

An Object-Specific Image-Texture Analysis of H-Resolution Forest Imagery

G. J. Hay,^{*} K. O. Niemann,^{*} and G. F. McLean[†]

A new structural image-texture technique, termed the *triangulated primitive neighborhood method (TPN)*, is employed to investigate the variable spatial characteristics of high-resolution forest objects, as modeled by a Compact Airborne Spectrographic Imager data set. Based on current psychophysical texture theory, this technique incorporates location-specific primitives and a variable-sized and shaped moving kernel to automatically provide object- and area-specific regularized images. These object-rich, but variance-reduced images allow a traditional classifier to be used on a complex high-resolution forest data set with improved accuracy. The robustness of this technique is evaluated by comparing the maximum likelihood classification accuracy of nine forest classes generated from a combination of the grey level cooccurrence matrix method, semivariance, and customized filters, against those derived from the TPN method. By including into the classification scheme an object-specific channel that models crown density, the highest overall classification accuracy (78%) from all techniques is achieved with the TPN method.

INTRODUCTION

When traditional remote sensing texture techniques are compared to those performed by human interpreters (Hay and Niemann, 1994), it quickly becomes apparent that the domain of texture still remains to be fully exploited within remote sensing. To investigate the variable spatial characteristics of forest objects in high-resolution digital imagery, a new structural image-texture

technique, termed the *triangulated primitive neighborhood method (TPN)* (Hay, 1993), has been designed. This method was initially conceptualized to instruct a computer vision system to *feel* the surface of a forest canopy, and describe the texture perceived. To develop this concept, the actions human interpreters perform while analyzing the texture of different scale forest imagery, and an understanding of current visual, haptic (touch), and remote sensing texture theory, were employed as a pattern, and will be briefly reviewed.

Casual experience with the phenomenon texture belies the complex and diverse nature of the domain designed to understand, define, and emulate what we intrinsically perceive it to be. Originally the term *texture* was adopted from textiles, where it refers to the weave or structure of various threads—loose or tight, even or mixed. Since then, it has evolved into a paradigm, that includes music (Kerman, 1980), artificial intelligence (Marr, 1982), vision (Bergen and Landy, 1991), psychophysics (Julesz, 1975), haptics (Lederman, 1982), pattern recognition (Ballard and Brown, 1982), and remote sensing (Haralick, 1979); and it is perceptible by all sensory modalities, though vision and haptics appear the most researched (Heller, 1989; Jones and O'Neill, 1985). Due in part to its perceptual nature, no singly accepted definition of texture exists (Haralick, 1979; Lederman et al., 1986); nevertheless, texture may be described as an intuitive scale specific characteristic of form perception (Uttal, 1983; 1988) capable of providing instantaneous structural information to the interpreter (Julesz and Bergen, 1983) regarding the size (Bergen and Adelson, 1988), shape (Klatsky et al., 1987), and spatial arrangement (Julesz, 1981) of the varying objects that constitute a three-dimensional scene, or model of that scene (i.e., a digital image).

By using one's hand as a texture sensor, texture perception is restricted to objects within the areal extent and haptic sensibility of its cutaneous receptors; similarly, visual texture perception is restricted to objects

^{*} University of Victoria, Department of Geography, Victoria, British Columbia

[†] University of Victoria, Department of Mechanical Engineering, Victoria, British Columbia

Address correspondence to G. J. Hay, University of Victoria, Department of Geography, P.O. Box 3050, Victoria, B.C., V8W 3P5, Canada.

Received 8 March 1995; revised 2 August 1995.

within the field of view, and resolving power of the lens. From these simple observations three important aspects of texture and sensors are hypothesized:

- Due to resolving power limitations (Slater, 1980), all sensors (visual, haptic, remote sensing, etc.) perform as spatial frequency filters, where patterns higher in frequency than the spatial resolution and lower in frequency than the spatial extent examined are inherently filtered out (Clark, 1990). As a consequence, texture analysis constitutes a form of contextual spatial pattern analysis, (dimensionally) limited by the spatial resolution of the sensor, which, in turn, dramatically effects the modeled, or perceived texture. Here, *context* refers to the relationship between a single object (or group of objects) and other objects within the scene (Jensen, 1986).
- Within the fixed minimum and maximum spatial resolutions of haptic and visual sensors, a dynamic range of contextual analysis occurs within a scene, indicating varying optimal spatial resolutions, and different-shaped optimal spatial regions (spatial extents) must also exist and should be assessed to adequately detect the different objects constituting the scene texture. This is critical as potential information about different objects at multiple scales exists within a single scene, with no one scale of observation ubiquitously capable of describing information from all others.
- Description of a scene's texture requires analysis of its components which exist on a relative geographic scale (Lam and Quattrochi, 1992) one spatial level higher than the texture of the object being described (i.e., large geographic scale = small area analysis in greater detail). For example, the texture of a forest stand cannot be described by the leaves and needles on a tree, as they in no way describe the immediate spatial relationship of trees comprising the stand. Rather, stand texture must be described by some relationship of the various sized, shaped, and spatially arranged trees, which represent the stand's constituent higher level components.

The terms *image-texture analysis* and *human-texture analysis* are used here to differentiate between remote sensing texture techniques and texture as perceived by humans. Traditionally, image-texture analysis involves statistical techniques that contextually evaluate the spatial variation of static sized and shaped, but tonally variable individual pixels, or tuples, within a fixed sampling range. As a consequence, image-texture is considered a measure of the spatial variation of image tone or

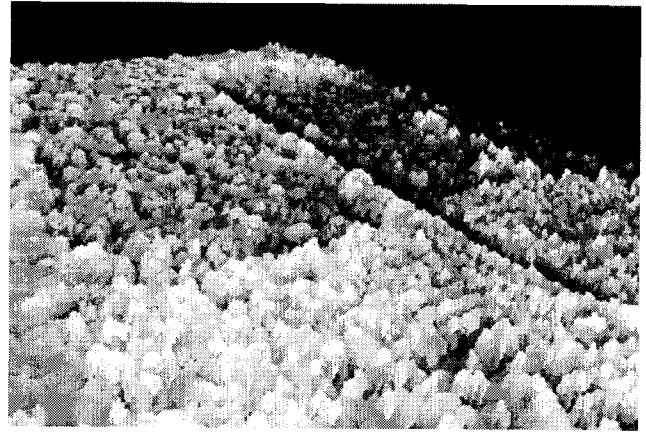


Figure 1. Perspective view of the study area. CASI Band 2 (NIR: 734.2–746.7 nm) is used as a pseudo-elevation channel, with Band 1 (Red: 675.1–685.9 nm) and Band 2 used as drapes. This results in the visual perception of individual tree crowns and stands (image-objects) appearing isolated from each other and their backgrounds.

intensity, with the grey level cooccurrence matrix method (GLCM) appearing the standard against which other techniques are evaluated (Weszka et al., 1976; Gu et al., 1989; Du Buff et al., 1988; Gong et al., 1992). In humans, a perceptual analysis of the variable size, shape, and spatial arrangement of individual *scene-* or *image-objects*, that is, tree crowns, naturally existing within the scene, or perceptually generated from H-resolution (Strahler et al., 1986) pixel groups, form a cognitive impression of texture.

Though not fully understood, humans possess the intrinsic ability to relate groups of pixels (image-texture primitives) of varying size, shape, and tone, to a plausible surface in three dimensions (Ballard and Brown, 1982; Julesz and Bergen, 1983), even though they represent only two Euclidean dimensions. This may in part be due to the height/topographic information integrated within the reflective signal of the image, which appears more dominant in higher, than in lower resolution forest imagery (Franklin and McDermid, 1993). Jarvis (1983) and Horne and Brooks (1989) suggest that visual depth information is readily available from a monocular image via: size perspective (diminution of size with distance), motion parallax, areal perspective (haziness, etc., associated with distance), occlusion effects, outline continuity (complete objects look closer), and surface shading variations. As a result, when high-resolution image-inherent topographical/height cues are visually exploited within a three-dimensional environment (Fig. 1), they produce a powerful visual perception of image-texture closely corresponding to the scene texture, though care must be taken in analysis, as no explicit height information exists in the digital imagery.

It is important to note that how well image-texture

visually relates to scene-texture is strongly dependent on the spatial resolution of the imagery. Woodcock and Strahler (1987) indicate that texture analysis must be performed on H-resolution imagery as numerous measurements are required for each element or class in order to allow the characteristic spatial texture to occur. Since the structural characteristics of scene-objects visually dominate in H-resolution imagery, image-objects geographically correlate exceptionally well to their corresponding scene-objects (Cohen et al., 1990), and the resulting image-texture strongly models the scene-texture. Conversely, when imagery is L-resolution, each pixel represents a regularization of an unknown number of scene-objects—defined by the point spread function of the sensor (Woodcock et al., 1988a), thus the sensor's instantaneous field of view (IFOV) dominates, potentially producing an image-texture completely unrelated to the scene elements it models (Marceau et al., 1994a,b).

