SPATIAL REASONING TO DETERMINE STREAM NETWORK FROM LANDSAT IMAGERY

R.M. Haralick, S. Wang, D.B. Elliott

Dept. of Electrical Engineering & Computer Science

Virginia Polytechnic Institute and State University

U.S.A.

ABSTRACT - In LANDSAT imagery, spectral and spatial information can be used to detect the drainage network as well as the relative elevation model in mountainous terrain. To do this, mixed information of material reflectance and topographic modulation in the original LANDSAT imagery must be separated first. From the material reflectance information, big visible rivers can be detected. From the topographic modulation information, ridges and valleys can be detected and assigned relative elevations. A complete elevation model can be generated by interpolating values for non-ridge and non-valley pixels. The small streams not detectable from material reflectance information can be located in the valleys with flow direction known from the elevation model. Finally, the flow directions of big visible rivers can be inferred by solving a consistent labeling problem based on a set of spatial reasoning constraints.

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 stream network and relative elevation model can be infered from a LANDSAT scene of mountainous and hilly terrain. The processing has a number of distinctly different steps. First to appropriately prepare the imagery for processing we must destripe it and perform haze removal. Destriping can be done by the Horn and Woodham [1979] technique. Haze removal can be done by the Switzer, Kowalik and Lyon [1981] technique. These two steps constitute the preprocessing and are not discussed in this paper.

To a first order effect, after preprocessing the cause of the intensity value at any pixel is due to the combined effect of 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 separating these two effects. For this purpose the Eliason, Soderblom and Chavez [1981] technique can be used to create two images from the one LANDSAT image. The first image is a reflectance image and the second image is a topographic modulation image and has information related to surface slope and sun illumination. The details are given in Section 2.

As discussed in Section 3, the reflectance image can be used by the Alfoldi and Munday [1978] procedure for identification of all areas of water. The topographic modulation image can be used to identify the ridges and the valleys. This is discussed in Section 4. 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 increases on the ridge path from the pixel to the saddle point where the ridge crosses a valley increases. The ridge valley elevation assignment procedure is discussed in Section 5. Once ridges and valleys have been located and assigned relative elevations, a complete elevation model can be generated by interpolating values for non-ridge and non-valley pixels. The interpolation procedures are discussed in Section 6. The final element of the spatial reasoning is the assignment of stream or river flow direction for those water bodies which were directly identified by reflectance properties. This is discussed in Section 7 and 8. In the remainder of this section, we review previous work done by saptial reasoning investigators.

CH1801-0/82/0000/0502 $00.75 © 1982 IEEE

Since the launch of the first Earth Resources and Technology Satellite (ERTS, later renamed LANDSAT) in July 1972, much work in remote sensing has been done by using pattern analysis and picture processing techniques for image classification, restoration and enhancement. Few people have tried the scene analysis or artificial intelligence approach to describe the image in terms of the properties of objects or regions in the image and the relationships between them. Ehrich [1977] found global lineaments by partitioning the image into windows and applying long, straight linear filters at different orientations in each window to extract local evidence. Dynamic programming [Montanari, 1971; Martelli, 1972] was then used to form complete global lineaments. VanderBrug [1976] tested various detectors to get linear features in satellite imagery. This was only at the local level. Later VanderBrug [1977a] used relaxation to reduce noise in the output. Finally VanderBrug [1977b] defined a merit function that can be used to select pairs of segments to be merged so that local line detector responses can be linked together into a global representation of the curves. His work is closely related to the Shirai [1973] technique which employed sequential line following to find edges in scenes containing polyhedra. Li and Fu [1976] used tree grammars to locate highways and rivers from LANDSAT pictures. The above investigations deal with the extraction of all the linear features from an image, but they do not deal with the interpretation of these linear features. In the following investigations, knowledge about the desired features are considered crucial in such analyses.

Bajcsy and Tavakoli [1975] argued that an image filter is not meaningful unless one has a world model, a description of the world one is dealing with. They recognized objects matching this description and filtered them out. This strategy is used to sequence the recognition of bridges, rivers, lakes, and islands from satellite pictures. Nagao and Matsuyama [1980] built an image understanding system that automatically located a variety of objects in an aerial photograph by using diverse knowledge of the world. It is one of the first image understanding systems that has incorporated very sophiscated artificial intelligence techniques into the analysis of complex aerial photographs. Fischler, Tenenbaum and Wolf [1981] designed a low-resolution road tracking (LRRT) algorithm for aerial imagery. The approach was based on a new paradigm for combining local information from multiple sources, map knowledge, and generic knowledge about roads. The final interpretation of the scene was achieved by using either graph search or dynamic programming.

Similarly, knowledge is important in our problem which requires analysis both at the local and global levels. Local level analysis will be discussed in Section 2 to 4; global level analysis will be discussed in Section 5 to 8.

2. Illumination model

The brightness and darkness in each band of LANDSAT images come from two main sources. First, they can be due to material properties. For example, in the spectral region (.8 - 1.1 μm) of band 7, water bodies absorb infrared radiation, so they appear as clearly delineated dark bodies; living vegetation reflects strongly in this portion of the infrared, so areas of living green vegetation appear as bright regions. Second, they may be due to topography and sun illumination angle effects. The mountain side facing to the sun appears as a bright region; the mountain side facing away from the sun may appear as a shadow or dark region. Unfortunately, the LANDSAT data values are some combination of these two effects. Eliason, Soderblom, and Chavez [1981] address this problem by defining an illumination model. In the following, their general theory about the brightness in LANDSAT imagery will be introduced, and extraction of material reflectance and topographic modulation information based on clustering on ratio images will be described.

The original LANDSAT image B' measuring the amount of light reflected from a surface at pixel (r, c) for wavelength w_i, i = 4, 5, 6, 7 is :

$$B'(r, c, w_i) = R(r, c, w_i) * Tp(r, c, p) + H(w_i)$$

where R is the brightness of the scene if the surface were flat, Tp is the modulation of the brightness introduced by topography, and H is the haze due to atmospheric scattering. The image of R is called material reflectance image, and the image of Tp is called topographic modulation image. p is the photometric function which depends on the phase angle, the incidence angle, and the angle of emergence [Wildey, 1975], but which does not depend on w_i. Two assumptions are made here:

1. Tp is independent of material properties and wavelength.
2. The photometric function p is independent of wavelength.

After H is calculated by the Switzer, Kowalik and Lyon [1981] technique, for each band, $H(w_i)$ is subtracted from $B'(r, c, w_i)$ at all pixels to get the haze-corrected image

$$B(r, c, w_i) = B'(r, c, w_i) - H(w_i) = R(r, c, w_i) * Tp(r, c, p).$$

Because Tp is independent of wavelength, in the ratio image of two bands with wavelength w_1 and w_2, each pixel (r, c) has gray level

$$B(r, c, w_1) / B(r, c, w_2)$$

$$= (R(r, c, w_1) * Tp(r, c, p)) / (R(r, c, w_2) * Tp(r, c, p))$$

$$= R(r, c, w_1) / R(r, c, w_2),$$

which is independent of Tp. Thus, the topographic information is removed. A simple demonstration of this theory is that, in the ratio image after all diffuse lighting has been removed, all shadows disappear.

