Geoinformation for European-wide Integration, Benes (ed.)

Analysis of spatial structure in images using local variance

Peder Klith Bøcher
Danish Institute of Agricultural Sciences, Department of Agricultural Systems, Research Centre Foulum, P.O. Box 50, DK-8830 Tjele, Denmark
Phone: +45 89991812; Fax: +45 89991200; Peder.Bocher@agrsci.dk

Keith R. McCloy
Danish Institute of Agricultural Sciences, Department of Agricultural Systems, Research Centre Foulum, P.O. Box 50, DK-8830 Tjele, Denmark

Keywords: local variance, image degradation, IKONOS, Landsat ETM+, Doi Inthanon, tropical montane forest

ABSTRACT: Local variance was defined as the average of local standard deviation values computed within a 3 by 3 window moving across the entire image by steps of one pixel (Woodcock & Strahler, 1987). This procedure was then repeated on successively coarsened versions of the original image to construct a graph of average local variance values. It was shown that the average local variance curve reaches a peak value at a specific resolution that is related to the size of scene objects, and then decreases again. This relation was reported to be at pixel sizes of ½ to ¾ of the object size.
The aim of this paper is to report on work being done to investigate the characteristics of local variance, and assess the significance for the interpretation of high-resolution multi-spectral forest images in Thailand. For this purpose four synthetic images were constructed, which were subsequently subjected to analysis by method of local variance.
We found numerous peaks curves related to scenes with only one object. One significant key peak at pixel size equal to object size plus a number of auxiliary peaks probably resulting from autocorrelation in the scene. For scenes composed of two object classes of distinct sizes, we found a key peak for each object at pixel sizes equal to object sizes. Auxiliary peaks were also found in these complex scenes.
We conclude that using a 2 by 2, moving window enhanced the essential peaks on the curve, making this a more informative method than the 3 by 3 window.
We found that all possible aggregated resolution levels are important for interpreting the local variance curve, not a subset of them. This was supported by the findings of numerous peaks as contrast to the one-peak curves reported by Woodcock and Strahler (1987).

1 INTRODUCTION

The results presented in this paper are part of an investigation into the potential use of remote sensing in mapping of a montane forest in northern Thailand. The work was conducted within the joint research programme *Forest and People in Thailand* under the auspices of the Danish Environmental Research Programme

1.1 *Study area*

The study area is the tropical montane cloud forest sited above 2,000 meters altitude within the Doi Inthanon National Park in Northern Thailand (figure 1).

Tropical montane cloud forests are dependent on a frequent low-hanging cloud cover (Hamilton et al., 1995; Santisuk, 1988).

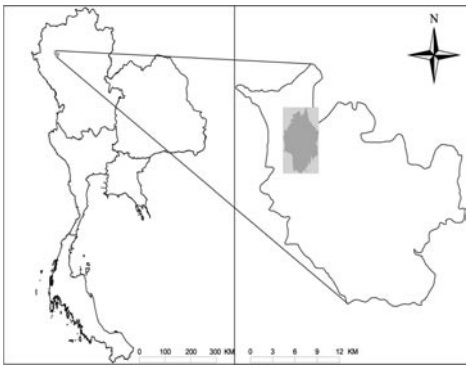

Figure 1. Doi Inthanon National Park is located in the northwestern part of Thailand, not far from the Burmese border. This study is restricted to the tropical montane forest above 2,000 meters altitude within the national park.

Like all of Thailand, the study site is influenced by a monsoon climate. This climate is characterised by distinct hot dry and rainy seasons. Three seasonal periods can be identified. The rainy season is from May to September when humid air masses are brought by the S-W monsoon. The N-E monsoon lasts from October until February and brings relatively cool, dry air. From late February the N-E monsoon declines causing the onset of the hot-dry season, which lasts until the onset of the S-W monsoon in late April. The hot dry season is the hottest period of the year.

Generally mean relative humidity ranges from 60% in the dry season to more than 80% in the rainy season. This scheme is of course fundamentally influenced by elevation and can vary a little in time from year to year.