One of the challenges in using high-spatial resolution remotely sensed imagery for classification is termed the *H-resolution problem*, and implies that as the spatial resolution of a sensor increases, so does the interclass spectral variability of surface features, resulting in a reduction of statistical separability with traditional classifiers and a consequent reduction in classification accuracy (Woodcock and Strahler, 1987; Marceau et al., 1990). A common practice for reducing interclass variability involves spatial filtering methods (Irons et al., 1985; Cushnie, 1987; Raffy, 1992), yet none of these techniques are object-specific, that is, none of them locate and analyze the interaction of scene-objects as modeled by discrete image-objects. Instead, analysis traditionally occurs on individual pixels which geographically represent (in the case of H-resolution imagery) an arbitrary portion of a single object, or (in the case of L-resolution imagery) unknown combinations of various scene-objects within an arbitrary sized contextual kernel equally (weighted) over the image. A thorough study by Marceau et al. (1990) indicates that as much as 90% of the image-texture classification variability is accounted for by the window size used, and Gong et al. (1992) report that poor class accuracies result from inappropriately sized kernels; yet little guidance is offered for selecting an appropriate sized kernel, even though this factor is critically important in defining the area, and or potential objects under analysis. Such an approach appears in total disregard to the concept of multiple object scales within a single scene.

To take advantage of the multiple object scales present in H-resolution forest imagery, and to provide object-rich, yet variance-reduced datasets for traditional classifiers, the TPN method involves a variety of object-specific steps. These include automatically identifying the crown center of individual trees, generating uniquely shaped primitives which locationally represent the spa-

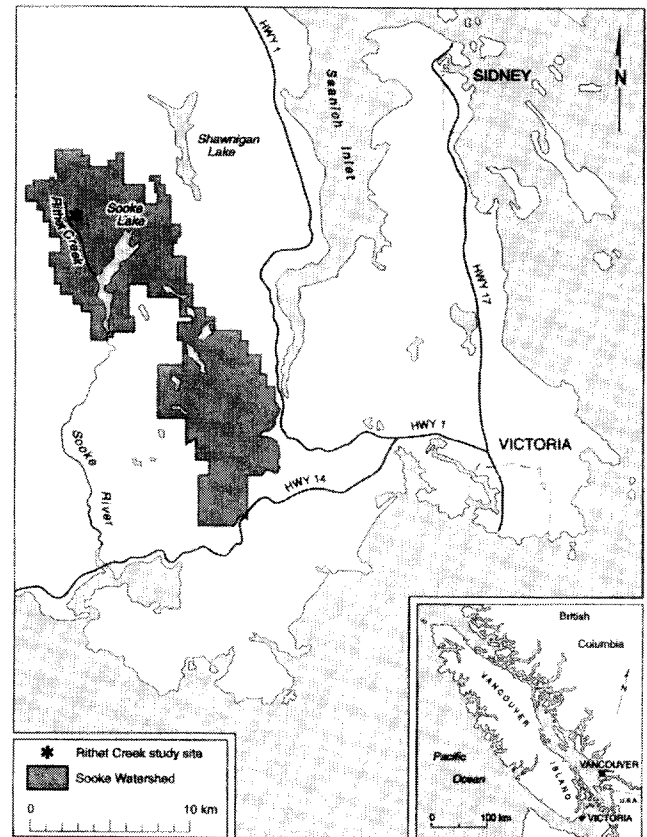


Figure 2. Map of study site.










tial association of tree groups, then contextually assessing these tree clusters within an object-specific variable sized and shaped moving kernel. These steps are modeled after current visual and haptic texture theory (Julesz and Bergen, 1983; Jones and O'Neil, 1985; Bergen and Adelson, 1988; Uttal, 1988), which suggests human-texture analysis occurs by an instantaneous global process defining object locations based on density differences. Scrutiny then proceeds at these locations in the form of a moving window (i.e., the resolving ability of the lens) that changes its scale according to the size of the features being sought, so that objects' spatial relationships may be determined.

The primary objective of this research is to evaluate the appropriateness and advantages of the TPN method, as an image-texture descriptor of the spatial interaction of individual trees in a complex, high-resolution remotely sensed digital forest image. Evaluation will be based on comparing the classification accuracy of classes in a thematic map generated by a combination of semivariance, custom filters, and the GLCM method, against those derived from the TPN method.

Site Characteristics

Due to its well-documented forest history and site accessibility, a 20.4 ha site was selected within the Sooke

Table 1. Class Characteristics

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Description	Mature	Mature	Mature	Mature	Young	Young	Immature	Immature	Non-veg*
Spatial Pattern	Clumpy 	Clumpy 	Clumpy 	Clumpy 	Uniform 	Clumpy 	Smooth 	Smooth 	Smooth 
Dominant Species	DF	DF	DF	DF	DF	Alder	DF	DF	-
Age (yrs)	141 - 250	141 - 250	141 - 250	141 - 250	27 - 30	27 - 30	4 - 7	4 - 7	-
Height (m)	55.5 - 64.4	46.5 - 55.4	37.5 - 46.4	28.5 - 37.4	10.5 - 19.4	12.5 - 15.4	2.5 - 4.4	0.0 - 2.4	-
Canopy Diameter (m)	6 - 12	6 - 10	5 - 10	5 - 8	2 - 6	6 - 8	2 - 4	0.0 - 2	-
Crown Closure (%)	56 - 65	56 - 65	56 - 65	56 - 65	66 - 75	36 - 45	16 - 25	0.0 - 15	-
Stems per/ha	130 - 140	150 - 170	160 - 175	215 - 230	690 - 720	60 - 70	-	-	-

* Non-veg class is composed of rock outcrops, gravel roads, and a dry gravel stream bed.

Watershed, west of the city of Victoria, British Columbia, Canada (Fig. 2). In this area, the very dry maritime Coastal Western Hemlock biogeoclimatic subzone dominates, though a small component of moist maritime Coastal Douglas-Fir subzone also exists. Due to its longevity, and the area history of repeated fire and wind-throw disturbances, successional processes on any site unit rarely proceed to a climax forest; as a consequence Coastal Douglas-Fir [*Pseudotsuga menziesii* (Mirb.) Franco var. *menziesii*] is the dominant seral tree species (GVWD, 1992).

With the aid of 1:10,000 forest inventory maps (GVWD, 1991, 1993), field surveys, and 1:12,000 color near-infrared (NIR) aerial photography, the moderately sloping site (15°) is well described by seven multistand conifer classes dominated by Douglas-Fir, a single deciduous class dominated by Red Alder (*Alnus rubra*), and one nonvegetated class consisting of rock outcrops, gravel roads, and a dry gravel stream bed (Table 1). Classes described as *immature* (7 and 8) consist of four sparsely vegetated sites with scattered trees growing in shallow depressions, and as small clusters adjacent to older stand edges.

Ground cover is dominated by short native grasses, low shrubs, and herbaceous vegetation which provide a visually smooth image appearance. Areas of bare soil, burn scars, and rocky outcrops are also visible. In the *young* classes (5 and 6), active thinning programs have produced four single layered, uniformly distributed

conifer stands with high crown closure, and a naturally generated deciduous stand scattered along a dry gravel creek bed. Within the four *mature* classes (1–4), a distinct “clumpy” structure exists due to wind, disease, and the self-pruning nature of Douglas-Fir which provides opportunities for a dispersed lower layer of tall shrubs (3–5 m), and clusters of immature Western Hemlock (4–6 m) [*Tsuga heterophylla* (Raf.) Sarg.]. Height appears as the dominant difference between these mature stands, most likely due to topographically related factors such as soil composition and an availability to water, and nutrients. Trees located near the more mesic valley bottom reach heights in excess of 64.5 m, while further up slope, smaller (28.5–37.5 m), equally aged, same species trees exist on more xeric sites. From a sample of 68 mature, and 140 young trees taken from a previous study in this area (Hay, 1993), a strong correlation (86.9%) exists between canopy diameter and tree height. In this site, canopy diameter is also strongly related to stems per hectare, where trees with larger crown diameters, tend to be taller, and have a lower stand density.

Image Acquisition

Compact Airborne Spectrographic Imager (CASI) data covering the study area were acquired in two bands over clear skies on 31 July 1992 at 21:45 h (GMT). CASI is a pushbroom sensor designed to operate from light aircraft and helicopters, with data capture capabilities based on

a two-dimensional frame transfer CCD array. This two-dimensional array allows for the sensor to function as both a multispectral imager (spatial mode), and an imaging spectrometer (spectral mode) sensitive in the visible and near-infrared portions of the electromagnetic spectrum (418–926 nm). In both modes, the system offers user programmable spectral bandsets (in terms of wavelength and width) with a sampling interval of 1.8 nm and a 12-bit dynamic range (padded to 16 bit). In spatial mode, the sensor provides a maximum capability of 15 programmable bands, while spectral mode offers the user a maximum of 288 bands in up to 39 different viewing directions across the swath (Gower et al., 1992).

