consisting of the border segments which are identified as ridges.

north—south boundary B_3. If the angle between B_1 and B_2 is smaller than the angle between B_1 to B_3, then we assign the labeling of boundary B_2 to B_1; otherwise, we assign the labeling of boundary B_3 to B_1. The results of ridge-valley finding are shown in Fig. 3.3.

3.2. Elevation growing

The detection of the ridge and valley segments as discussed in the last section only assigns a ridge or valley label to them and does not assign relative elevations to them. This section describes how to estimate their relative elevations. First, a model called elevation growing is used to assign initial estimated elevations for all ridge and valley pixels. Next, interpolation is used to assign elevations for non-ridge and non-valley pixels. After this, the topographic modulation image is used to improve the result.

The cross-sections of valleys are V-shaped, and the cross-sections of ridges are A-shaped. If one looks at topographic maps, the elevation contours of valleys such as those shown in Fig. 3.4 can be frequently found. Thus, if one draws a line ab perpendicular to the valley Va, the elevations are increasing from point 0 to point a, and also from point 0 to point b. If the end point of a valley segment of smaller order is encountered during the growing, it is deduced that this end point is the lower end of this smaller valley segment. However, if a ridge point is encountered during the process, the increasing has to stop because the elevation starts to decrease. Based on this knowledge, an "elevation growing" model can be created.
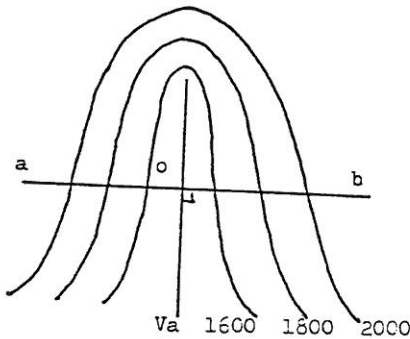

Fig. 3.4. The elevation pattern of valleys and its relation to elevation growing.

Three different local slopes are assigned to three classes of non-ridge and non-valley pixels: pixels which are close to ridges, pixels which are close to valleys, and other pixels. A large local slope 0.4 m is assigned to pixels within 5 pixel distances to ridges so that one has steep hillsides; a small local slope 0.02 m is assigned to the four-neighbors of valley pixels so that one has a wider valley bottom, and a medium local slope 0.1 m is assigned to the rest pixels.

For a valley pixel, if it is within 10 pixel distances to a peak junction, then it is assigned a large local slope 0.4 m; otherwise, it is assigned a small local slope 0.02 m.

Using estimates of local slopes and assuming the elevations of visible rivers which can be detected by the Alfoldi and Munday technique (1978), are lowest in a small area, the elevation growing algorithm can be defined as follows:

(1) Trace the border segments of visible rivers and give all the pixels elevation E_0. E_0 can be an arbitrary constant or a datum read from the map.

(2) Repeat until all the pixels are elevation labeled.

(2.1) Growing. If a pixel p has elevation $E_1(p)$, give its unassigned neighbors elevations $E_1(p)$ + DELTA(p) (DELTA(p) is the assigned local slope) unless: (a) an image boundary is encountered; or (b) a ridge is encountered.

(2.2) Including new valley segments. If any elevation-unlabeled valley segment is touched by an elevation-labeled pixel pe resulting from the growing, assign elevations to all the pixels of this segment. The end touching pe will have the same elevation as pe. Then, starting from this end, trace the whole segment and give every pixel linearly increasing elevation with some constant slope.

(3) For a ridge pixel, take the maximum elevation value from its 4-neighbors' as its elevation.