1.2 *Imagery*

Four satellite images were acquired to cover the study area at two different dates and with two different sensor types. It was expected that the distinct dry season from November/December to April/May would influence the phenology of the majority of the tree species. Thus, two of the images were acquired near the end of the dry season early in 2000 and two were acquired at the beginning of the dry season in late 2000 (table 1). By this means one growing-season was covered with satellite images.

Table 1. Imagery used in this study

Satellite	Extent	Date of acquisition	Rectification basis
Ikonos	17x17 km	22nd Mar. 2000	GCP
-	11x11 km	08th Nov. 2000	GCP
Landsat7 ETM+	45x49 km	05th Mar. 2000	Ikonos / Map
-	45x49 km	18th Dec. 2000	Ikonos / Map

The images were acquired in pairs from two different satellites: Ikonos and Landsat7 ETM+, with different spatial but similar spectral resolution in the first four bands (table 2).

Table 2. Spectral and spatial specifications for the two satellite sensors used in this study

Satellite	Band	Spectral resolution	Spatial resolution
Ikonos	1	0.45 – 0.52 µm	4 m
-	2	0.52 – 0.60 µm	4 m
-	3	0.63 – 0.69 µm	4 m
-	4	0.76 – 0.90 µm	4 m
Landsat7 ETM+	1	0.45 – 0.52 µm	30 m
-	2	0.52 – 0.60 µm	30 m
-	3	0.63 – 0.69 µm	30 m
-	4	0.76 – 0.90 µm	30 m
-	5	1.55 – 1.75 µm	30 m
-	6	10.4 – 12.5 µm	60 m
-	7	2.08 – 2.35 µm	30 m

1.3 *First observations and hypotheses*

2-D scattergrams of pair-wise band-combinations were produced for the four satellite images in turn. Strong linear scattering was detected between the red and near infrared channels in all the images (figure 2). In fact all the band combinations involving the near infrared channel resulted in linearity in the 2-D scattergram for all images. On the other hand, all combinations of channels only from within the visible spectra showed lumpy scattering.

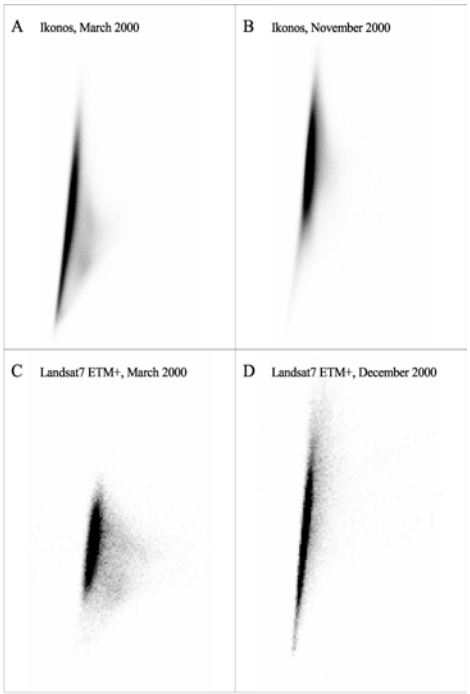

Figure 2. 2-D scattergrams of band-wise pixel-values from Ikonos and Landsat images. In each case the X-axis indicates the digital values in the red band, and the Y-axis indicates digital values in the near-infrared band. The scales of the axes in the two scattergrams are identical for each image pair from a sensor, but differ between the sensors.

It was hypothesised that the linear scattering of pixel-values primarily is due to structural features within the scenes. The causes of these structural features are terrain, canopy effects, and gaps in the canopy.

Shannon's Sampling Theorem states that a function band limited at a frequency of W can be reconstructed exactly from samples taken at a frequency of $W/2$ (Nyquist, 2002; Shannon, 1948). The implication for this study is that the sensor with a low spatial resolution will not detect the structural patterns occurring at the canopy level.

Figure 3. False colour (NIR, red, green) Ikonos image dated March 2000. The study area is the forest within this scene. At the centre of the image a military radar station can be seen together with a road coming up from the southern side.

Shadowing and shading in the image due to the terrain and its interaction with the sun's position will cause discrete spatial patterns in the scene. The size of the patterns is expected to be several tens of meters in this scene.