Eliason, Soderblom, and Chavez estimate the material reflectance image R for each w_i by clustering using different ratio images as features. The result is a cluster image Cl (r, c). For each w_i, the average brightness value of all the pixels in each cluster is taken to represent R for their respective cluster. Two basic assumptions are inherent:

1. All image elements that group in a ratio cluster represent a single material.
2. The topographic slopes of all elements in a cluster are symmetrically distributed toward and away from the sun, such that their average brightness can be used to estimate the brightness of that material on a flat surface.

The material reflectance images will be used in detecting visible rivers described in Section 3.

From the cluster image, four material reflectance images $R(r, c, w_i)$ can be estimated because, for each cluster, four average brightness values can be calculated from the four bands, i.e., for i = 4, 5, 6, 7, if Cl (r, c) = k at pixel (r, c), then

$$R(r,c,w_i) = \frac{\sum_x \sum_y \underset{Cl(x,y)=k}{} B(x,y,w_i)}{\sum_x \sum_y \underset{Cl(x,y)=k}{} 1}$$

For any pair of the haze-corrected image $B(r, c, w_i)$ and material reflectance image $R(r, c, w_i)$, i = 4, 5, 6, 7, the topographic modulation image Tp can be calculated simply by taking the ratio of B over R. This topographic modulation image is an image whose tonal variation is unambiguously identified with surface slope and sun illumination angle.

3. Detection of Visible Rivers

Because one pixel in the LANDSAT image represents approximately a 57 meter by 80 meter area on the ground, the resolution is low. For the most part, it is not possible to directly observe the drainage network of the LANDSAT data. If rivers or lakes are visible to the humans, they can be detected by spectral information as described in this section. We call such rivers (including big lakes) "visible rivers." On the other hand, if the streams are not visible, they can only be detected by spatial information and we call them "invisible streams." For any window over the LANDSAT image, if one can detect some visible rivers by using spectral information and detect ridges and valleys by spatial pattern, then it is possible to continue to look for invisible streams by using spatial information. Examples of visible rivers and invisible streams are shown in Figure 1. The image was taken in April, 1976 over areas in Nicholas County, W. Va. and neighboring counties.

Figure 1 - LANDSAT scene in W. Va.
a indicate a visible river, and
b, c indicate invisible streams

Once the material reflectance image is created by the technique in last section, it can be used to identify visible rivers. In the spectral region (.8 - 1.1 μm) of band 7, water bodies absorb infrared radiation, so visible rivers appear as dark curves, and lakes appear as dark regions. In the material reflectance image of band 7, these dark features become more clear because shadows are removed. However, not all dark features are water bodies; the real water bodies can be identified by the following process [Alfoldi and Munday, 1978].

(1) A band 4 green coefficient x of every pixel is calculated as the ratio of

the radiance of band 4 over the radiance sum of bands 4, 5 and 6. Similarly a band 5 red coefficient y is calculated for every pixel. x and y are called LANDSAT chromaticity coordinates.

(2) In this coordinate system, Munday [1974] has determined a curve (Figure 2) which is the locus of the positions of chromaticity values of water bodies. If, for some pixels, the x, y values calculated in 1 are close to this curve, then those pixels can be identified as portions of water bodies.

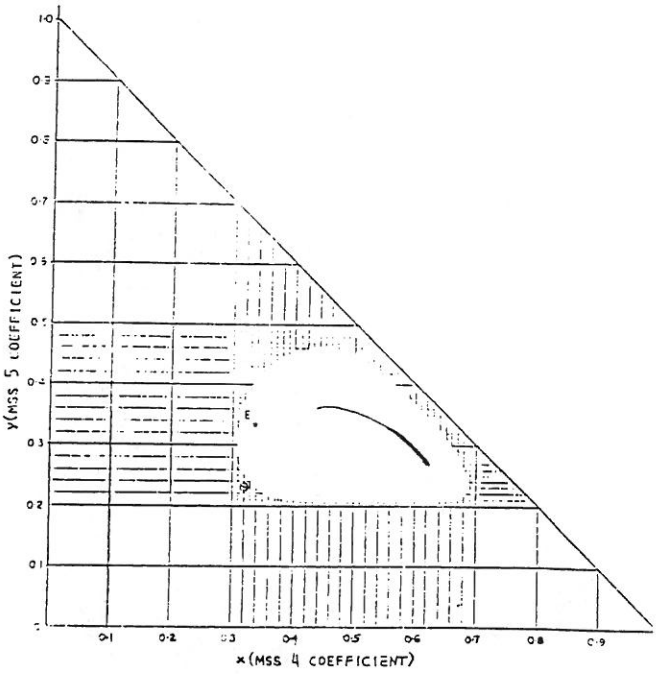

Figure 2 - Chromaticity plot

4. Ridge-Valley Map and Invisible Streams

After visible rivers are detected by both spectral and 2D spatial information, invisible streams can be detected by 3D spatial information. In this mountainous area, water flows through valleys, so that the drainage network of invisible streams is a substructure of the valley network. Therefore, the approach suggested here is to first get a relative elevation model, then extract the valley network from this elevation model, and finally extract the drainage network of invisible streams from the valley network. In this section, we describe how to extract shadowed and bright areas, create linear features on the borders between these areas, and classify these linear features into ridge and valley segments. In the next two sections, we discuss how to generate a relative elevation model. The extraction of valley network and the network of invisible streams will be discussed in Section 9.

In the topographic modulation image, bright areas indicate that the surfaces are facing to the sun, dark areas indicate shadows. In order to detect valleys and ridges, it is necessary to first segment the image into regions of shadowed and bright areas because valleys and ridges exist on the borders between these regions. Gray level thresholding can be used to determine shadowed and bright areas. We use the Watanabe [1974] technique in a recursive way to select thresholds. The details are given in Wang and Haralick [1982]. The connected components of them are shown in Figure 3.