At an altitude of 914 m above the study site, data were acquired in the red spectrum (Band 1) at 675.1–685.9 nm, and the near-infrared (NIR) spectrum (Band 2) at 734.2–746.7 nm with a nominal spatial resolution of 1.2 m in both along and across track. Bandwidth positions correspond to portions of the trough and peak of the red and NIR curves respectively. At the time of data acquisition, a Global Positioning System / Inertial Navigation System was unavailable, although data from a strap-down two-axis gyro and flux gate compass allowed the imagery to be geometrically corrected for roll and yaw error (Zacharias et al., 1994). A 330×512 pixel subimage representing the study site was selected from the image swath. No further geometric or radiometric corrections were performed, as the post-processed nadir scene visually correlated exceptionally well to the available forest inventory maps, airphotography, and field notes.

IMAGE PROCESSING

This section briefly describes the TPN and GLCM method, custom filters, and the use of semivariance to determine a best fit custom window for their analysis.

The TPN Method

The TPN method involves a hierarchical methodology that involves identifying the geographic location of individual trees, generating a tree specific shape grammar, regularizing spectral data based on the spatial association of neighboring trees in an object specific variable shaped moving window, and then employing the transformed contextual output as additional logic channels to assist in thematically delineating and classifying stand characteristics. A more detailed explanation of these processes follows.

Tree Peaks Defined

Based on an understanding of the spectral response curves and spatial structure of the trees imaged in the NIR band (Guyot and Guyon, 1989; Chavez, 1992; Avery and Berlin, 1992), visual interpretation of the scene both in two-dimensional image planes and as a three-

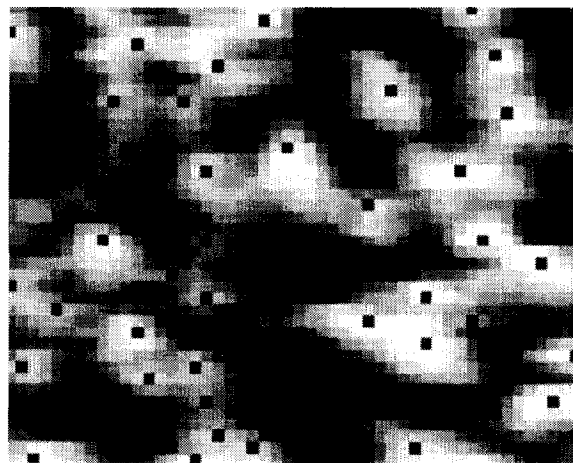


Figure 3. Magnified portion of original Band 2, illustrating peak locations (black) on individual tree crowns (lighter tones). These locations also represent pixels used in training and testing the classifier.

dimensional forest image (Fig. 1) indicates a relatively simple crown architecture, where the apex of each crown appears as the brightest pixel, or peak, per crown area. This occurs as each peak is the most directly lit, and less likely to contain shading (Gougeon, 1995); therefore, NIR peak locations (equivalent to the highest digital value) within crowns are considered geographically representative of crown centers.

Rather than manually identifying the spatial location of each tree composing the scene, a moving local maxima search applied to the original NIR image was used to automatically define the central pixel (within a 3×3 pixel kernel) as a peak if and only if it was surrounded by pixels representing lower digital values. This kernel size represents the smallest nonbiased contextual neighborhood available for analysis. Although others (Watson et al., 1985) have used spatial filters (i.e., average or gaussian) to smooth the data before peak searching, visual analysis of peak locations graphically overlaid on the original NIR image plane, indicate an excellent fit when derived from a 3×3 kernel search of the original non-smoothed NIR data (Fig. 3). Averaging filters (3×3 and 5×5) applied to the imagery, then assessed for peak locations within 3×3 and 5×5 kernels, resulted in a poorer visual correlation of peak locations when compared to corresponding crown locations.

The peak finding algorithm automatically discriminated 10,925 individual peak locations from the 330×512 NIR band. To ensure these locations corresponded only to living vegetation, a *vegetation specific mask* was generated from a normalized difference vegetation index (NDVI) image of the red and NIR CASI data. From this mask, only live tree peak locations were extracted for further analysis. An NDVI image is the result of a band

ratioing technique which relies on the inherent differences between healthy green vegetation, dry soil, and senescent vegetation in the visible and near infrared portions of the electromagnetic spectrum (Jensen, 1986). Healthy green vegetation will generally yield high index values due to relatively high NIR and low red reflectance, while similar reflectances in both bands from rock, bare soil, and dead vegetation result in values near zero (Avery and Berlin, 1992). Since NDVI statistics range from -1 to $+1$ (Tucker, 1979), the resulting NDVI image plane values were linearly scaled to 16-bit data, then clustered into 20 classes by an unsupervised K-means classifier (PCI, 1994). The resulting classes, representing various states of photosynthetically active vegetation (Tucker, 1979; Wiegand et al., 1991), were then manually aggregated to form a binary classification scheme representing living vegetation versus nonvegetated areas of reflectance based on field notes and aerial photography. The resulting binary image was then converted into a bitmap, where 5247 vegetation specific peak locations—ideally, representing individual trees—were extracted from the (168,960 pixel) image.

Image Texture Primitives

The task of modeling the texture of trees in a forest was originally visualized as feeling the grains of sand on sandpaper. As a hand moved over the grainy surface, the skin on the hand would deform to fill the voids between the grains, producing a perception of texture. As texture does not exist for a single point, and since at least three points are required to define a plane to model the forest / sandpaper surface, Delaunay triangulation, created to determine the most compact division of space of successive points on a map (McCullagh and Ross, 1980; Davis, 1986; Tipper, 1991) was employed as a shape grammar to regionalize the newly created “peak” surface. Shape grammars are high-level primitives that closely correspond to the elements that compose the texture (Ballard and Brown, 1982). By triangulating all 5247 crown centers (and 90 periphery points included at 10-m intervals to define the site perimeter), 10,574 triangular regions were generated. Taken together, these facets produce a surface with a one-to-one correspondence to the underlying image surface, visually represent the unique spatial arrangement and density differences among groups of three trees, and for the purpose of this study are considered *image texture primitives* (TPs) (Fig. 4).

By applying a point in polygon routine (Tipper, 1991) to the spectral data beneath the triangulated surface, three different algorithms were applied to each TP. The first provided an area measure, the second, the average spectral (grey-tone) value, and the third, a measure of the spectral variance of the pixels located within each TP. Variance and average calculations were based on standard algorithms (Davis, 1986). For each algorithm

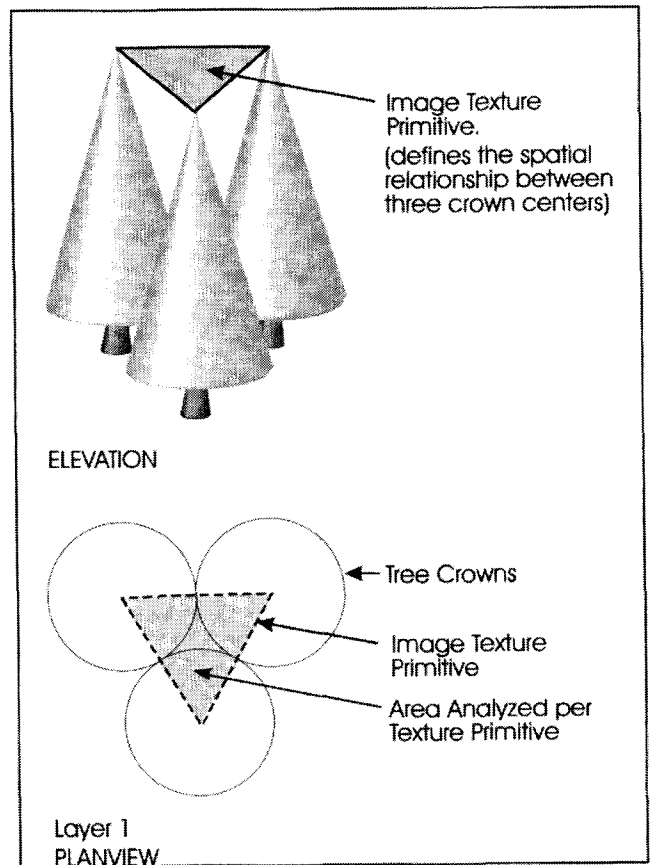


Figure 4. Image texture primitives (TPs) described in relation to tree crowns and area analyzed.

used, the single value derived from each TP was then assigned back to its component pixels, essentially normalizing each primitive based on the pixels within its boundary. This computation represents an analysis of how three parts of three trees and the nontree objects (the background) between crowns interact (Fig. 4).