These terrain effects should be readily detected in the ETM image data, because the sizes of the topographic elements are several times larger than the image resolution. These elements should also be detected in the Ikonos data with its finer resolution. They are obvious in the Ikonos image (Figure 3).

For the Landsat ETM+ sensors with a spatial resolution of 30 meters, these terrain effects will be detected, due to their relatively large size comparative to image resolution, and hence separation of them theoretically possible. This will logically also apply to the Ikonos sensors because the spatial resolution is 4 meters on the ground and hence of a finer resolution than the ETM+ sensors.

The canopy effects on the other hand will most likely be of a size of 1 – 10 meters (Nagendra, 2001), and hence much smaller than the size of the Landsat pixels. This will consequently be averaged out in the Landsat ETM+ data, leaving the larger terrain effects to dominate the data scattergrams.

In contrast the Ikonos sensor is expected to be capable of detecting elements within the range of the presumed canopy effects additionally to the terrain effects. The data produced with the Ikonos sensor is hence expected to be the result of a complex spatial pattern of mostly shadowing and shading due to variation in both the terrain and the canopy. Given that the per-pixel effect of shadowing is the same independent of the source of the effect, the outcome on the scattergram will be the same for the canopy and terrain effects.

1.4 *Objectives for this study*

These observations led to the general hypothesis for this work: That the detected linear scattering almost similar in the two sensor-types, Ikonos and Landsat ETM+, is due to a combination of terrain and canopy effects.

The observations presented above supports the theory of local variance (Woodcock & Strahler, 1987). The model hypothesises that local variance, or to be more exact, *local standard deviation*, reaches a peak that is functionally related to the object/pixel ratio.

The present study aims at investigating whether the concept of local variance is capable of detecting variation present in a scene at more than one spatial scale and if so then to examine the object/pixel ratio for this compound situation.

Local variance values were calculated using a 3 by 3 pixel window moving over the entire image by steps of one pixel (Woodcock & Strahler, 1987). Accordingly, it is an objective of this study to consider what influence the size of the analytical window as well as the length of each step has on the capability to analyse the object/pixel interactions.

The results found from this study will then be applied to the Ikonos and Landsat ETM+ images acquired over the Doi Inthanon area. The purpose of this application will be to investigate whether the hypothesised causes of the linearity in the scattergrams are supported by the method of local varicance.

2 METHODS

2.1 *The Concept of Local Variance*

In 1987 the concept of local variance analysis was introduced (Woodcock & Strahler, 1987). Graphs of local variance in images as a function of spatial reso-

lution were used to measure spatial structure in images. Construction of these graphs was achieved by degrading the image under study to successively more coarse spatial resolutions, while measuring the local variance value at each of these resolutions.

Local variance was measured as the mean of all the values of *standard deviation* within a 3 pixels by 3 pixels wide window moving across the entire image with steps of size one pixel. For each step the standard deviation of the nine values located inside the window was computed. The mean of all local standard deviation values over the entire image were then used as an indication of the local variability contained by the image.

Degradation of the image to coarser resolutions was accomplished by averaging of a block of the low-resolution pixels to be combined into one single, larger resolution pixel. This entails a limitation to the number of times an image is able to be degraded and still contain enough pixels to determine local variance values. This restriction is obviously linked to the size of the analytical window as well.

The length of each step on the local variance curve, indicated as a multiple of fine-resolution pixels, were not specified. Thus, the curves were not produced by means of degrading the images into all possible levels but only into a quite arbitrary subset of these, approximating to some degree a doubling of pixel size at each aggregation level.

All results presented by Woodcock and Strahler showed either none or a single distinct peak on an otherwise smooth curve. For the no-peak situations the curves were then always sharply declining in the beginning. The peak, when present, was in all cases directly related to the size of the objects within the scene with pixel size being within the range of ½ - ¾ of the object size. Thus, the peaks did not occur exactly at the size of the objects in the scene. This observation was argued to be linked to the *Shannon Sampling Theorem* (Shannon, 1948), which states that in order for a band-limited signal to be reconstructed fully, it must be sampled at a rate at least twice as fast as the signal frequency, i.e. the *Nyquist Frequency* (Nyquist, 2002).