Figure 3 - a. Connected components of dark regions

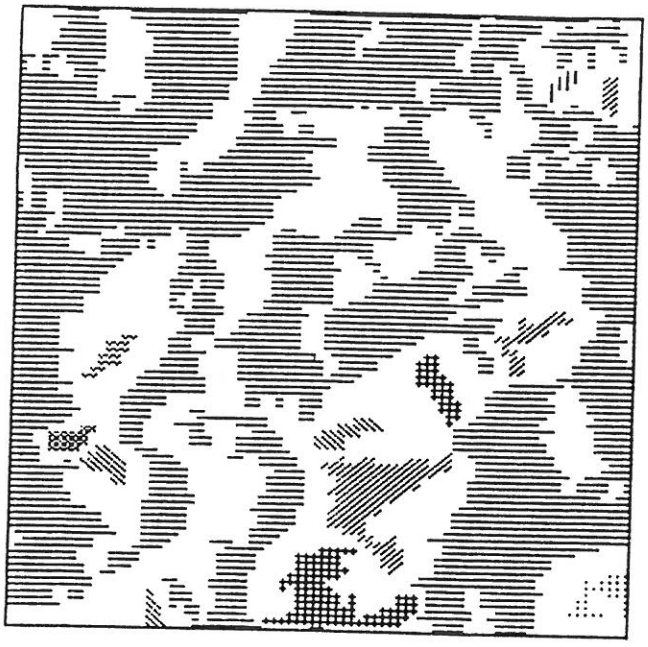

Figure 3 - b. Connected components of bright regions

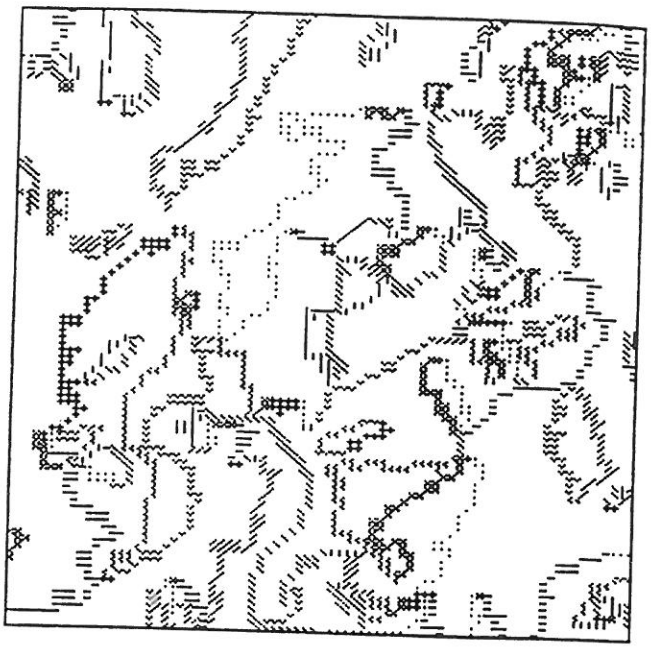

Figure 4 -Border Segments

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 Figure 1, each border segment is further broken into several pieces according to orientation: all the east-west parts can be separated from the north-south parts. The result is shown in Figure 4.

As the sun illumination comes from east, those border segments which are valley segments or ridge segments can be identified according to the brightness of their left and right regions. Because most of the trees in this area in April are unfoliated, the strongest region boundaries are shadow boundaries rather than tonal 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.

For east-west region boundaries, the above ridge-valley inference mechanism fails. Where east-west boundaries exist, some are ridges and some are valleys. To classify these east-west border segments correctly, it requires elevation information. As shown in Figure 5, if end a of the valley segment V1 is higher than end b, X can be determined to be a valley. Also, if end a is lower than end b or about the same, X can be determined to be a ridge. The results of ridge-valley finding are shown in Figure 6. Assignment of relative elevation to ridge and valley is discussed in the next section.

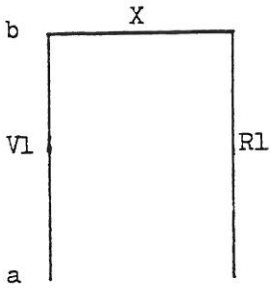

Figure 5 - Classifying east-west border segments. V1 is a valley segment; R1 is a ridge segment

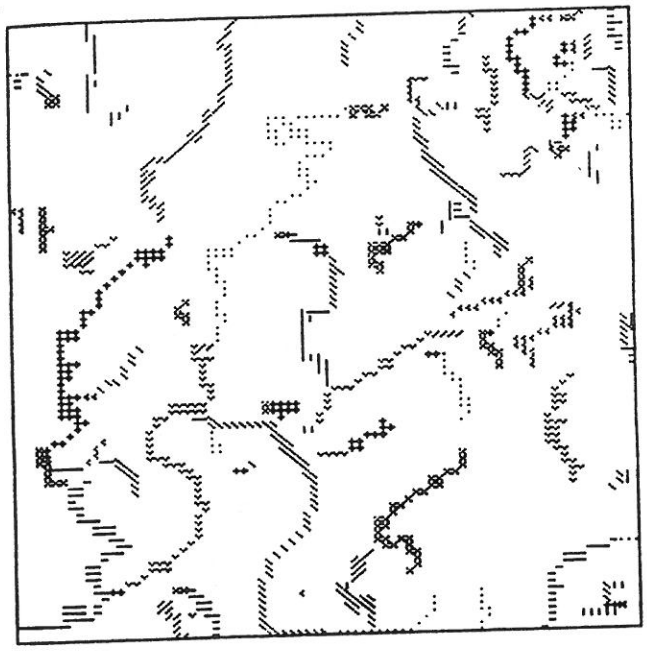

Figure 6 - a. Valley map consisting of the border segments which are identified as valleys.

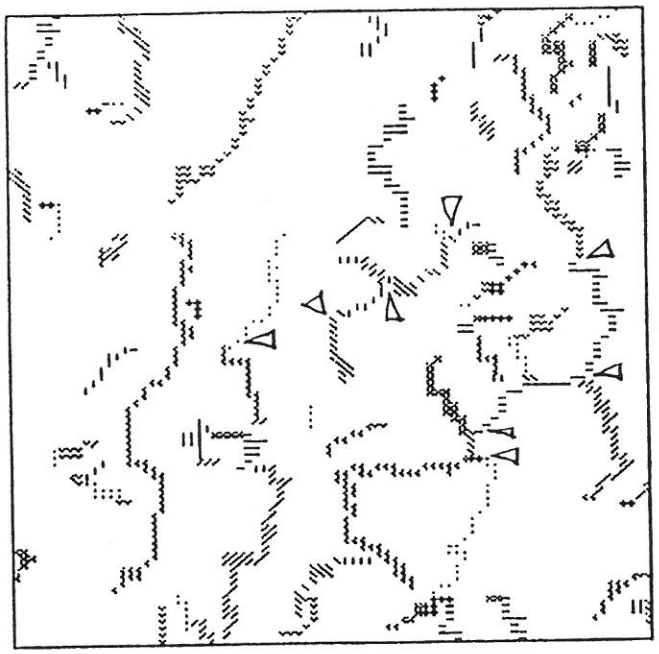

Figure 6 - b. Ridge map consisting of the border segments which are identified as ridges.

5. Relative Elevations of Ridge and Valley Segments

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. In this section, we describe how to estimate their relative elevations. First we will describe a model which can do the elevation assignment job, then we will give the equations of elevation assignment.