To model how all portions of these three trees spatially interact with their neighbors, a topological search was initiated on a single (seed) TP, over each of the newly generated normalized datasets (described above). This required locating all TPs sharing at least one common vertex (Fig. 5), and averaging their normalized values (including that of the seed) and assigning it back to the seed TP. This process was then repeated over all remaining TPs. By so doing, each topological search identified a unique number of different-sized and -shaped image texture primitives, which, as a topologically defined neighborhood, combine to form a unique, variable shaped object-specific kernel, with a shape, size, number of neighbors, and digital value based on the structural scene characteristics of the tree crowns, as modeled by the high spatial resolution imagery. This analysis resulted in three neighborhood data sets per channel of imagery: triangulated area (TArea), triangulated variance [TV (1-

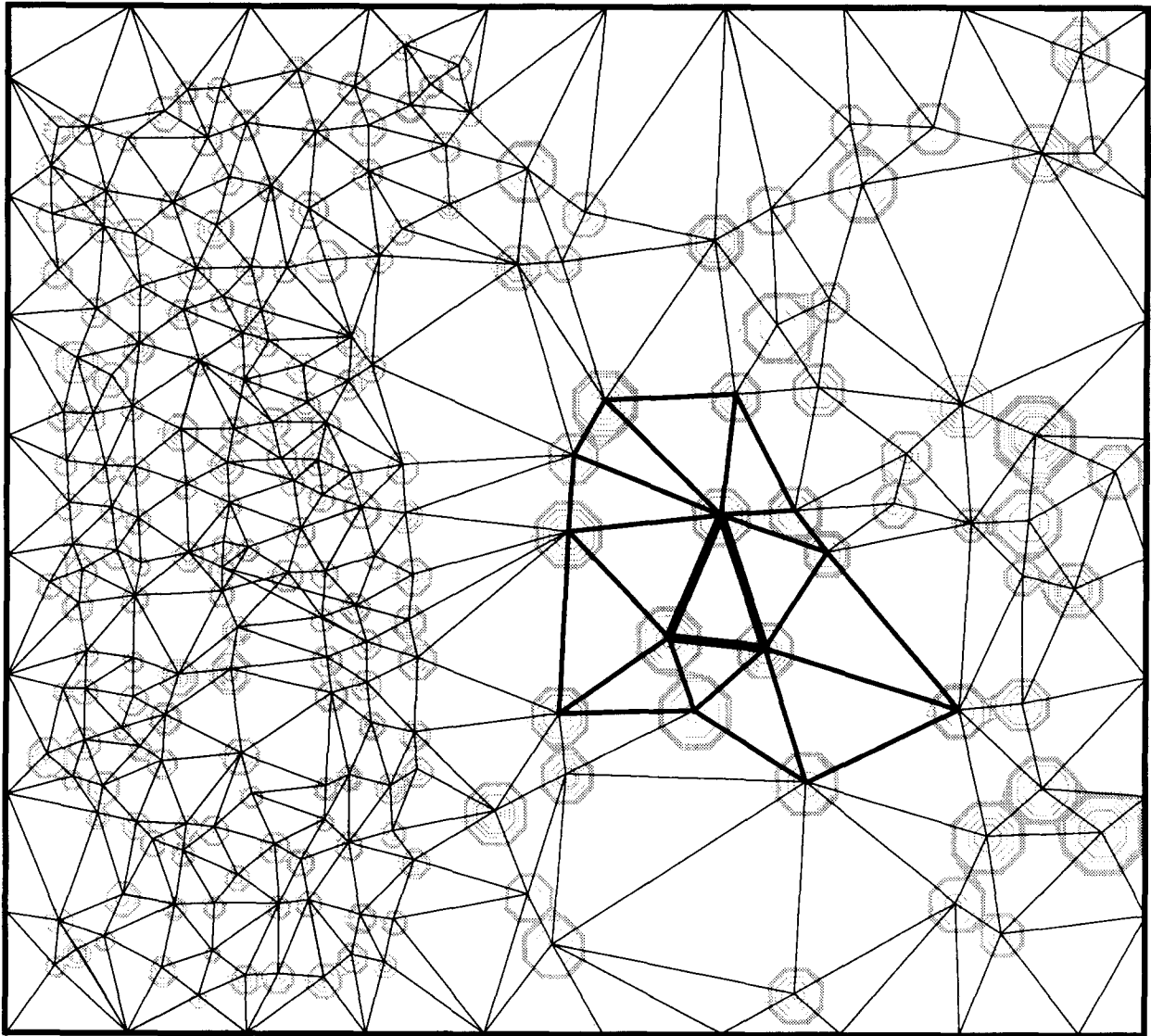


Figure 5. Within an image texture primitive (TP) neighborhood (defined by the darker triangular boundaries), structural information assigned to TPs is averaged, and then reassigned to the central TP (seed). Octagonal shapes model the varying size and location of individual tree crowns.

3)] and triangulated average [TA (1–3)]. The area dataset remains constant, regardless of the spectral channel used, as area measures are based on NIR crown locations due to a higher information content in this channel (Chavez, 1992). These resulting data sets were then used as additional logic channels in classification strategies. Triangulation was performed using commercially available software, while TP, point-in-polygon, and peak-finding analysis was generated from in-house software written by the authors in the C programming language.

Semivariance, GLCM, and Custom Filters

To provide an objective approach for selecting the best fit GLCM and custom filters, measures of semivariance were calculated for 14 vertical transects over the nine

classes within each of the three images (total 42 transects). These measures were then employed as a guide for selecting the most appropriate window sizes for analysis (Franklin and McDermid, 1993). Due to restrictions imposed by the semivariance program used, all further analysis, involving TPN and custom filters, was performed on 8-bit linear transformations of the original data.

Semivariance originates from the geostatistical theory of regionalized variables (after Matheron, 1971) where it is considered a measure of the degree of spatial dependence between samples along a specific support (Davis, 1986). From a remote sensing perspective, semivariance may be used to estimate the variability of radiance values as a function of their spatial separation (Woodcock et al., 1988b; Cohen et al., 1990; McGwire

et al., 1993). This function, described by the variogram $\gamma(h)$ (Curran, 1988; Oliver et al., 1989), is a graphical representation used to provide an unbiased description of the scale and pattern of spatial variability within the sample, and may be formally calculated as

$$\gamma(h) = \frac{1}{2(n-h)} \sum_{i=1}^{n-h} [Z(x_i + h)]^2, \quad (1)$$

where h is the lag (a vector of specific direction and distance) over which semivariance (γ) is measured, n is the number of observations used in the estimate of $\gamma(h)$, and Z is the spectral value at spatial position (x_i) denoted in one, two, or three dimensions (Cohen et al., 1990; Oliver et al., 1989). It is important to note that the accuracy of $\gamma(h)$ decreases with increasing lag because n decreases with increasing h ; therefore, the maximum lag used in this research to calculate $\gamma(h)$ was one-third the total transect length (Webster, 1985).

When calculated over a specific support, the shape of the variogram tends to approximate one of three basic models, with characteristic shapes defined by the range, sill, and nugget (Curran, 1988). Although techniques exist to fit observed semivariance to such models (Woodcock et al., 1988a; Oliver et al., 1989; Jupp et al., 1989), and thereby obtain quantitative measures of form, they are beyond the intent of this research. Of primary importance in this study were the range and sill measures. The range is an estimate of the distance (lag) at which the maximum variability (the sill) is reached, and can be used as a measure of the spatial dependency of image elements as they relate to scene structure (Jupp et al., 1988; Woodcock et al., 1988a; Cohen et al., 1990; van der Meer, 1994).

The GLCM method may be defined as a matrix of relative frequencies in which two neighboring pixels, separated by distance d and angle a , occur within a fixed sized moving kernel, one with grey tone i , the other with grey tone j (Haralick et al., 1973). For excellent reviews see Haralick (1979) and Marceau (1989). Though considered one of the most powerful image-texture methods (Weszka et al., 1976; Gu et al., 1990; Du Buff et al., 1988; Gong et al., 1992), it assesses image-texture based on outdated psychophysics theory (Hay and Niemann, 1994), and it requires important user decisions which traditionally have been selected in an ad hoc site specific manner (Marceau et al., 1990), making replication at different sites difficult. Such decisions include selecting the spectral bands for analysis, the quantization level of the image (which determines matrix size), the size of the moving window, the interpixel distance and angle during cooccurrence computation, and the texture statistics used.

The custom filter used in this analysis was a common *low-frequency* or *low-pass* spatial filter (Jensen, 1986) that involved calculating the average of the grey tone values within a fixed sized window and assigning this value to

the center pixel. The window is then moved one pixel, and the process is iterated over the entire data set, until a new regularized image is produced. The images generated by this process will be referred to here as *low pass channels*.

Based on a visual analysis of the variograms produced, a custom window size of 9×9 was selected for the NDVI and red channels and 7×7 for the NIR channel. Within these static sized kernels, 12 GLCM channels were generated, as were three low-pass channels. GLCM analysis was conducted on 6-bit linear transformations of the original 16-bit data using measures of angular second moment (ASM)—a measure of homogeneity, entropy (ENT)—a measure of randomness, inverse different moment (IDM)—a measure of lack of variability, and variance (VAR)—a measure of variability. Quantization level scaling was required to reduce data storage, and increase computational efficiency during matrix calculations. The distance between pixels during matrix computations remained constant at 1, and the average of the four main interpixel angles (0° , 45° , 90° , and 135°) was used, based on the assumption that no cover exhibited a preferential directionality. The resulting 12 channels were then linearly transformed to 8-bit data for analysis. Initial evaluation of these channels (Table 2) revealed a strong correlation between ASM, ENT, IDM, and VAR. Since VAR (see algorithm 2) well represented the output from the other three algorithms, and in order that the dimensionality of similar data be reduced, measures of variance were retained, while the other measures were discarded from further analysis:

$$\text{GLCM variance} = \frac{\sum_i \sum_j P^2(i,j) - [\sum_i \sum_j P(i,j)]^2 / R}{R}, \quad (2)$$

where P is the spatial cooccurrence matrix and R is the frequency normalization for the selected orientation.