Cases of no peak were interpreted as an indication of the image resolution being too coarse to be capable of picking up the scene-variance. This statement was essentially supported by the decline at the beginning of these curves.

Woodcock and Strahler reported no curves with more than one peak; hence no mixed situations were detected (Woodcock & Strahler, 1987).

2.2 *Methods*

Four images of four different types of patterns were synthesized for the investigation presented here, cf. figure 3.

Image A consists of a systematic alternation of white and black square elements; 24 by 24 elements. The size of each element is 10x10 pixels, coded either 0 (=white) or 250 (=black). The image hence is composed of 240 by 240 pixels in total. This image then is representing a condition of only one object size.

Image B, on the other hand, is constructed to be indicating a situation with two object classes of different sizes. It consists of a regular arrangement of 24 by 24 square elements of 10 by 10 pixels in size plus 6 by 6 square elements of 40 by 40 pixels in size. The total size of image B is equal to image A. Both images have abrupt boundaries between the elements.

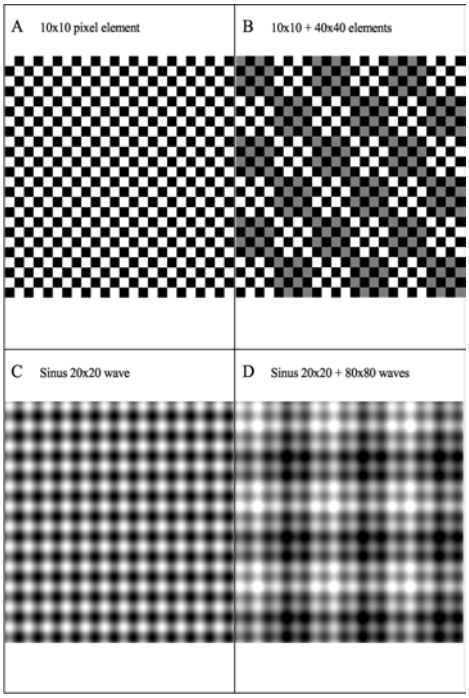

Figure 4. Synthetic images. A is having one element of size 10x10 pixels. B is having two elements of size 10x10 pixels and 40x40 pixels respectively. C is the result of a sine wave with a wavelength of 20 pixels in both directions. D is the result of two sine waves: one with a wavelength of 20 pixels in both directions and one with a wavelength of 80 in both directions.

Image C is constructed on the basis of a sine wave with a wavelength of 20 pixels in both directions. This image matches image A in the spatial repetition of elements but doesn't have sharp boundaries between the elements. In a way this pattern is the result of a number of local gradients which may be more realistic compared to the clear-

cut change within image A and B. Finally image D is constructed with two sine waves of 20 pixels and 40 pixels wavelengths respectively, hence matching image B.

The synthetic images were subsequently subjected to analysis by means of local variance. For this purpose a scheme was set up to investigate the four images per se but also to investigate into response to alteration of method-variables. Consequently the methodology of local variance (Woodcock & Strahler, 1987) was followed with some exceptions: three different window sizes were investigated; 2 by 2, 3 by 3, and 5 by 5. Additionally the window was either moving by the original steps spanning one pixel in size (termed *moving*) or by steps spanning one window in size (termed *jumping*). Table 3 presents an overview of the denomination of the different combinations of analysis variables.

Table 3. Denomination of the set-up of variables for analysis of local variance

Index	Description
Window size	
2	Window size is 2 by 2 pixels
3	Window size is 3 by 3 pixels
5	Window size is 5 by 5 pixels
Movement of the window	
M	Moving by size one pixel (*moving*)
J	Moving by size one window (*jumping*)
Images	
A	Image A, cf. figure 3A
B	Image B, cf. figure 3B
C	Image C, cf. figure 3C
D	Image D, cf. figure 3D

3 RESULTS

The results from the analysis are displayed in figure 5, 6, and 7. Of the 24 possible combinations of analysis variables (cf. table 3) only the 12 considered most important are presented here. The results presented are orientated similar to the images in figure 4.