Assuming that we have a stream network in a mountainous area, and we know where the biggest rivers are, we can trace the network, starting from the biggest rivers, to find the flow directions of all the stream segments because water always flows from higher locations to lower locations. In other words, if the valley segments detected in the last section formed a network, then starting from the visible rivers detected in Section 3, we can trace the network and assign relative elevations to all the segments. Unfortunately, the observed valley segments do not form a network; there are many gaps. As shown in Figure 7, if it is dark on the right and bright on the left of stream Vb, then Vg cannot be detected due to the shadow on the right of Vb, and a gap exists between Vb and a smaller stream Vs.

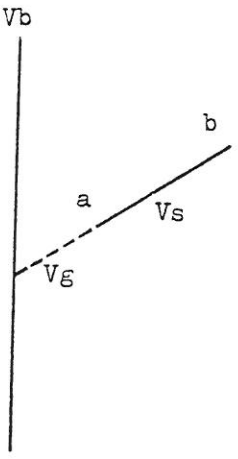

Figure 7 - The gap between a smaller and a larger stream

The knowledge that the cross-sections of valleys are V-shaped can be used to bridge the gaps. If one looks at topographic maps, the elevation contours of valleys such as in Figure 8 can be frequently found. Thus, if one draws a line ab perpendicular to the valley Va, the elevations are increasing from point o to point a , and also from point o to point b. However, if a ridge point is encountered during the process, the increasing has to stop because the elevation starts to decrease. Thus the route of growth is directed both by the valleys and by the ridges, in other words, by global information.

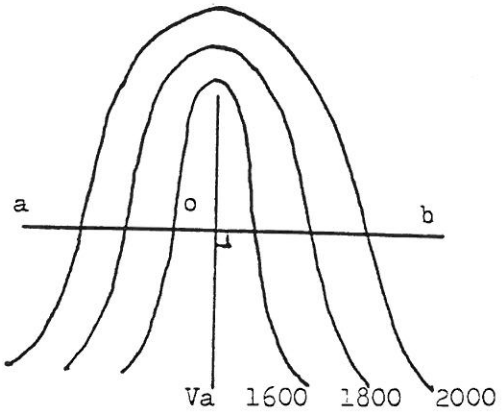

Figure 8 - The elevation pattern of valleys and its relation to elevation growing

Applying this idea to Figure 7 and assuming that growing propagates away from valley segment Vb, the end a of valley segment Vs will be touched first by this growing, and it is deduced that end b of Vs must be higher than end a. This is the basic idea for determining the higher-lower ends of all the valley segments. The elevations of all the points in one segment can be calculated if we know its slope. On the other hand, ridges get elevations when the growing stops at them. Now, we will give the simple equations of elevation assignment.

Our elevation growing model simply assumes that elevation increases monotonically from valleys to ridges or along valley segments from rivers to the saddles where a valley crosses a ridge. It can be used for assigning initial relative elevations to each pixel. Because no attempt is made to realistically account for the topographic shape of the hillsides from the valley to the ridge, the initial relative elevations will be more accurate for the ridge or valley labeled pixels than the non-ridge and non-valley labeled pixels. Section 6 discusses a more realistic procedure for hillside elevation estimation using the ridge valley elevations calculated in this section.

There are two ways a pixel can get assigned an elevation depending on whether the pixel belongs to a valley segment or whether the pixel does not belong to a valley segment. Let U be the set of valley segments. Two slopes are associated with each valley segment Vs in U: Sv(Vs) and Sp(Vs). Sv(Vs) is the slope along Vs itself. Sp(Vs) is the slope of lines outside of Vs and perpendicular to Vs.

The elevation growing model constructs the elevation function El: Zr X Zc -> Ip, where Zr is the set of row coordinates, Zc is the set of column coordinates, and Ip is the set of zero and positive integers. If p is a pixel belonging to a valley segment Vs and pl is the lower end pixel identified as in Figure 7, then

El(p) = El(pl) + Sv(Vs) * Dist(p, pl)

where Dist is the Euclidean distance between two pixels.

If p does not belong to any valley segment, and its elevation is originated from pixel pr of valley segment Vs, then

El(p) = El(pr) + Sp(Vs) * Dist(p, pr).

In a small area, one can assume the elevations of visible rivers are lowest. Assigning some initial elevation values to the pixels of the valley segments classified as visible rivers, the elevations of all the other pixels in the image window can be related to the initial elevations of visible river segments by repeatedly using the above two equations. The relative heights of valley segments created by elevation growing model are indicated by arrows in Figure 9, and the ground truth is shown in Figure 10.

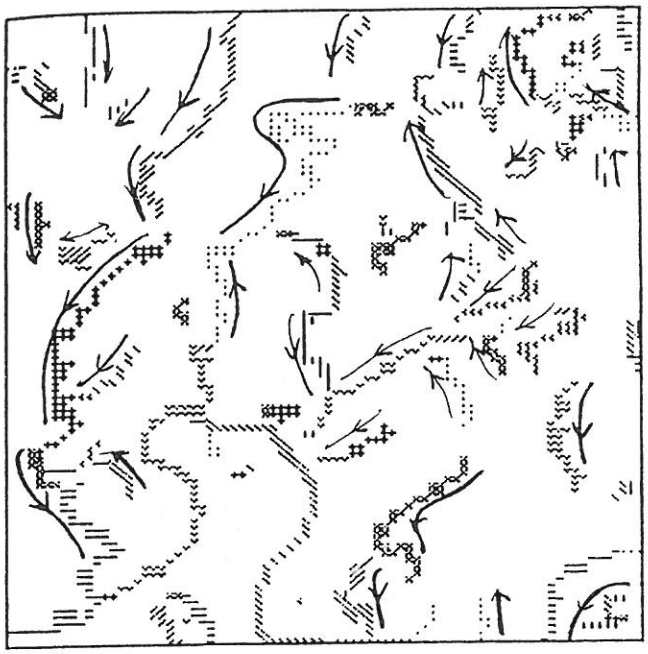

Figure 9 - Relative elevations of valley segments. The arrow are from high ends to low ends.