CLASSIFICATION

A supervised multispectral classification was performed using individual tree peaks as training and test data. Tree peaks were automatically defined by the peak finding algorithm on the NIR channel, and then randomly selected from within masked areas corresponding to the different forest classes described in Table 1. Masked areas ensured peaks were selected from the interior of each polygon, rather than from stand edges where reflectances from different vegetation classes could influence the signal, and individual peak selection ensured training data was not biased by contiguous pixel selection within a single image-object. A total of 659 individual peaks were selected for classes 1–6. Due to limited numbers of trees composing each class, 38 class 1, 63 class 4, and 41 class 6 trees were available for training. Because of their small size, individual trees representing classes 7 and 8, were

Table 2. Correlation analysis of GLCM-VAR channels to ASM, ENT, and IDM

CHANNELS	NIR-ASM	NIR-ENT	NIR-IDM	NIR-VAR	NDVI-ASM	NDVI-ENT	NDVI-IDM	NDVI-VAR	R-ASM	R-ENT	R-IDM	R-VAR
NIR-ASM	1.00											
NIR-ENT	-0.89	1.00										
NIR-IDM	0.53	-0.55	1.00									
NIR-VAR	0.91	-0.91	0.53	1.00								
NDVI-ASM	-0.16	0.16	-0.19	-0.18	1.00							
NDVI-ENT	0.12	-0.12	0.17	0.15	-0.95	1.00						
NDVI-IDM	-0.16	0.16	-0.11	-0.17	0.82	-0.80	1.00					
NDVI-VAR	-0.16	0.16	-0.19	-0.18	1.00	-0.95	0.82	1.00				
R-ASM	0.34	-0.45	0.24	0.35	0.28	-0.34	0.22	0.28	1.00			
R-ENT	-0.36	0.46	-0.23	-0.36	-0.27	0.35	-0.17	-0.27	-0.93	1.00		
R-IDM	0.16	-0.24	0.26	0.19	0.29	-0.29	0.28	0.28	0.72	-0.63	1.00	
R-VAR	0.34	-0.45	0.24	0.35	0.28	-0.34	0.22	0.28	1.00	-0.93	0.72	1.00

difficult to visually discern within the imagery, yet more than 800 vegetation specific peak locations were computationally defined within these class areas suitable for training and testing. Suitability for all training and test data was based on field experience, field maps, and aerial photography.

Due to its robustness and ubiquity, a maximum likelihood classifier (MLC) was applied to various spatial and spectral channel combinations (Gong et al., 1992). This classifier assumes that training data statistics for each class in each band are gaussian in nature. When this occurs, MLC is well suited for providing optimal classification accuracies using a limited number of variables (Peddle, 1993). Accuracy was evaluated visually from thematic map output, and from confusion matrices derived from independent test data. Test data contained over 70 samples per class, except for classes 1, 4, and 6, which (due to a limited number of trees) had 34, 40, and 25 samples, respectively. In all classes, test and training data conform to standard sample size recommendations (Congalton, 1991). Two measures of accuracy were defined, user's and overall (Congalton, 1991). The "user's accuracy" is a measure of "commission error" or "reliability," indicative of the probability that a pixel classified on the map/image actually represents that category on the ground, while the "overall accuracy" is a ratio of the total correct (sum of main diagonal) by the total number of pixels in the error matrix. The null class value was also included to provide a measure of how well the various channel combinations actually defined the study area. Null class values represent the areal percentage of the imagery unable to be classified under the current classification scheme.

RESULTS AND DISCUSSION

Variogram Analysis

Although a significant body of theory supports the use of semivariance for structural analysis, it is important to

recognize that a generous amount of user subjectivity is involved in the generation and visual analysis of variograms. Our experience with semivariance in this complex forest environment suggests that at least two or three (if possible multidirectional) transects per class, or group of classes, be selectively chosen, as several identical length transects located over the same class, yet separated by only one or two pixels, can produce entirely different measurements. Such anisotropic results may be due to structural differences within the canopy, illumination, viewing angle, radiometric noise, and/or calibration error in the imagery.

Visual analysis of the Band 2 variogram ranges for each class (Fig. 6) strongly represent field measures of crown diameter, although it should be recognized that selection of a unique class range can be complex, as the range for a single class varies depending on the wavelength assessed. For example, a range of five pixels represents class 1 and Band 1, while a range of eight pixels represents the same class in Band 2. Unique measures for each class indicate that different optimal kernel sizes are required to properly assess the structural differences between each class. Analysis of the TPN area channel (TArea) over the transect locations at which variograms were sampled (Table 3) also indicates that a large amount of spatial variation occurs within each class, which is not adequately described by semivariance, or a single kernel size. In the selection of custom kernel sizes for GLCM and low-pass channels, emphasis was placed on measures which sampled three to four classes along the entire length and width of the image, rather than on a single class, as multiclass sampling more fully represented the entire image structure.

Though range was the primary element used in kernel selection, the height of the variograms also provide interesting structural information. By ranking all classes in both channels based on their sill height, Band 1 variograms appear inversely related to crown closure and the amount of understory, while Band 2 variograms are re-

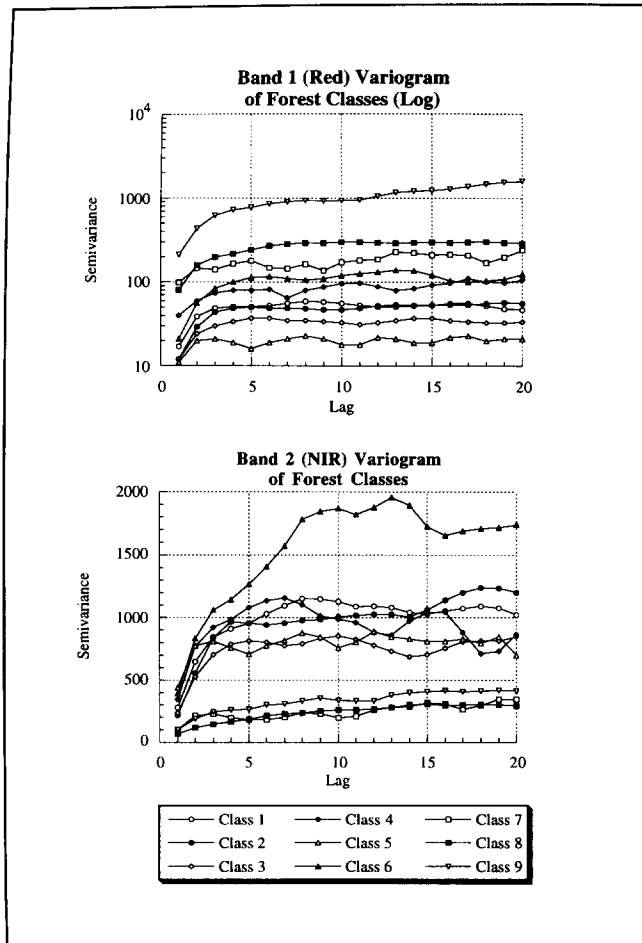


Figure 6. Variograms describing forest classes within Bands 1 and 2.

lated to canopy height, and age, with obvious discrimination potential between deciduous and coniferous stands (Fig. 6).

Within Band 1, dry soils, gravel, and dead vegetation have a higher reflectance than healthy vegetation; therefore, stands with low crown closure and low stand density will have a higher measure of semivariance than more densely vegetated stands, with a lower visible soil reflectance. This is illustrated by the red variograms, where class 9 (nonveg) has the highest measure of semivariance with a crown closure of 0.0%, followed by classes 8, 7, 6, 4, 1, 2, 3, and 5 (young fir), which has the lowest measure of semivariance, yet the largest measure of crown closure (66–75%). When Band 2 conifer classes are ranked, class 4 is first, followed by classes 1, 2, 3, 5, 7, and 8, which corresponds strongly to site measures of height and age as defined in Table 1. In both sets of variograms, class 4 appears out of rank when compared with its structural characteristics (Table 1). Even though this class is found on forest inventory maps, its variogram height, range, and subsequent poor classification results

suggest that this class designation may be more artificial than real.

Classification Results

Various feature combination strategies have been tested on 16 different channels with channel combinations kept to a maximum of seven, as multivariate normality assumptions for the MLC tend to be violated by higher dimension datasets, resulting in a reduction of classification accuracy (Peddle, 1993; PCI, 1994). The 16 bands analyzed are represented in Table 4 by three main groups, with Table 5 providing an overview of 29 specific channel combinations. Classification accuracy represents the effect, or relative importance of specific logic channels. Measures of user accuracy per class and overall accuracy are provided, as is a null class measure, to present the reader an unbiased summary of results. The first group in Table 4 contains the two original CASI images, a single NDVI channel, and three lowpass channels. The second group contains three GLCM variance channels generated from custom kernels, and the third group contains three TPN channels generated from an average measure of TPs within a variable sized and shaped neighborhood; three channels generated from a grey-level variance measure; and a single channel (TArea) representing the average area within these unique neighborhoods. TArea (Fig. 7) is a unique measure that models the (area) spatial relationship between three trees and their nearest neighbors, thus representing the crown density of all vegetation large enough to be discriminated by the sensor IFOV imaged in the NIR wavelength. Dark tones correspond to three trees within a small area, thus a high crown density, while brighter tones represent three trees within a larger area, thus a lower crown density.