Figure 5 displays the result for a 2 by 2 window moving by steps of one pixel in length. Curve 2MA displays a distinct and significant peak at pixel size 10. This peak is the largest on the curve. It reflects the situation at which the image-resolution corresponds exactly to the object size, and therefore all objects are contained in one pixel only. This gives the maximum variance possible in a two by two pixel window.

In contrast to any of the curves presented by Woodcock and Strahler (1987), the 2MA curve has several peaks. The second largest peak is at a pixel-size of 5. In this situation each object is comprised of four pixels. Consequently ¾ of the standard dev-

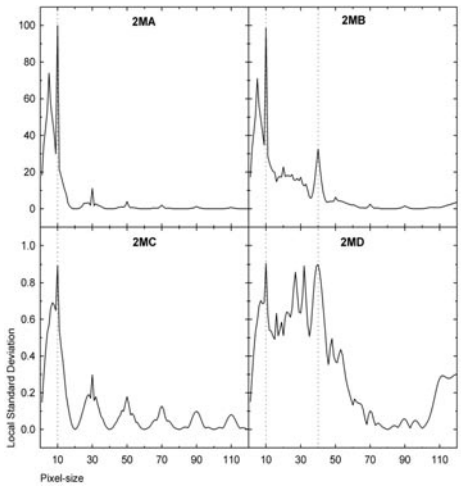

Figure 5. Local variance curves based on analyses of the synthetic images using a 2 by 2 -pixel moving window. Denomination of the curves follows the outline given by table 3. The vertical, dotted lines indicate the position of the object according to the pixel/object ratio. Pixel-size is indicated relative to the initial pixel-size.

-iation values detected by the moving window will be at maximum value and ¼ will result in a value of zero. This results in an overall value for the image, which is 75% of the value at the pixel-size of 10. This was in fact the result.

Additionally a number of auxiliary peaks are present at larger pixel-sizes. These sizes are 30, 50, 70 etc. For the situation with pixel-size being 30 each pixel will contain an average of pixel-values from exactly three by three objects. The result is therefore that every second pixel will have a value that is the average of five white objects and four black objects. The other pixels will contain the average of four white and five black. The remainder of small peaks can be explained by the same logic. This alteration in phase at certain intervals of larger scales is a result to be explained by the strict autocorrelation in the data.

In 2MB two significant and distinct peaks occur at pixel-size of 10 respectively of 40. It is worth noting, that the peak at 5 endures with the same magnitude as for 2MA and thus is larger than the peak at 40 in 2MB. At pixel-size of 20 a peak is present with a magnitude of ¾ of that at 40. The cause to this peak is the same as to the peak at 5; only the peak at 20 is related to the peak at 40.

The small peaks present in 2MA at positions 30, 50, 70 etc are still present in 2MB due to the same reason. For this reason a peak at size 120 was expected in 2MB since at this scale there will be a change in phase of the large objects (40 by 40 pixels) similar to that explained for the small objects.

Curve 2MC is similar to 2MA, but with some significant differences. The auxiliary peak at 5 pixels in 2MA has shifted to 7 pixels in 2MC, and decreased, since the strict autocorrelation in 2MA has been lost in 2MC. The auxiliary peaks at 30, 50, 70 etc. pixels continue to exist, but are much more profound, and with shoulders that are not so obvious in 2MA.

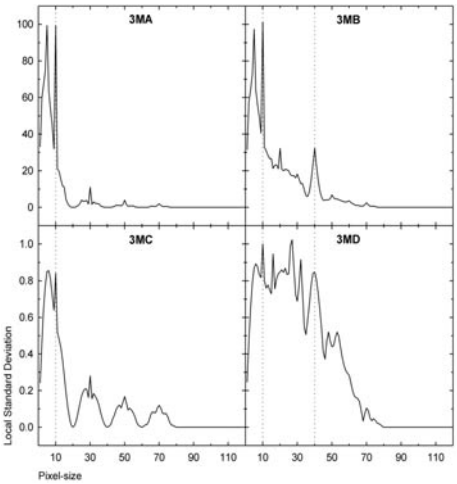

Figure 6. Local variance curves based on analyses of the synthetic images using a 3 by 3 pixel, moving window. Denomination of the curves follows the outline given by table 3. The vertical and dotted lines within the graphs indicate the position of the expected peak(s) according to the pixel/object ratio. ratio. Pixel-size is indicated relative to the initial pixel-size.