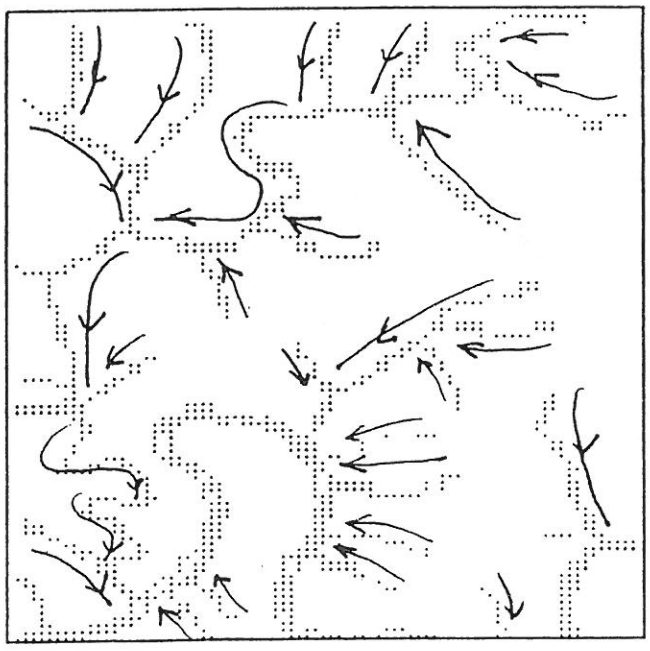

Figure 10 - Stream map created from ground truth.

5.1 Identification of Peak Junctions

When several valleys and ridges point toward a junction, very often this juction is a peak (peak at junction). The peak itself is formed by the junction of several ridges that radiate outward from the peak. (The idealized situation represented in Figure 11 shows four symetrically oriented ridges; in our area, real peaks are often formed by junctions of two or three ridges.) Ridges of course are separated by valleys, so the higher tips of valley segments tend to point toward peaks. The ridge segments intersect to form a peak, whereas valley segments tend to point towards peaks, without actually joining. In this subsection, we discuss the criteria which can be used to identify peak junctions.

Because ridge segments are the major features of peaks, we make the constraint that the number of ridge segments at a junction is larger than the number of valley segments. For many situations, it seems reasonable to relate the heights of peaks to the lengths of ridges that form the peaks. For our class of topographic forms (for example), it is unlikely that very high peaks can be formed by the intersection of very short ridges. As a result, to exclude very low peaks and false peaks from consideration, we impose a rather arbitrary constraint opon definitions of peaks. Currently, we define a peak junction as a junction composed of four border segments, with the number of its ridge segments larger than the number of valley segments, and the length of its longest ridge segment longer than 800 meters. The peaks thus located in Figure 1 are marked as triangles in Figure 6.b. The correspondence between this result and the topographical map is suprisingly good.

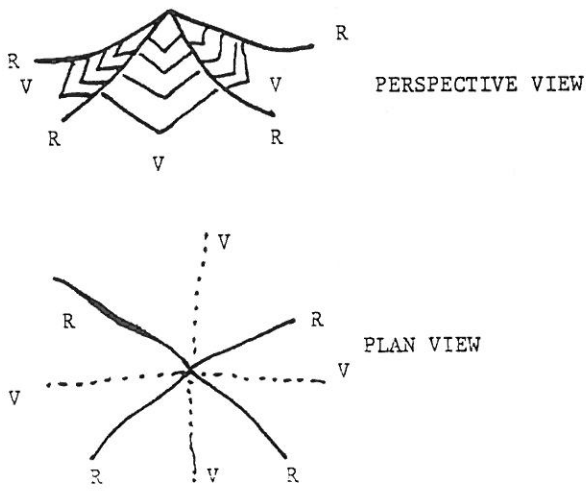

Figure 11 - Idealized relationships between peaks, valleys, ridges.

6. Interpolating Between Ridges and Valleys

In the last section all pixels were assigned elevations, but because realistic shape of the hillsides from valleys to ridges were not taken into account, only the relative elevations of the ridges and valleys are held to be accurate. In this section we describe a few interpolation procedures which permit more realistic elevation assignment to non-valley and non-ridge pixels.

The first interpolating surface has the given elevation values at ridges and valleys and has a 3 X 3 digital Laplacian of zero at all non-ridge and non-valley pixels. This will be referred to as the Laplacian surface. The system of linear equations which this constraint gives rise to can be written as

$$A\ x = b.$$

The vector x is the solution and represents the values to be assigned to each "variable" (non-ridge non-valley) pixel in the elevation model. The A matrix is defined by applying the digital Laplacian mask operator (Figure 12) to each variable pixel. A mask operator is applied to a pixel by placing the mask over the image so that the central (large positive) mask value is directly over the pixel whose value is to be computed. The pixel value is changed to make the sum of the mask values times the corresponding image values under them equal to zero. For the Laplacian surface only, Neumann boundary conditions are enforced along the outside rows and columns of the elevation model image. That is, the outermost row or column is repeated so that the mask operator can be applied to the outside pixels. There is one row in A for each variable pixel in the elevation model and one coefficient value in that row for each variable. A is a sparse matrix since no variable is constrained by more than four other variables (due to the definition of the digital Laplacian mask operator). The b vector is the right hand side of each of the linear equations in the system. The constants on the left hand side of each equation (that result from applying the Laplacian operator to a variable pixel that has a known pixel 4-adjacent to it) are carried to the right hand side and appear in b. For equations representing variable pixels not 4-adjacent to known pixels, the corresponding b element is zero.

$$\begin{array}{ccc} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{array}$$

Figure 12 - A digital Laplacian mask

The second interpolating surface has the given boundary values and minimizes the quadratic variation of the resulting surface [Grimson, 1981]. The boundary conditions with which the surface must agree are depth values along the zero-crossings. If the surface elevation function is E and subscripts denote partial differentiation, then the final surface E minimizes

$$\iint (E^2_{xx} + 2\,E^2_{xy} + E^2_{yy})\,d_x\,d_y$$

Since the surface function can be converted to a discrete grid format, the differential operators can be converted to difference operators, and the double integral can be converted to double summation, the solution of the above function can be formed by setting up a discrete corresponding set of linear equations

$$Q\ x = b.$$

The x and b vectors have the same meaning as in the Laplacian case and are constructed similarly. The Q matrix is likewise similar to the A matrix of the Laplacian. Instead of using Neumann boundary conditions at the edge of the image, the quadratic variation surface is defined by using special masks to fit the rows and columns near the outside edges. The six masks (Figure 13) are rotated as necessary and applied to the only appropriate variable pixels of the elevation image to define Q. Mask two is applied to corner pixels, mask three is applied to pixels in the outside row or column that are adjacent to a corner pixel, mask four is applied to other pixels in the outside rows and columns, mask five is applied to pixels in the next-to-the outside row and columns that are 8-adjacent to corner pixels, mask six is applied to other pixels in the next to the outside rows and columns, and mask 1 is applied to all other variable pixels in the image.

```
        2                          2
      4 -16   4                  -8
  2 -16  40 -16   2               8  -8   2
      4 -16   4
        2
```

(1) (2)

```
        2                            2
      4 -12   4                    4 -12   4
   -8  20 -12   2               2 -12  22 -12   2
```

(3) (4)

```
        2                            2
      4 -16   4                    4 -16   4
  -12  36 -16   2              2 -16  38 -16   2
      4 -12   4                    4 -12   4
```

(5) (6)

Figure 13 - Six masks for the quadratic variation method.