The relative heights of valley segments created by the elevation growing model are indicated by arrows in Fig. 3.5a, and the ground truth is shown in Fig. 3.5b. Because realistic shape of the hillsides from valleys to ridges were not taken into account in the raw elevation growing, only the relative elevations of the ridges and valleys are held to be accurate. Haralick et al. (1982)
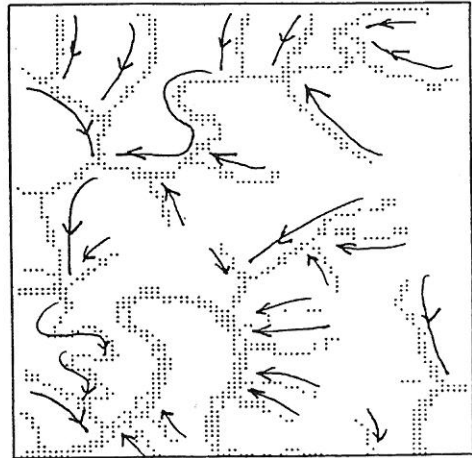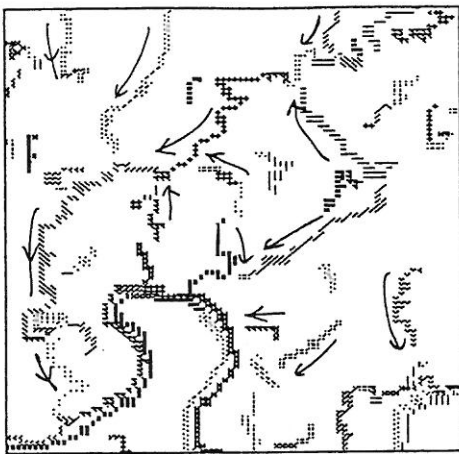

Fig. 3.5a. Relative elevations of valley segments. The arrows are from high ends to low ends.
Fig. 3.5b. Stream map created from ground truth.

describe a few interpolation procedures which permit more realistic eleva-
tion assignment to non-valley and non-ridge pixels. The interpolation makes
all non-ridge or valley pixels be the average of their north, south, east, and
west neighbors and makes ridge or valley pixels keep the values produced by
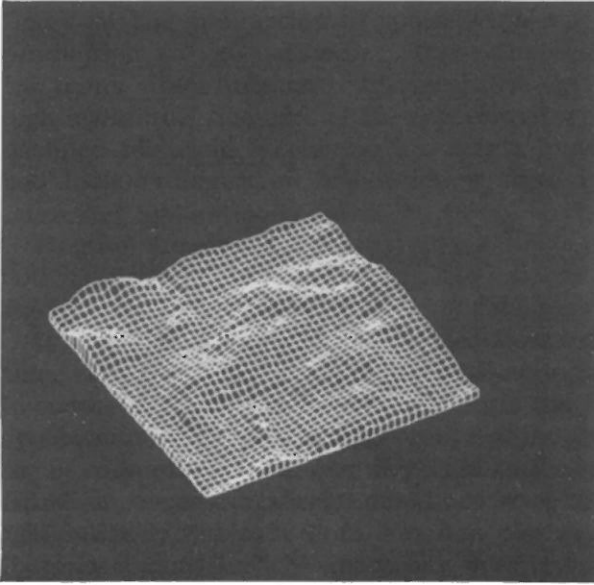the elevation growing model. This interpolation is shown in Fig. 3.6.

Fig. 3.6. Elevation model by using Laplacian mask.

3.3. Use of topographic modulation image to improve the result

It is impossible to obtain first-order partial derivatives f_x, f_y of the eleva-
tion model $f(x, y)$ from the topographic modulation image because the in-
formation of slope and aspect angle is mixed in the topographic modulation
image. The relationship between f_x, f_y, and $\cos\theta$ is contained in the follow-
ing equation:

$$\cos\theta = \frac{f_x S_x + f_y S_y - S_z}{\text{SQRT} (f_x f_x + f_y f_y + 1)} \tag{3.1}$$

where (S_x, S_y, S_z) is the known unit vector of sun illumination direction.