Curve 2MD contains almost the same peaks as 2MB, but with some differences. The auxiliary peak at 5 has shifted to 7 pixels in 2MD like it happened in 2MC. A trough in 2MD replaces the minor but distinct peak at 20 pixels in 2MB. This may be due to a right-shift of the peak, like the described right-shift for the peak at 5 pixels. Another obvious difference to 2MB is the number of auxiliary peaks between the 10 and 40 pixel peaks. The two most significant of these peaks are occurring at 27 and 32 pixels respectively. It seems like these double peaks occurs at intervals of 20 starting from 30, as for the auxiliary peaks in 2MA and 2MC. Thus the explanation may be that these are a resonance effect due to the autocorrelation, like it was explained in 2MC, but somehow enhanced by effects from the large objects. The build-up of the curve by the end can be explained as due to a shift in phase of the large object pattern similar to that explained for 2MB.

Curves 3MA and 3MB (figure 6) are very similar to 2MA and 2MB respectively, except that the auxiliary 5-pixel peak is much higher. At this level each object in the image is composed of 2 by 2 pixels. With a window size of 3 by 3 pixels, the local variance value will in all positions be a maximum value equal to that of one pixel per object.

Curves 3MC and 3MD are also very similar to 2MC and 2MD respectively, again with some interesting differences. The auxiliary peak has gone from 7 pixels in 2MC to 5 pixels in 3MC, and the auxiliary peaks between 10 and 40 pixels in 3MD have increased in value relative to these two peaks.

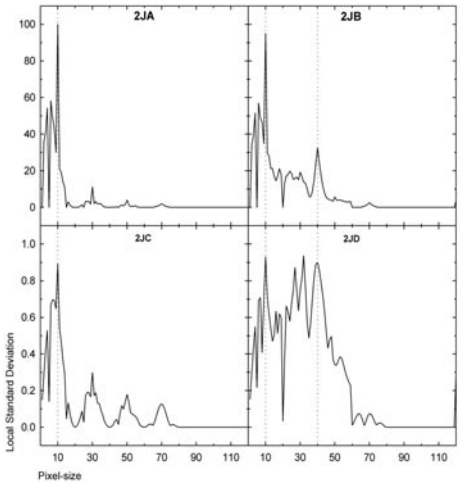

Figure 7. Local variance curves based on analyses of the synthetic images using a 2 by 2 pixel, jumping window. Denomination of the curves follows the outline given by table 3. The vertical and dotted lines within the graphs indicate the position of the expected peak(s) according to the pixel/object ratio. ratio. Pixel-size is indicated relative to the initial pixel-size.

The curves that result from a jumping window (figure 7) are similar to those of a moving window (figure 5), except for two interesting phenomena. The curves have a distinct trough at 5 pixels, actually with an average standard deviation value of zero in 2JA and 2JB. Again this can be explained with the fact, that each object is composed of 2 by 2 pixels, which exactly match the window size. Consequently, the window will contain only pixels of equal values in each step resulting in standard deviation values of zero. This is therefore an analytical artefact influenced by the strict autocorrelation. If the jumping window started by skipping just one pixel, then the result would have been a peak as significant as the peak at 10, because all standard deviation values would then have been at maximum.

The jumping curves moreover tend to be noisier than the moving curves, possibly because they do not overlap the same pixels more than once when computing local variance values to be averaged. The analogy between the moving window and a moving average is obvious, and the smoothing effect on a moving average curve is well known.

4 DISCUSSION

For the simplest situation with objects of only one size (image A & B in figure 4) a significant, distinct but very narrow peak occurs in all cases at pixel size exactly matching object size in concert with several auxiliary peaks, which often are much broader than the key peak (figure 5 and 6). This is in contrast to the results reported by Woodcock and Strahler (1987), which in all cases found only one broad peak at pixel size ½ - ¾ the size of the objects. We believe this difference is partly due to the close intervals on our curves as compared to the large intervals on the curves presented by Woodcock and Strahler (1987). Making large intervals on the curve will filter out many details, leaving only the broader patterns.