The third kind of interpolation surfaces can be created without using any mask. For each non-boundary pixel, we can first find its distances to the nearest valley pixels and nearest ridge pixels. From these distances and the elevations at these nearest valley pixel and nearest ridge pixel, either a linear, cubic, or fifth order fit interpolation can be used to calculate the elevation of this non-boundary pixel. If cubic fit is used, the first order derivative is zero at ridge and valley pixels. If fifth order fit is used, both the first and second order derivatives are zero at ridge and valley pixels. The resulting surface plots of these elevations are shown in Figure 14.

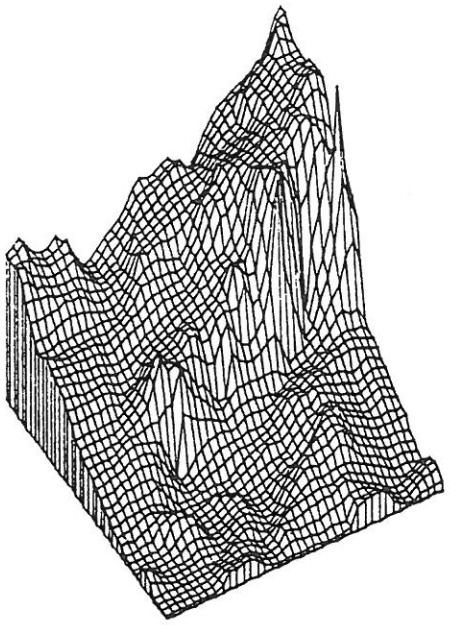

Figure 14a. Elevation Model by Method 1, Laplacian Mask

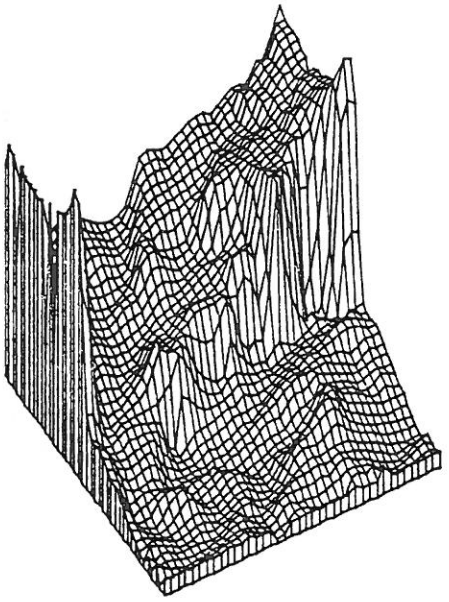

Figure 14b. Elevation Model by Method 2, Quadratic variation

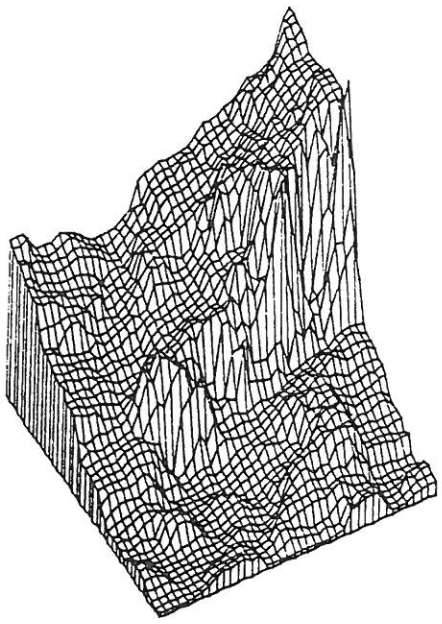

Figure 14c. Elevation Model by Method 3, Linear fit

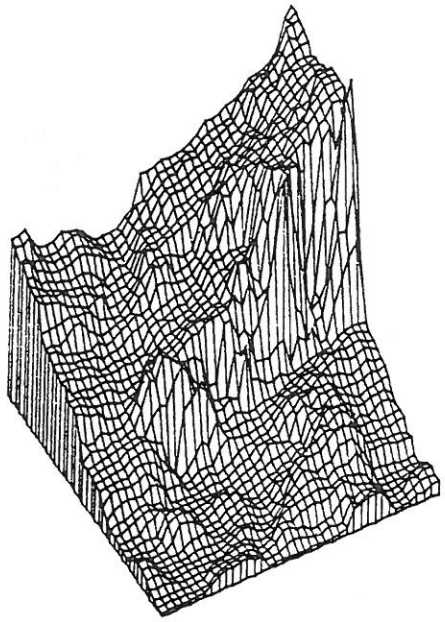

Figure 14d. Elevation Model by Method 3, Cubic fit

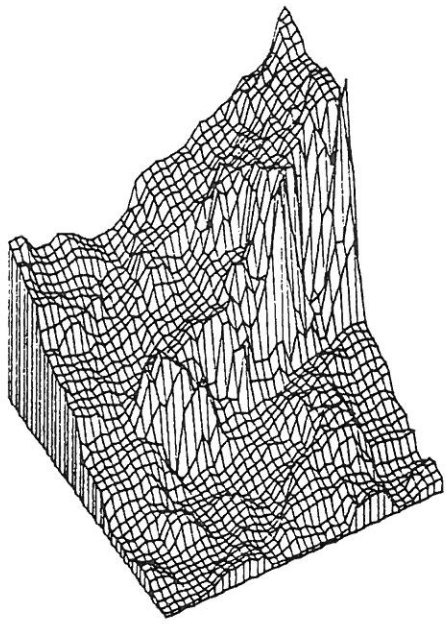

Figure 14e. Elevation Model by Method 3, Fifth order fit

Thus far, we have found the relative elevations of valley segments and generated several interpolation surfaces. More about invisible streams is discussed in Section 9. In the next two section, we discuss the assignment of flow directions to visible rivers detected in Section 3.

7. Flow Directions of Visible Rivers and Constraints at Junctions

In Section 5, an elevation growing model was used to find relative elevations of invisible streams in valleys. It remains to find the flow directions of visible rivers which are assigned constant elevations in the elevation growing model. This problem of assigning labels of {upstream, downstream} to the visible river segments is very much like the Waltz [1975] problem of labeling edges of polyhedra objects, and we need to find constraints applicable to streams.

It is believed that several stream segments joint at a junction with certain orientation and length patterns. The most obvious and important one is the configuration similar to Figure 7. It is plotted as Figure 15. When a smaller stream S2 flows into a larger stream S1 S3, very often the angle between S2 and S1 is less than 90 degrees. General rules about flow directions at junctions are given in Table 1.