The two CASI bands alone (C1-2) yield an overall

Table 3. Comparison of variogram and TArea distance measures

Class	* §		* §		
	Red Lag	NIR Lag	TArea min	Transects max	Transects avg
1	5	8	4	18	11.1
2	5	4	5	21	11.2
3	5	5	5	17	10.3
4	5	7	3	27	8.5
5	3	3	2	9	4.7
6	6	10	4	19	10.11
7	5	3	57	255	150.6
8	10	15	3	92	33.9
9	10	9	5	27	14

* All values represent numbers of pixels.

§ Sample size (n) = 100 pixels.

Table 4. Key to symbols used in classification results

Symbol	Description	
C1	CASI	- Band 1: (675.1- 685.9 nm) Red
C2		- Band 2: (734.2 -746.7 nm) NIR
C3		- NDVI from C1 & C2. See algorithm (1)
CA1		- Average, 9 x 9 kernel, CASI Band 1
CA2		- Average, 7 x 7 kernel, CASI Band 2
CA3		- Average, 9 x 9 kernel, NDVI
GV1	GLCM	- Variance, 9 x 9 kernel, CASI Band 1
GV2		- Variance, 7 x 7 kernel, CASI Band 2
GV3		- Variance, 9 x 9 kernel, NDVI
TA1	TPN	- Average spectral value, CASI Band 1
TA2		- Average spectral value, CASI Band 2
TA3		- Average spectral value, NDVI
TArea		- Average area measure
TV1		- Variance, CASI Band 1
TV2		- Variance, CASI Band 2
TV3		- Variance, NDVI

accuracy of 65%, for the nine classes. With an addition of the NDVI channel (C3), overall accuracy decreased 4%, accompanied by a 20% increase in the null class, indicating classifier confusion. Further analysis and enhancement of the NDVI channel revealed image artifacts resulting from roll correction, and the additive characteristics of noise found in ratio images (Lillesand and Kiefer, 1987). Roll correction artifacts could not be removed as preroll corrected imagery were unavailable. Though moderately noisy, this channel was retained, and used in subsequent filtering strategies which reduced its noise

component while offering unique vegetation information. On their own, the three low-pass (CA1-3), GLCM variance (GV1-3), TPN average (TA1-3), and the three TPN variance channels (TV1-3) produced an overall accuracy of 70%, 35%, 69%, and 40%, respectively. When TArea was added to CASI Bands 1 and 2, accuracy increased 5% to an overall accuracy of 70%. When applied to the low-pass, the GLCM variance, the TPN average, and the TPN variance channels, overall accuracy increased 4% to 74%, 39% to 74%, 9% to 78%, and 4% to 44%, respectively.

On its own, the low-pass technique combined with custom sized kernels reduces the H-resolution problem by increasing classification accuracy to 70% with a null class of 8%. With the addition of the GLCM channels, overall accuracy dropped 6%, and the area of the image unable to be classified increased to 31%. In all cases, the addition of the original CASI and NDVI channels, and the GLCM and TPN variance channels—in various combinations—tends to result in lower classification accuracies and higher null-class percentages. In several cases where these channels were combined with TArea, overall accuracy was over 70%, but the null class was greater than 30%. With the addition of TArea to all but one combination, overall class accuracy improved, and the percent null class tends to be reduced.

The importance of spatial information within the H-resolution imagery is particularly apparent from the combination of GLCM and TArea with an overall accuracy of 74%, and TPN Average and TArea which ties for the highest overall accuracy of 78%, with a combination of the red and NIR-TPN average channels, the NIR and NDVI-TPN variance channels, and the TPN area channel. In these three cases the lowest null class measure is 2%.

Tree Top Sampling

It has been suggested (Franklin and McDermid, 1993) that selecting tree-only pixels will result in training data which does not reflect the true nature of the stand, thus providing poor performance in larger area mapping and or classification. This has not proven to be the case in this research. Classification of the original CASI data based only on two bands and single tree-top pixels for eight classes (class 9 has no tree-top pixels) has produced a reasonably high class accuracy (65%). If one considers that a single pixel actually represents a regularization of spectral data around each pixel, not just within the individual pixel, then depending on the size of the tree crown in relation to the area represented by a single pixel, this form of training and test data selection may indeed be representative. When used in combination with the TPN method, each tree-top pixel no longer represents a single tree crown; rather it is a regularized measure of portions of three trees, their immediate neighbors, and the visible understory between them. When combined with appropriate masks, tree top selection also offers a

Table 5. Maximum likelihood classification results

Variables	#vars	Percent of user-accuracy by class										O-All	Null
		1	2	3	4	5	6	7	8	9			
C1-2	2	73	51	49	27	58	100	88	87	80	65	7	
C1-2, TArea	3	63	59	54	39	75	94	96	95	94	70	12	
C1-2, GV1-3	5	53	62	65	36	56	86	86	80	75	61	44	
C1-2, GV1-3, TArea	6	52	66	67	41	79	92	89	88	94	64	51	
C1-3	3	80	51	46	25	53	97	90	75	80	61	27	
C1-3, TArea	4	72	59	54	33	74	97	90	87	96	67	31	
C1-3, GV1-3	6	61	66	66	34	54	82	87	76	74	57	54	
C1-3, GV1-3, TArea	7	59	65	69	43	81	91	88	87	97	60	60	
CA1-3	3	68	74	30	45	59	84	94	92	84	70	8	
CA1-3, TArea	4	78	64	53	52	77	95	93	97	88	74	14	
CA1-3, GV1-3	6	74	67	57	69	71	94	98	94	79	64	31	
CA1-3, GV1-3, TArea	7	75	67	56	52	87	93	98	97	92	61	39	
GV1-3	3	28	30	50	20	37	31	65	46	67	35	7	
GV1-3, TArea	4	91	81	79	49	81	63	76	91	81	74	4	
GV1-3, TV1-3	6	27	27	38	0	40	27	67	78	93	42	16	
GV1-3, TV1-3, TArea	7	33	49	59	19	74	19	75	95	94	50	20	
TA1-3	3	71	69	57	50	65	74	77	70	93	69	2	
TA1-3, TArea	4	77	77	62	56	83	74	87	99	87	78	2	
TA1-3, TV1-2, TArea	6	84	77	73	72	84	80	91	94	90	76	11	
TA1-3, TV1-3, TArea	7	88	76	77	56	87	75	83	95	91	72	16	
TV1-3	3	20	10	0	24	16	11	0	68	87	40	1	
TV1-3, TArea	4	18	34	0	16	49	15	53	94	81	44	2	
C2-3, TV1-3	5	70	49	0	33	48	66	90	100	80	55	34	
C2-3, GV1-3	5	56	65	66	37	58	93	88	79	68	62	46	
C2-3, TA1-2, TArea	5	77	75	68	52	88	95	95	80	74	60	60	
C2, TA1-3, TArea	5	68	78	60	52	88	89	94	100	86	75	35	
C2, TA1-3, TV2	6	79	84	67	59	90	87	94	96	90	74	53	
TA1-2, TV1-2, TArea	5	80	76	76	50	75	48	81	94	96	72	6	
TA1-2, TV2-3, TArea	5	82	76	81	69	78	94	97	96	83	78	5	

