IMAGE THRESHOLDING BASED ON SPATIAL VARIATION ATTRIBUTE SIMILARITY

M. Beauchemin* and Ko B. Fung
Natural Resources Canada
Canada Centre for Remote Sensing
588 Booth Street, Ottawa
Canada K1A 0Y7
*tel: (613) 943-8839, fax: (613) 947-1383,
e-mail: mario.beauchemin@ccrs.nrcan.gc.ca

Abstract

According to a recent study, image thresholding can be categorized into six groups of methods that are based on histogram shape, clustering, entropy, attribute, spatial, and local information. In this paper, we describe two algorithms for image binarization that are based on attribute similarity relying on spatial measures. The rationale of the method is to binarize an image in such a way that it best reproduces the spatial variation of the original image across several scales. Two different measures that characterize image spatial variation have been selected to pursue that objective: semivariance and lacunarity. Semivariance measures the spatial variation of a variable at a given scale. Lacunarity is a measure of translational invariance, at a given scale, and is often refer to as a measure of 'gappiness'. In both approaches, the threshold is selected so that the scale-dependant measure in the bi-level image best approximate, in the least square sense, the ones of the original image. Both methods are illustrated with remote sensing images of high spatial resolution. The results are compared with some other popular thresholding techniques.

Introduction

According to a recent survey (Sankur and Sezgin 2001), image thresholding can be categorized into six groups of methods that are based on histogram shape, clustering, entropy, attribute, spatial, and local information. In this paper, we describe two algorithms for image thresholding that are based on attribute similarity relying on spatial measures. The multiscale approach has become pervasive over a wide range of image processing techniques (e. g. image compression, image registration, image segmentation, noise filtering, texture analysis, edge detection). In remote sensing, it has been used, among others, to examine the spatial structure of images (see Marceau and Hay 1999 for a review on scale issues in remote sensing). In particular, semivariance and lacunarity analysis has been utilized to study the spatial characteristics of images (Cohen *et al*. 1990; Sun and Ranson 1998). The rationale of the method proposed in this paper is to binarize an image in such a way that it best reproduces the spatial variation of the original image across several (spatial) scales. For the evaluation of several popular thresholding techniques, see Sezgin and Sankur (2001, and references therein).

Method

THRESHOLDING BASED ON SEMIVARIANCE

Semivariance measures the spatial variation of a variable at a given scale (Milne and Cohen 1999). It is defined as:

$$\gamma(h) = \frac{1}{2\,N(h)} \sum_{i=1}^{N(h)} (DN_i - DN_{i+h})^2 \quad , \tag{1}$$

where DN_i is the digital number value at location i and $N(h)$ is the number of pairs of observations separated by the same distance lag, h. A variogram is obtained when $\gamma(h)$ is plotted against h.

Let us binarize an image by setting all pixel values below a given threshold t to z_0, and all pixels that are equal or above that threshold to z_1. The number of pixels associated with levels, z_0 and z_1, are respectively N_0 and N_1. The rationale of the proposed algorithm is to threshold the original image and then adjust the binarized values, z_0 and z_1, such that the semivariance of the bilevel image best approximates the ones of the original image over a range of values of h. Let $\gamma^o(h)$ denotes the semivariance of the original image at lag distance h and $\gamma^b_{\mathrm{zo,z1}}(h, t)$ the semivariance of its binarized version. The sum of the square of the difference between the semivariances of both images, Δ^2, can be used as a measure of 'goodness-of-fit' to quantify how well the binarized image reproduces the spatial characteristics of the original image:

$$\Delta^2 = \sum_h [\gamma^{o}(h) - \gamma^b_{\mathrm{zo,z1}}(h,t)]^2 \, . \tag{2}$$

It is thus the variogram of both images that are actually compared. It can be shown that for fixed values of h and t:

$$\gamma^b_{\mathrm{zo,z1}}(h, t) = (z_1 - z_0)^2 \gamma^I(h, t), \tag{3}$$

the semivariance of an image binarized with values z_0 , z_1 ($z_1 > z_0$), is equal to $(z_1 - z_0)^2$ times the indicator semivariance, γ^I. The latter is obtained by assigning values of 0 and 1 to the bilevel image. After substitution of (3), (2) can be rewritten as:

$$\Delta^2 = \sum_h [\gamma^{o}(h) - k(t)^2 \gamma^I(h,t)]^2 \, , \tag{4}$$

where $k = (z_1 - z_0)$. For a given value of t, z_0 and z_1 must be selected to minimize Δ^2. By setting $\partial \Delta^2 \big/ \partial(k^2) = 0$, the value of k^2 that minimizes Δ^2 is found:

$$k^2(t) = \sum_h \gamma^{o}(h) \gamma^I(h,t) \Big/ \sum_h [\gamma^I(h,t)]^2 \, . \tag{5}$$

Here, k represents the difference in grey levels that best reproduce, in the least-square sense, the semivariance of the original image at the given threshold value t. The optimal threshold, t_s*, is the threshold t_s that minimize (4) with k given by (5):

$$t_s^* = \underset{t \in DN}{\operatorname{Arg\,Min}} \left\{ \sum_h [\gamma^{o}(h) - \gamma^{I}(h,t) \sum_h \gamma^{o}(h)\, \gamma^{I}(h,t) \Big/ \sum_h [\gamma^{I}(h,t)]^2\,]^2 \right\} , \qquad (6)$$

where DN is the set of integers that represent the digital number values in the image.