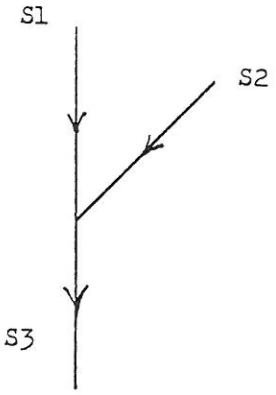

Figure 15 - One pattern of a stream junction

Angle (S1,S3)	(S1,S2)	(S2,S3)	Upstream	Downstream
=180	=90	=90		
|S1| = |S2|	<	|S3|	S1 and S2, S3	
|S1| = |S3|	>	|S2|	S2	S1 or S3
|S1| = |S2|	>	|S3|	S2 and S3, S1	
|S1| = |S2|	=	|S3|	S2	S1 or S3
=180	<90	>90	S1 and S2, S3	
=180	>90	<90	S2 and S3, S1	
<180	>90	>90	S1 and S3, S2	
>180	>90	<90		
|S1| = |S2| = |S3|			S2 and S3, S1	
|S1| = |S3| > |S2|			?	?
|S2| = |S3| > |S1|			S2 and S3, S1	
|S1| = |S2| > |S3|			?	?
=120	=120	=120		
|S1| = |S2| = |S3|			?	?
|S1| = |S3| > |S2|			S1 and S3, S2	
|S2| = |S3| > |S1|			S1 and S3, S2	
|S1| = |S2| > |S3|			S1 and S2, S3	

Table 1 - Rules of flow directions at junctions

Two constraint relations can be formally stated on the basis of Table 1 as follows. One is about all 3-tuples of stream segments that constrain each other because they meet in a junction; the other is about all 6-tuples of segment-label pairs where the stream segments meet in a junction and the labels are possible for that type of junction. Let a stream junction be a vertex at which 3 stream segments meet and at least one segment is identified as visible rivers. Let J be the set of stream junctions, S = {1, ..., M} be the set of stream segments at stream junctions, L = {upstream, downstream}, and X be the set of junction patterns in Table 1. Let f: J -> X be the function that assigns labels to functions, and h: X -> L^3 be the function that specifies 3-tuples of labels that can meet at each type of junction. For each x in X, we define Tx and Rx as follows.

$Tx = \{(s_1,s_2,s_3) \mid s_1,s_2,s_3$ meet in junction type x$\}$

$Rx = \{(s_1,l_1,s_2,l_2,s_3,l_3) \mid (s_1,s_2,s_3) \in Tx$ and $(l_1,l_2,l_3) \in h(x)\}$

Let $T = U_{s \in S} T_{f(s)}$ and $R = U_{s \in S} R_{f(s)}$.

T consists of all 3-tuples of stream segments that constrain each other, and R is the corresponding constraint relation. The labeling problem can be described by a compatibility model (S, J, T, R), which is a particular instance of the general consistent labeling problem [Haralick and Shapiro, 1979 and 1980]. Because we believe that there are many spatial inference problem which are instances of consistent labeling problem, in the next section we describe the form of the general consistent labeling problem as given by Ullmann, Haralick and Shapiro [1982].

8. Consistent Labeling and Spatial Reasoning Model

Let U be a set of objects called units, and L be a set of possible labels for those units. Let $T \subseteq \{f \mid f \subseteq U\}$ be the collection of those subsets of units from U that mutually constrain one another. That is, if $f = \{u_1, u_2, \ldots, u_k\}$ is an element of T, then not all possible labelings of $u_1,\ldots,u_k$ are legal labelings. Thus there is at least one label assignment $l_1, l_2, \ldots, l_k$ so that u_1 having label l_1, u_2 having label l_2, ..., u_k having label l_k is a forbidden labeling. T is called the unit constraint set. Finally, let $R \subseteq \{g \mid g \subseteq U \times L$, g single-valued, and $Dom(g) \in T\}$ be the set of unit-label mappings in which constrained subsets of units are mapped to their allowable subsets of labels. If $g = \{(u_1,l_1),(u_2,l_2), \ldots, (u_k,l_k)\}$ is an element of R, then $u_1, u_2, \ldots, u_k$ are distinct units, $\{u_1,u_2, \ldots, u_k\}$ is an element of T meaning $u_1, u_2, \ldots, u_k$ mutually constrain one another, and u_1 having label l_1, u_2 having label l_2, ..., and u_k having label l_k are all simultaneously allowed.

In the consistent labeling problem, we are looking for functions that assign a label in L to each unit in U and satisfy the constraints imposed by T and R. That is, a consistent labeling is one which when restricted to any unit constraint subset in T yields a mapping in R. In order to state this more precisely, we first define the restriction of a mapping. Let h:U-->L be a function that maps each unit in U to a label in L. Let $f \subseteq U$ be a subset of

the units. The restriction $h|_f$ (read h restricted by f) is defined by $h|_f$ = {(u,l) ∈ h | u ∈ f}. With this notation, we define a consistent labeling as follows.

A function h:U --> L is a consistent labeling if and only if for every f ∈ T, $h|_f$ is an element of R.

An example is given below. Suppose the inputs to the problem are as follows:

U = {1, 2, 3, 4, 5}
L = {a, b, c}
T = { {1}, unary constraint
{1, 2}, binary constraints
{2, 5},
{1, 3, 4}} ternary constraint
R = { {(1,a)},}(1,b)}, unary constraint
{(1,a),(2,a)},
{(1,a),(2,b)},
{(1,b),(2,b)}, binary constraints
{(2,a),(5,a)},
{(2,b),(5,c)},
{(1,a),(3,a),(4,c)}, ternary constraints
{(1,b),(3,a),(4,a)}}

Then h = {(1,a) (2,a) (3,a) (4,c) (5,a)} is a consistent labeling. To see this note that $h|_{\{1\}}$ = {(1,a)}, $h|_{\{1,2\}}$ = {(1,a),(2,a)}, $h|_{\{2,5\}}$ = {(2,a),(5,a)}, and $h|_{\{1,3,4\}}$ = {(1,a),(3,a),(4,c)} are all elements of R.

If having $(l_1, \ldots, l_N)$ applied to $(u_1, \ldots, u_N)$ when $(u_1, l_1, \ldots, u_N, l_N)$ is not in R is allowed with a penalty, the process is called inexact consistent labeling [Shapiro and Haralick, 1981]. In order to include these mappings, an error function Ew is defined. Let Ew : $T \times L^N$ -> [0, 1] be the error weighting function. Ew $(u_1, \ldots, u_N, l_1, \ldots, l_N)$ is the error which occurs when labels $(l_1, \ldots, l_N)$ are applied to $(u_1, \ldots, u_N)$. The mapping h : U -> L is an inexact consistent labeling if for all $(u_1, \ldots, u_N)$ in T, the summations of Ew $(u_1, \ldots, u_N, h(u_1), \ldots, h(u_N))$ is within some upper bound.