However, once a raw estimated elevation model f' is computed by the
elevation growing, the estimated partial derivatives f'_x, f'_y can be calculated
from f' using first differences. Then from the initial estimated f'_x and f'_y and
the observed value $\cos\theta'$ in the topographic modulation image, an a-posteriori
estimate f_x, f_y can be determined so that f_x and f_y maximize:

$$P(f_x f_y \mid \cos\theta') = \frac{P(\cos\theta' \mid f_x f_y)\, P(f_x f_y)}{-\int P(\cos\theta' \mid f_x f_y)\, P(f_x f_y)} \tag{3.2}$$

Because one wants $\cos\theta$ in eq. 4.1 to be as close as possible to the observed value $\cos\theta'$, $P(\cos\theta' \mid f_x f_y)$ can be written as:

$$k_1 e^{k_2 \cos^2(\theta - \theta')} \tag{3.3}$$

Assuming that the true derivative f_x, f_y minus the estimated derivatives f_x', f_y' are independent normals with mean zero, and standard deviation d, then f_x', f_y' calculated from the estimated elevation model, then:

$$P(f_x f_y) = k_3\, e^{-1/2\, [(f_x - f_x')/d]^2}\, e^{-1/2\, [(f_y - f_y')/d]^2} \tag{3.4}$$

where d is the standard deviation of the normal distribution.

Putting eqs. 3.2, 3.3, and 3.4 together, one wants to find f_x and f_y which maximize:

$$k\, e^{k_2 \cos^2(\theta - \theta')}\, e^{-1/2\, [(f_x - f_x')/d]^2}\, e^{-1/2\, [(f_y - f_y')/d]^2} \tag{3.5}$$

under condition (4.1) and $k = k_1 k_3$.

The maximizing f_x, f_y can be determined by exhaustive search over the range $[-2m, 2m] \times [-2m, 2m]$ of $(f_x - f_x', f_y - f_y')$ using an interval of 0.1 m. The resulting f_x and f_y is then fed back to the elevation growing procedure to start another iteration of estimating the elevation model. This processing can be repeated until the difference between the elevation averages of two consecutive iterations is small. The resulting elevation image as well as surface plot are shown in Fig. 3.7.
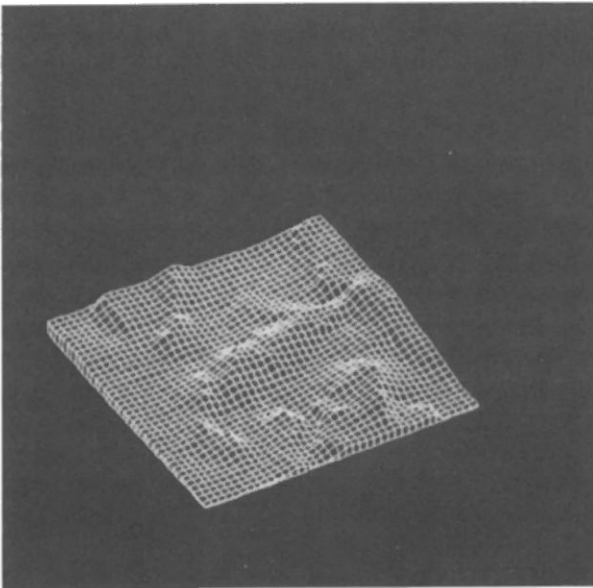

Fig. 3.7. Optimal elevation image.

In the initial elevation growing process, valleys are assigned small constant derivatives. However, if the end of a valley segment is located close to a ridge segment, then the initial estimated f'_x or f'_y will be high. When this high value is fed back to eq. 3.5, the final computed derivatives f_x and f_y will be much higher than the old small constant derivatives used in the elevation growing process. As can be seen in Fig. 3.7, the reconstructed surface obtained after feedback is more consistent at locations where valleys are close to ridges.

The image and surface plot of the elevation data read from Digital Terrain Tape (INCIC, 1980) for the area of Fig. 1.1 are shown in Fig. 3.8. The comparison of the minimum, maximum, and mean elevation values between the reconstructed elevation model and tape data for three test areas is listed in Table 3.1.
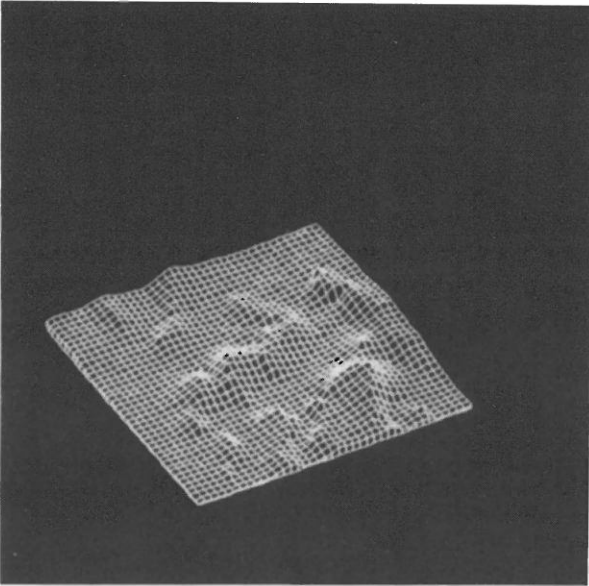