Further, we have found algebraic support for our observations that average local variance curves have a peak at a pixel size equal to object size for the method using a 2 by 2 moving window (McCloy & Bøcher, 2002). Comparing the different methods for image A and B (figure 5, 6 and 7) we conclude that the 2 by 2 moving window is the most informative method.

The average local variance curves are affected by autocorrelation in the scene. This effect is manifested as auxiliary peaks on the curves. The effect is smaller for the sine wave than for the square wave. For the simple situations with only one-sized objects, all the auxiliary peaks have been explained as a kind of resonance between the objects, and the pixels at the different resolutions, and the size of the analytical window (cf. the results). The auxiliary peak occurring left of the key peak is positioned at a pixel size ½ (2MA) or ¾ (2MC) of the position for the key peak. Since this peak is much broader than the key peak, one could ask the question whether this is the peak Woodcock and Strahler (1987) have reported on. This is of course not the entire explanation, since we have been working on synthetic data with all objects having the same size, and they reported on exposures of real scenes where objects will vary in sizes.

The auxiliary peaks occurring at multiples of a key peak may help to confirm the importance of the key peak [if having got a peak, check for autocorrelation]. Auxiliary peaks will occur if it is a key peak and may not if it is a combination of auxiliary peaks.

It seems to us that all details in the curve are important and therefore required for interpretation. Hence we conclude that all aggregation-levels are needed to construct the curve, not a subset of them.

For the combined scenes, where objects of two distinct sizes occur (image C and D in figure 4) we found a key peak for each object on the local variance curve together with several auxiliary peaks. In 2MB (figure 5) the peaks at 10 was as significant as it was in 2MA with the same auxiliary peak to the left. We also detected auxiliary peaks at the same multiples as in 2MA. The key peak at 40 in 2MB was much less significant, being only 1/3 as high as the peak at 10. This probably has to do with the intensity of these objects, having less sharp contrast in the image compared to the small objects. In 2MB we found an auxiliary peak at the same relative position to the left of the key peak as for the peak at 10. This auxiliary peak seemed to interfere with the signal from the auxiliary peak at 30 related to the peak at 10, making this signal complex for interpretation. Else no other peaks were found in 2MB as compared to 2MA.

For the sine wave (2MD in figure 5) the resulting curve is very complex. Two key peaks at position 10 and 40 were displayed. Auxiliary peaks were found to the left of both key peaks at position 7 and 22 respectively. We believe these peaks are analogous to the auxiliary peaks found left of the key peak in the simple images (image A and B in figure 4). In 2MD auxiliary twin-peaks were found at position 27 32, 48 53, etc. We believe these are analogous to the auxiliary peaks found to the right of the key peak in the simple images. Finally in 2MD a build-up was observed at the end of the curve towards 120, which is the triple distance to the key peak at 40. This is analogous to the first auxiliary peak to the right of the key peak at 10.

Even though the curves are much more complex for the combined scenes, it look as if the method of 2 by 2 moving window has the capacity of revealing the key peaks. A good rule of thumb appears to be to look for peaks having auxiliary peaks at positions multiples of the peaks. If it is a key peak, then it will have these auxiliary peaks, else it may be a combination of auxiliary peaks.

REFERENCES

Hamilton, L. S., Juvik, J. O., and Scatena, F. N. 1995. Tropical montane cloud forests. 407.

McCloy, K. R. and Bøcher, P. K. 2002. The fundamentals of local variance (In prep.).

Nagendra, H. 2001. Using remote sensing to assess biodiversity. Inter.

Nyquist, H. 2002. Certain topics in telegraph transmission theory (Reprinted national Journal of Remote Sensing. 22 (No. 12): 2377-2400 from Transactions of the A. I. E. E., February, pg 617-644, 1928). Proceedings of the IEEE. 90 (No. 2): 280-305.

Santisuk, T. 1988. An account of the vegetation of northern Thailand. (No. 5):

Shannon, C. E. 1948. A mathematical theory of communication (Reprinted with corrections). The Bell System Technical Journal. 27 379-423-623-656.

Woodcock, C. E. and Strahler, A. H. 1987. The factor of scale in remote sensing. Remote Sensing of Environment. 21 311-332.