In spatial reasoning problems, many spectral and geometrical properties can be detected for the locally detected units. Some frequently used properties are average gray level, size, and shape descriptors. Let P be the set of properties. The spatial reasoning model is (U, P, L, T, R, Ew). U, L, T, Ew have the same meanings as before; however, the elements in R now have the form $(u_1, p_1, l_1, \ldots, u_N, p_N, l_N)$ where p_i is the list of property value ranges for all the properties in P for unit u_i, i = 1 to N. It means that if the property values of u_i are within the ranges specified by p_i for i = 1 to N, and $(u_1, \ldots, u_N)$ is contained in T, then it is legal to label l_1 to u_1, ..., l_N to u_N at the same time.

The spatial reasoning model (U, P, L, T, R, Ew) can be applied to find the flow directions of visible rivers. U contains the units of visible rivers plus the units of invisible streams intersecting the visible rivers at junctions. P contains all the properties detectable from the border segments. L is {Upstream = 1, Downstream = 2}. T contains the junction relations. R contains the relations of legal flow directions defined in Table 1. Ew is the number of times inconsistency occurs at junctions normalized by the total number of junctions.

9. Conclusion

To detect stream network in LANDSAT, both visible rivers and invisible streams need to be distinguished. Visible rivers can be detected by both spectral and 2D spatial information. However, the detection of invisible streams needs 3D spatial information.

For invisible streams, ridge and valley segments must first be detected and then an elevation growing model can be used to assign relative elevations to them. Interpolation can generate surface elevation at all locations from the known values at ridge and valley segments. From this elevation surface a valley network can be generated easily. Another way to form a valley network is to create gap units as in Figure 7 during elevation growing. Local information including rules in Table 1 and other knowledge can be used to determine the invisible stream network as a subset of the valley network. The flow directions of invisible streams come directly from the relative elevations of valley segments.

For the visible rivers part, the consistent labeling based spatial reasoning model can be used to find the flow directions of visible rivers whose units are assumed to have all constant elevations in the elevation growing model.

Based on the consistent labeling model, two types of spatial reasoning models can be formulated. If one is only interested in the classification or labeling of the two dimensional space so that roads, buildings or other ground objects can be identified, the model is called a 2D spatial reasoning model. One example is the model discussed in Section 8. If, in addition to the classification, some feature values such as terrain elevations are needed over the two dimensional space, it is called a 3D spatial reasoning model. In other words, in the 2D spatial reasoning model, the output specifications are symbolic; in

the 3D spatial reasoning model, the output specifications are numerical.

An application of the 2D spatial reasoning model which we hope to report on soon is to recognize the ground objects in an urban area by segmenting the aerial image into regions, measuring the properties of these regions, formulating constraints in T and R, and applying the model. An application of the 3D spatial reasoning model which will be discussed in Wang's forthcoming dissertation will be to find the best set of segment slopes so that the estimated relative valley ridge elevations are as accurate as possible.

Acknowledgement

The authors want to thank J. Campbell, R.W. Ehrich, and L.G. Shapiro for their helpful suggestions and discussions.

References

T.T. Alfoldi and J.C. Munday, Jr., Water Quality Analysis by Digital Chromaticity Mapping of LANDSAT Data, Canadian J. of Remote Sensing, 4 (2): 108 - 126, 1978.

R. Bajcsy and M. Tavakoli, Image Filtering - a Context Dependent Process, IEEE Trans. on Circuits and Systems, May 1975.

R.W. Ehrich, Detection of Global Edges in Textured Images, IEEE Trans. on Computers, Vol. C-26, June 1977, pp. 589-603.

P.T. Eliason, L.A. Soderblom, and P.S. Chavez, Jr., Extraction of Topographic and Spectral Albedo Information from Multispectral Images, Photogrammetric Engineering and Remote Sensing, Vol. 48, No. 11, Nov. 1981, pp. 1571-1579.

M.A. Fischler, J.M. Tenenbaum and H.C. Wolf, Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique, Computer Graphics and Image Processing, Vol. 15, 1981, pp. 201-223.

W.E.L. Grimson, An Implementation of a Computational Theory of Visual Surface Interpolation MIT AI Memo, 1981.

R.M. Haralick and L.G. Shapiro, The Consistent Labeling Problem: Part 1, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-1, No.2, April 1979.

R.M. Haralick and L.G. Shapiro, The Consistent Labeling Problem: Part 2, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 3, May 1980.

B.K.P. Horn and R.J. Woodham, Destriping LANDSAT MSS Images by Histogram Modification, Computer Graphics and Image Processing, Vol. 10, 1979, pp. 69-83.

R.Y. Li and K.S. Fu, Tree System Approach for LANDSAT Data Intepretation, Proceedings of the IEEE Symposium in Machine Processing of Remotely Sensed Data, June 1976.

A. Martelli, Edge Detection Using Heuristic Search Methods, Computer Graphics and Image Processing, Vol.1, pp.169-182, 1972.

U. Montanari, On the Optimal Detection of Curves in Noisy Pictures, CACM, Vol.18, pp.335-345, 1971.

J.C. Munday, Jr., Lake Ontario Water Mass Determination from ERTS-1, Proc. 9th Intl. Symp. Remote Sensing of Environment, ERIM, Ann Arbor, pp. 1355-1368, 1974.

M. Nagao and T. Matsuyama, A Structural Analysis of Complex Aerial Photographs, Plenum Press, 1980.

L.G. Shapiro and R.M. Haralick, Structural Description and Inexact Matching, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-3, No.5, Sep. 1981.

Y. Shirai, A Context Sensitive Line Finder for Recognition of Polyhedra, Artificial Intelligence, Vol.4, pp.95-119, 1973.

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

G.J. Vanderbrug, Line Detection in Satellite Imagery, IEEE Trans. on Geoscience Electronics, Vol. GE-14, Jan. 1976, pp. 37-44.

[1977a] G.J. VanderBrug, Experiments in Iterative Enhancement of Linear Features, Computer Graphics and Image Processing, Vol. 6, 1977, pp. 25-42.

[1977b] G.J. Vanderbrug, Curve Representation and Mapping, Computer Science Technical Report TR-561, U. of Maryland, Aug. 1977.

D. Waltz, Understanding Line Drawings of Scenes with Shadows, in The Psychology of Computer Vision, edited by P.H. Winston, McGraw-Hill, 1975.

S. Wang and R.M. Haralick, Automatic Multi-Threshold Selection, Computer Science Technique Report, 1982, Virginia Polytechnic Institute.

S. Watanabe and the CYBEST Group, An Automated Apparatus for Cancer Prescreening: CYBEST, Computer Graphics and Image Processing, 3, 1974, pp. 350-358.

R.L. Wildey, Generalized Photochinometry for Mariner 9, Icarus, Vol.75, pp.613-626, 1975.

J.R. Ullmann, R.M. Haralick, and L.G. Shapiro, Computer Architecture for Solving Consistent Labeling Problem, Proceedings of SSST, 1982.