Fig. 3.8. Elevation data from digital terrain tape.

A pixel by pixel comparison between the reconstructed elevation model after scaling and the tape data are also listed in this table. The means of the absolute difference and the root mean square differences of elevation, and slope, are computed.

In Assumption 2 of Section 2, the spatial average of $\cos\theta$ is assumed to be the same for all material clusters. After the elevation model is reconstructed, all Xc values for the test area are calculated and listed in Table 3.2.

TABLE 3.1

Comparison of reconstructed elevation model and tape data

	Low	High	Mean
Tape data	950	2125	1350
Reconstruction	950	2033	1320

	Means of absolute difference	RMS difference
Elevation (ft)	115	143.76
Slope (ft)	0.117	0.154

TABLE 3.2

Listing of Xc values

Test Imagery 1

Cluster	1	2	3	4	5
Xc	0.7193	0.3735	0.7495	0.7734	0.7042

3.4. Reconstruction of the LANDSAT imagery

From eq. 2.6, the LANDSAT imagery can be reconstructed by the following steps:

(1) Illuminate the elevation model of Fig. 3.7 by an artificial sun at specified azimuth and elevation angles.

(2) Multiply the image of Step 1 by the reflectance image of Fig. 2.9.

(3) Add diffuse light image of Fig. 2.5 to the image of Step 2.

The resulting imagery is shown in Fig. 3.9.

For evaluation, the original and reconstructed imagery are scaled to the same graytone range [1, 64] and the difference image between them is computed. The means and variances of the difference imagery for three test areas are listed in Table 3.3.

4. CONCLUSION

Two problems are addressed in this paper: definition of an illumination model and computation of an elevation model. First, an illumination model was defined and a series of steps was used to extract the shadow image, the diffuse light imagery, the reflectance imagery, and the topographic modula-
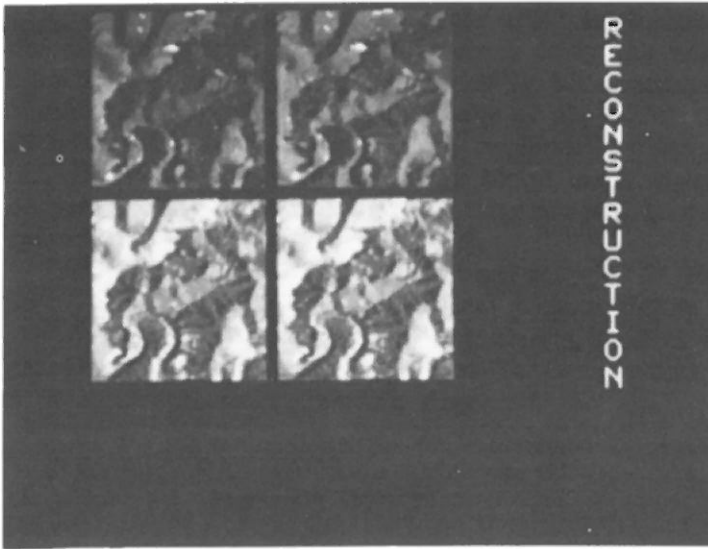
Fig. 3.9. Reconstructed LANDSAT imagery.

TABLE 3.3

Mean and variance of the difference imagery between original and reconstructed
LANDSAT imagery

Test Image 1

	Mean	Variance
Band 4	4.1	18.9
Band 5	3.62	16.2
Band 6	5.13	32.2
Band 7	8.27	43.3

tion image. The success of the process depends largely on clustering.