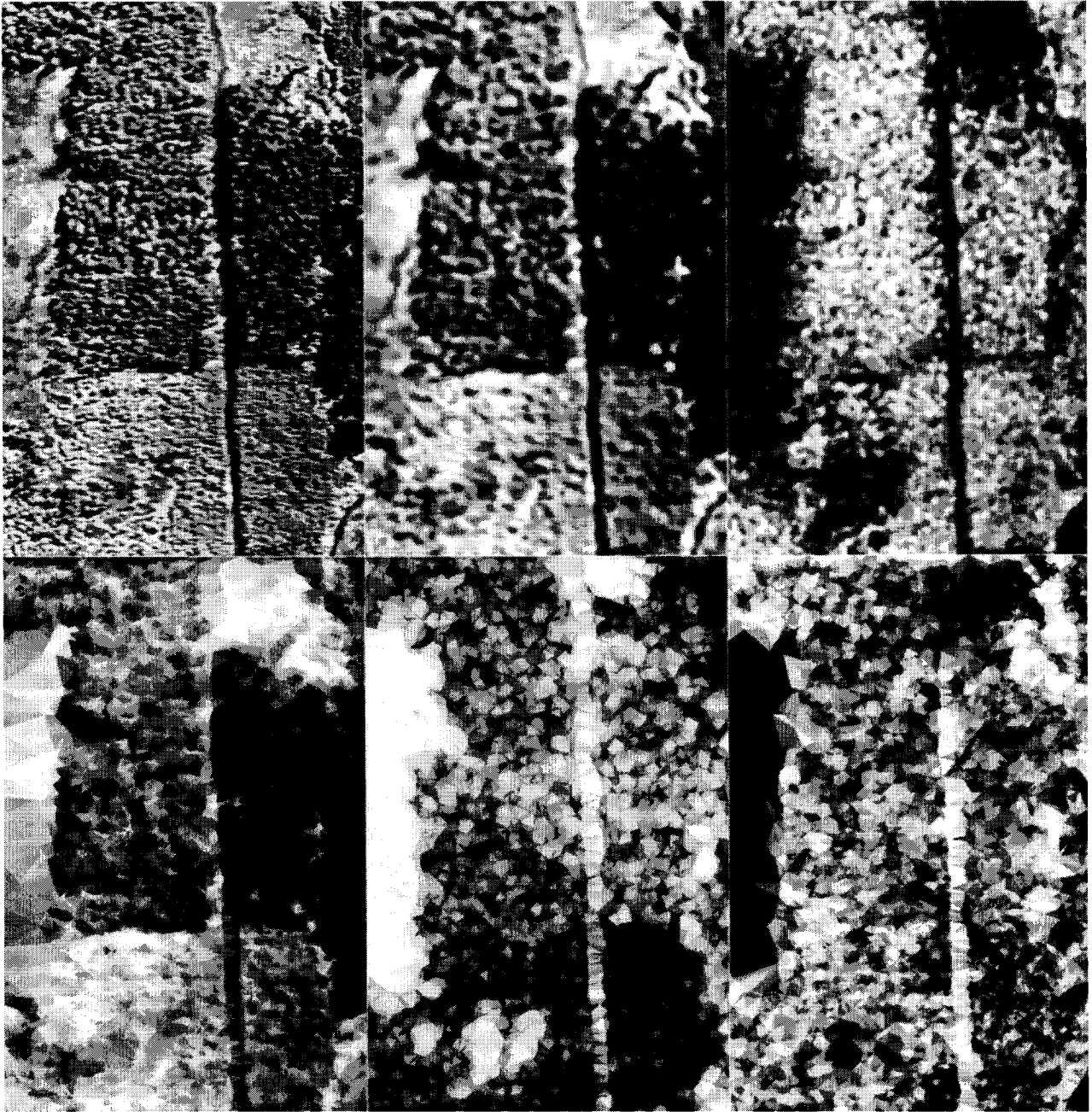


Figure 7. Study site mosaic of NIR channels used for classification. Each image represents a 330×512 scene composed of 1.2 m^2 pixels. Beginning at top left, to bottom right, CASI Band 2 (C2), lowpass (CA2), GLCM variance (GV2), TPN average (TA2), TPN Area (TArea), TPN Variance (TV2).

relatively straight forward, partially automated technique for selecting training and test data within H-resolution imagery.

Though seldom discussed in the literature, selecting appropriate training and test data from H-resolution imagery are considerably more difficult than selection within L-resolution imagery. In selecting an H-resolution image-object for training and / or testing, which appears to the user to occupy a small area, it is, in fact, more likely to be composed of many pixels; thus the problem

of inadvertently biasing the classifier to a unique image object within a class is possible, resulting in poor class accuracies.

CONCLUSIONS

The principal objective of this study was to evaluate the functionality and advantages of the TPN method, as an image-texture descriptor, capable of evaluating the spatial interaction of individual trees within a complex H-

resolution remotely sensed CASI image. Based on an understanding of human- and image-texture analysis, and an evaluation of the classification results of forest classes generated by the TPN method against those derived from a combination of semivariance, custom filters, and the GLCM method, the following conclusions are made:

- Based on MLC results of 29 different channel combinations, the TPN method produces a higher overall classification accuracy (78%) than classifications generated from customized GLCM (35%) and low pass filters (70%)—either individually or in combination with other channels.
- Texture represents a quantification of the spatial variation of primitives as modeled by the contextual limits of the sensor. For humans, this is achieved by a dynamic process, which incorporates the unique spatial variation of different-sized and -shaped objects constituting a scene's texture—as different object scales and object combinations are potentially resolvable within a single scene, each capable of producing unique textures. In digital remote sensing, texture analysis traditionally occurs within a static sized moving window, disregarding the concept of multiple object scales.
- Spatial filtering is an effective method of reducing the H-resolution problem, with object-specific, variable-shaped, and -sized kernels increasing classification accuracy (8–42%) over customized static sized kernels.
- TArea is a unique structural measure which models the crown density of all resolvable trees. Due to a strong correlation (86.9%) in this study area between canopy diameter and tree height, and how these characteristics relate to tree age and stems/ha, the addition of this logic channel improves overall accuracy (4–35%) among classes which differ primarily in height and crown diameter.
- The need for variable sized and shaped kernels capable of resolving the unique spatial structures within natural scenes is supported by semivariance and TArea measures (Table 3), which indicate that structural variance within and between classes is high, but cannot be accounted for by a static sized kernel.
- Although semivariance provides unique structural information, generation of meaningful data is highly subjective and not yet capable of being fully incorporated or exploited in remote sensing analysis. By generating GLCM and spatial filter measures, the user is also faced with a variety of subjective decisions, the most important of which involves selecting an appropriate

static window size to evaluate the spatially varying structure within the scene. In the TPN method this important decision is automatically evaluated based on the spatial relationship of crown locations. The only subjective decisions involve selecting a channel for defining crown centers, and the relatively trivial task of generating a binary vegetation versus nonvegetated, bitmap.

- Automatic peak selection offers an effective and efficient method for selecting training and test data from H-resolution data, which removes the ability of the user to bias the classifier to a unique image object within a class as only a single pixel per image object is selected.

The TPN method is a novel image-texture technique, capable of evaluating the spatial interaction of individual trees within a complex H-resolution remotely sensed CASI image. This method provides a reduction of user required subjective decisions, the automation of an object-specific variable sized kernel, the capability of exploiting the different object scales present in H-resolution forest imagery to provide object-rich, yet variance-reduced datasets for mitigating the H-resolution problem, and an higher classification accuracy than those derived from a combination of semivariance, custom filters, and the GLCM method.

This research was supported by scholarships and grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), Forestry Canada, and a Presidents Research Scholarship from the University of Victoria. Special thanks is extended to Gordon Joyce and Peter van de Mortel at the GVWD for access to the site and helpful field assistance, to Gregory J. McDermid for the semivariance code, and to Ken Josephson and Ole Heggin for assistance with graphics.

REFERENCES

- Avery, T. E., and Berlin, G. L. (1992), *Fundamentals of Remote Sensing and Airphoto Interpretation*, 5th ed., MacMillan, p. 472.
- Ballard, D. H., and Brown, C. (1982), *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ, 519 pp.
- Bergen, J. R., and Adelson, E. H. (1988), Early vision and texture perception, *Nature* 333:363–367.
- Bergen, J. R., and Landy, M. S. (1991), Computational modeling of visual texture segregation, in *Computational Models, Visual Processing* (M. S. L. Movshow and J. A. Movshow, Eds.) MIT Press, Cambridge, pp. 253–271.
- Chavez, P. S., Jr. (1992), Comparison of spatial variability in visible and near-infrared spectral images, *Photogramm. Eng. Remote Sens.* 58(7):957–964.
- Clark, C. D. (1990), Remote sensing scales related to the frequency of natural variation: an example from Paleo-ice-flow in Canada, *IEEE Trans. Geosci. Remote Sens.* 28(4):503–508.
- Cohen, W. B., and Spies, T. A. (1992), Estimating structural