Next, knowing the sun azimuths and shadow image the ridges and valleys were found. An elevation growing process from valleys to ridges was found to be efficient in reconstructing the elevation model. Then an iterative method improved the elevation model by making it as consistent as possible with the topographic modulation image.

The techniques of this research work best for areas having big shadow areas. If water areas cannot be found to help identify the lowest valley locations, the elevation understanding problem will be more complicated. Other applications of this technique include the refinement of a given coarse digital elevation model using higher resolution multispectral imagery. In this kind of

application, the given digital elevation model essentially calibrates the elevation growing process so that the resulting refinement constitutes a smart interpolation process.

ACKNOWLEDGEMENT

REFERENCES

Alfoldi, T.T. and Munday, J.C., Jr., 1978. Water quality analysis by digital chromaticity mapping of LANDSAT data. Can. J. Remote Sensing, 4(2): 108—126.

Bryant, J., 1979. On the clustering of multidimensional pictorial data. Pattern Recognition, II: 115—125.

Campbell, J., Ehrich, R.W., Elliott, D., Haralick, R.M. and Wang, S., 1981. Spatial Reasoning in Remotely Sensed Data. Proc. 15th Int. Symposium on Remote Sensing of Environment, May 11—15, 1981, Ann Arbor, Michigan.

Duda, R.O. and Hart, P.E., 1973. Pattern Classification and Scene Analysis. Wiley, New York, N.Y., 482 pp.

Eliason, P.T., Soderblom, L.A. and Chavez, P.S., Jr., 1981. Extraction of topographic and spectral albedo information from multispectral images. Photogramm. Eng. Remote Sensing, 48: 1571—1579.

Goldberg, M. and Shlien, S., 1978. A clustering scheme for multispectral images. IEEE Trans. on Systems, Man, and Cybernetics, 8 (2): 86—92.

Haralick, R.M. and Wang, S., 1983. Relative elevation determination from Landsat imagery. Proc., COMPCON, September, 1983, Washington, D.C.

Haralick, R.M., Wang, S. and Elliott, D.B., 1982. Spatial reasoning to determine stream network from Landsat imagery. 6th Int. Conf. Pattern Recognition, 1982, Munich, Germany.

Kuchler, A.W., 1964. Potential Natural Vegetation of the Conterminous United States. Am. Geograph. Soc. Spec. Publ., No. 36.

Moik, J.G., 1980. Digital processing of remotely sensed images. NASA sp - 431, 1980.

Mulder, N.C., 1982. Methodology of color coding MSS and other data. Paper presented at Inter-Congress Symposium, Commission III (Mathematical Models, Accuracy Aspects, and Quality Control), International Society of Photogrammetry and Remote Sensing, 1982, Otaniemi, Finland.

National Cartographic Information Center, 1980. Digital terrain tapes, user guide, 2nd Ed.

Raines, G.L., Offield, T.W. and Santos, E.S., 1978. Remote sensing and subsurface definition of facies and structure related to uranium deposits, Powder River Basin, Wyoming. Econ. Geol., 73: 1706—1723.

Switzer, P., Kowalik, W.S. and Lyon, R.J.P., 1981. Estimation of Atmospheric Path-Radiance by the Covariance Matrix Method. Photogramm. Eng. Remote Sensing, Oct.: 1469—1476.

Thornbury, W.D., 1967. Regional Geomorphology of the United States. Wiley, New York, N.Y., 609 pp.

Vincent, R.K., 1973. Ratio Maps of Iron Ore Deposits, Atlantic City District, Wyoming. Proc. Symp. on Significant Results Obtained from the ERTS-1, pp. 379—386.

Wang, S., Elliott, D.B., Campbell, J.B., Ehrich, R.W. and Haralick, R.M., 1983. Spatial Reasoning in Remotely Sensed Data. IEEE Trans. Geoscience and Remote Sensing, GE-21 (1): 94—101.

Wildey, R.L., 1975. Generalized Photometry for Mariner 9. Icarus, 75: 613—626.