- attributes to Douglas-fir / western hemlock forest stands from Landsat and SPOT imagery, *Remote Sens. Environ.* 41: 1–17.
- Cohen, W. B., Spies, T. A., and Bradshaw, G. A. (1990), Semivariograms of digital imagery for analysis of conifer canopy structure, *Remote Sens. Environ.* 34:167–178.
- Congalton, R. G. (1991), A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sens. Environ.* 37:35–46.
- Curran, P. J. (1988), The semivariogram in remote sensing: an introduction, *Remote Sens. Environ.* 24:493–507.
- Cushnie, J. L. (1987), The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies, *Int. J. Remote Sens.* 8(1):15–19.
- Davis, J. C. (1986), *Statistics and Data Analysis in Geology*, 2nd ed., Wiley, New York, 646 pp.
- Du Buff, J. M. H., Kardan, M., and Spann, M. (1988), Texture feature performance for image segmentation, *Pattern Recognition* 23(3/4):291–309.
- Franklin, S. E., and McDermid, G. J. (1993), Empirical relations between digital SPOT HRV and CASI spectral response and lodgepole pine (*Pinus contorta*) forest stand parameters, *Int. J. Remote Sens.* 14(12):2331–2348.
- Gong, P., Marceau, D. J., and Howarth, P. J. (1992), A comparison of spatial feature extraction algorithms for land-use classifications with SPOT HRV data, *Remote Sens. Environ.* 40: 137–151.
- Gougeon, F. A. (1995), Comparison of possible multispectral classification schemes for tree crowns individually delineated on high spatial resolution MEIS images, *Can. J. Remote Sens.* 21(1):1–9.
- Gower, J. F. R., Borstad, G. A., Anger, C. D., and Edel, H. R. (1992), CCD-based imaging spectroscopy for remote sensing: the FLI and CASI programs, *Can. J. Remote Sens.* 18(4): 210–223.
- Gu, Z. Q., Duncan, C. N., Renshaw, R., Mugglestone, M. A., Cowan, C. D. N., and Grant, P. M. (1989), Comparison of techniques for measuring cloud texture in remotely sensed satellite meteorological image data, *IEE Proc.* 136(5):236–248.
- Guyot, G., and Guyon, D. (1989), Factors affecting the spectral response of forest canopies: a review, *Geocarto Int.* 3:3–18.
- Greater Victoria Water District (GVWD) (1991, 1993), *Greater Victoria Water District Watershed Management Forest Cover Classification*, Sooke Lake Watershed. Scale 1:10000, Hugh Hamilton Ltd., Vancouver, British Columbia.
- GVWD (1992), *Greater Victoria Water District Land Use Technical Review Summary Report*, TERRASOL Environmental Consulting Division, 97 pp.
- Haralick, R. M. (1979), Statistical and structural approaches to texture, *Proc. IEEE* 67(5).
- Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973), Textural features for image classification, *IEEE Trans. Syst. Man Cybern.* SMC-3 (Nov.):610–621.
- Hay, G. J. (1993), Visualizing 3-D texture: a three dimensional structural approach to model forest texture, M.Sc. thesis, University of Victoria, 1993 (unpublished).
- Hay, G. J., and Niemann, K. O. (1994), Visualizing 3-D texture: a three dimensional structural approach to model forest texture, *Can. J. Remote Sens.* 20(2):90–101.
- Heller, M. A. (1989), Texture perception in sighted and blind observers, *Perception Psychophys.* 45(1):59–54.
- Horn, B. K. P., and Brooks, M. J., Eds. (1989), *Shape from Shading*, MIT Press, Cambridge, 577 pp.
- Hubel, D. H., and Wiesel, T. N. (1968), Receptive fields and functional architecture of monkey striate cortex, *J. Physiol.* 195:215–243.
- Irons, J. R., Markham, B. L., Nelson, R. F., et al. (1985), The effects of spatial resolution on the classification of Thematic Mapper data, *Int. J. Remote Sens.* 6(8):1385–1403.
- Jarvis, R. A. (1983), A perspective on range finding techniques for computer vision, *IEEE Trans. Pattern Analy. Machine Intell.* PAMI-5(2):122–139.
- Jensen, J. R. (1986), *Introductory Digital Image Processing. A Remote Sensing Perspective*, Prentice-Hall, Englewood Cliffs, NJ, 379 pp.
- Jones, B., and O'Neil, S. (1985), Combining vision and touch in texture perception, *Perception Psychophys.* 37(1):66–72.
- Julesz, B. (1975), Experiments in the visual perception of texture, *Sci. Am.* 233(4):34–43.
- Julesz, B. (1981), Textons, the elements of texture perception, and their interactions, *Nature*:91–97.
- Julesz, B., and Bergen, J. R. (1983), Textons, the fundamental elements in preattentive vision and perception of textures, *Bell System Tech. J.* 62:1619–1646.
- Jupp, D. L. B., Strahler, A. H., and Woodcock, C. E. (1988), Autocorrelation and regularization in digital images. I. Basic theory, *Trans. Geosci. Remote Sens.* 26(4):463–473.
- Jupp, D. L. B., Strahler, A. H., and Woodcock, C. E. (1989), Autocorrelation and regularization in digital images. II. Simple image models, *Trans. Geosci. Remote Sens.* 27(3):247–258.
- Kerman, J. (1980), *Listen*, Worth, New York, 561 pp.
- Klatzky, R. L., Reed, C., and Lederman, S. (1987), There's more to touch than meets the eye: the salience of object attributes for haptics with and without vision, *J. Exp. Psychol. Gen.* 116(4):356–369.
- Lam, N. S.-N., and Quattrochi, D. A. (1992), On the issues of scale, resolution, and fractal analysis in the mapping sciences, *Prof. Geogr.* 44(1):88–98.
- Lederman, S. J. (1982), The perception of texture by touch, *Tactile Perception: A Sourcebook* (W. Schiff and E. Foulke, Eds.), Cambridge University Press, Cambridge, pp. 130–167.
- Lederman, S. J., Jones, B., and Thorne, G. (1986), Perception of texture by vision and touch: multidimensionality and intersensory integration, *J. Exp. Psychol. Hum. Perception Performance* 12(2):169–180.
- Marceau, D. (1989), A review of image classification procedures with special emphasis on the grey-level cooccurrence matrix method for texture analysis, ISTS-EOL-TR89-007, Earth Observations Laboratory Institute for Space and Terrestrial Science, 59 pp.
- Marceau, D. J., Howarth, P. J., Dubois, J. M. M., and Gratton, D. J. (1990), Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery, *IEEE Trans. Geosci. Remote Sens.* 28(4):513–519.
- Marceau, D. J., Howarth, P. J., and Gratton, D. J. (1994a), Remote sensing and the measurement of geographical entities in a forested environment. The scale and spatial aggregation problem, *Remote Sens. Environ.* 49:93–104.

- Marceau, D. J., Gratton, D. J., Fournier, R. A., and Fortin, J. P. (1994b), Remote sensing and the measurement of geographical entities in a forested environment: the optimal spatial resolution, *Remote Sens. Environ.* 49:105–117.
- Marr, D. (1982), *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman, San Francisco, 396 pp.
- Matheron, G. (1971), *The Theory of Regionalized Variables and Its Applications*, Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau., No. 5, Ecole Nationale Supérieure des Mines de Paris.
- McCullagh, M. J., and Ross, C. G. (1980), Delaunay triangulation of a random data set for isorhythmic mapping, *Cartogr. J.* 17(2):93–99.
- McGwire, K., Friedl, M., and Estes, J. E. (1993), Spatial structure, sampling design and scale in remotely-sensed imagery of a California savanna woodland, *Int. J. Remote Sens.* 14(11): 2137–2164.
- Oliver, M., Webster, R., and Gerrard, J. (1989), Geostatistics in physical geography. Part 1: Theory, *Trans. Inst. Br. Geogr. N. S.* 14:259–269.
- PCI (1994), *Using PCI Software*, Version 5.2 EASI / PACE Image Analysis System Manual, Perceptron Computing, Toronto.
- Peddle, D. R. (1993), An empirical comparison of evidential reasoning, linear discriminant analysis, and maximum likelihood algorithms for alpine land cover classification, *Can. J. Remote Sens.* 19(1):31–45.
- Raffy, M. (1992), Change of scale in models of remote sensing: a general method for spatialization of models, *Remote Sens. Environ.* 40:101–112.
- Slater, P. N. (1980), *Remote Sensing Optics and Optical Systems*, Addison-Wesley, Reading, MA. 574 pp.
- Strahler, A. H., Woodcock, C. C., and Smith, J. A. (1986), On the nature of models in remote sensing, *Remote Sens. Environ.* 20:121–139.
- Tipper, J. C. (1991), Fortran programs to construct the planar Voronoi diagram, *Comput. Geosci.* 17(5):597–631.
- Tucker, C. J. (1979), Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sens. Environ.* 8:127–150.
- Uttal, W. R. (1983), *Visual Form Detection in 3-Dimensional Space*, Lawrence Erlbaum, Hillsdale, NJ, 163 pp.
- Uttal, W. R. (1988), *On Seeing Forms*, Lawrence Erlbaum, Hillsdale, NJ, pp. 187–199.
- van der Meer, F. (1994), Quantification of particle shape and texture using mathematical morphology and geostatistical techniques, *ITC* 1:13–22.
- Watson, L. T., Laffey, T., and Haralick, R. M. (1985), Topographic classification of digital image intensity surface using generalized splines and the discrete cosine transformation, *Comput. Vis. Graph. Image Process.* 29:143–167.
- Webster, R. (1985), Quantitative spatial analysis of soil in the field, *Adv. Soil Sci.* 3:1–70.
- Wesza, J. S., Dryer, C. R., and Rosenfeld, A. (1976), A comparative study of texture measures for terrain classification, *IEEE Trans. Syst. Man Cybernet.* SMC-6(4):269–285.
- Wiegand, C. L., Richardson, A. J., Escobar, D. E., and Gergermann, A. H. (1991), Vegetation indices in crop assessments, *Remote Sens. Environ.* 35:105–119.
- Woodcock, C. E., and Strahler, A. H. (1987), The factor of scale in remote sensing, *Remote Sens. Environ.* 21:311–332.
- Woodcock, C. E., Strahler, A. H., and Jupp, D. L. (1988a), The use of variograms in remote sensing. I. Scene models and simulated images, *Remote Sens. Environ.* 25:323–348.
- Woodcock, C. E., Strahler, A. H., and Jupp, D. L. (1988b), The use of variograms in remote sensing. II. Scene models and simulated images, *Remote Sens. Environ.* 25:349–379.
- Zacharias, M., Borstad, G., and Kerr, R. (1994), Three dimensional geo-referencing of airborne multispectral data for use in the GIS, in *Eighth Annual Symposium on Geographic Information Systems*, Polaris Conferences, Vancouver, BC, Vol. 1. pp. 399–404.