THRESHOLDING BASED ON LACUNARITY

Lacunarity emerged from the study of fractals (Mandelbrot 1983). It was introduced to describe the distribution of gap sizes of a geometric object. It is a measure of translational invariance. Plotnick *et al*. (1996) have presented lacunarity analysis as a general technique for the analysis of spatial patterns. Many measures have been suggested to quantify lacunarity. Here, we used the gliding box algorithm introduced by Allain and Cloitre (1991). Consider a binary image where object pixels are attributed a value of one and zero otherwise. A square window or box of size r (pixels) is first placed in one corner of the image (say upper left) and the number of pixels of value equal to one within that window is counted. This number represents the box mass. The window is then moved one pixel in the horizontal direction, and the box mass is determined again. This process is repeated for each pixel line until the whole image is scanned. This algorithm is usually applied for values of r ranging from r=1 to one half the dimension of the image. For every value of r, the mean μ and variance s of all box masses are computed and the lacunarity $\Lambda(r)$ is obtained from:

$$\Lambda(r) = s(r) / \mu(r)^2 + 1. \qquad (7)$$

Lacunarity is a dimensionless number. Plotnick *et al*. (1996) have indicated that lacunarity can also be computed on quantitative data. For raster images, the box mass is equal to the sum of the digital number values within the gliding box (Sun and Ranson 1998). In this paper, we will consider the first term in (7) involving the ratio of the statistical measures, that we define as $\lambda(r) = \Lambda(r) - 1$ or,

$$\lambda(r) = s(r) / \mu(r)^2 \ . \qquad (8)$$

Similarly to the treatment applied to semivariance, a set of equations equivalent to equations (2) to (6) can be derived for the lacunarity-related measure $\lambda(r)$:

$$\delta^2 = \sum_r [\lambda^{o}(r) - \lambda^{b}_{zo,z1}(r,t)]^2 , \qquad (9)$$

$$\lambda^{b}_{z_o,z_1}(r,t) = \left[\left(\frac{z_1}{z_o} - 1\right) \Big/ \left(\frac{N_o}{N_1} + \frac{z_1}{z_o}\right) \right]^2 \lambda^{b}_{0,1}(r,t) \ , \qquad (10)$$

$$\delta^2 = \sum_r [\lambda^o(r) - g(t)^2 \lambda^b{}_{0,1}(r,t)]^2 , \tag{11a}$$

$$g = (\frac{z_1}{z_o} - 1)(\frac{N_o}{N_1} + \frac{z_1}{z_o})^{-1} , \tag{11b}$$

$$g^2(t) = \sum_r \lambda^o(r) \lambda^b{}_{0,1}(r,t) \Big/ \sum_r [\lambda^b{}_{0,1}(r,t)]^2 , \tag{12}$$

$$t_l^* = \underset{t \in DN}{\mathrm{Arg\,Min}} \{ \sum_r [\lambda^o(r) - \lambda^b{}_{0,1}(r,t) \sum_r \lambda^o(r) \lambda^b{}_{0,1}(r,t) \Big/ \sum_r [\lambda^b{}_{0,1}(r,t)]^2]^2 \} . \tag{13}$$

In (9) to (13), $\lambda^o(r)$ is the lacunarity-related measure of the original image, $\lambda^b{}_{zo,z1}(r,t)$ is the lacunarity-related measure of a binary image with values equal to z_o and z_1, $\lambda^b{}_{0,1}(r,t)$ is the lacunarity-related measure of the binary image with values equals to 0 and 1, and finally δ^2 is the sum of the square of the difference between the lacunarity-related measure of the original image and its binarized version. δ^2 is the goodness-of-fit criterion that quantifies how well the binarized image reproduces the spatial characteristics of the original image.

Experimental Results

SYNTHETIC IMAGE

The application of the two proposed algorithms to a synthetic image is shown in Figure 1, along with the results of the methods of Kittler and Illingworth (1986), Kapur *et al.* (1985) and Tsai (1985). These methods fall under the clustering, entropy and attribute similarity categories respectively. The synthetic image consists of disks that are spatially distributed at random (uniformly) over a constant background. Overlapping disks have the same grey level intensity as single disks. Uncorrelated Gaussian white noise has been added to each pixel after the requested number of disks was reached. The image was then coded into 8-bits. The image in Figure 1 is 128 x 128 pixels in size and the number of generated disks, all having a diameter of 9 pixels, is 180. The noise-free disks and background have respectively gray level values of 106 and 100. The standard deviation of the noise generator was 3. The frequency count histogram is shown in Figure 2. Clearly, a valley between two successive peaks cannot be distinguished. Threshold values (t^*) displayed in Figure 1 indicate that the values computed from the proposed algorithms are, overall, in agreement with the other methods.

REAL IMAGES

Figure 3 and 4 present two near-infrared images of forested areas. They have been extracted from an image acquired by a multispectral video camera. The images are 128 x 128 pixels in size (8-bits, 1 meter pixel size). Figure 3 indicates how the thresholds (t^*) vary from one method to another for the image in the upper left corner of Figure 3. The method based on semivariance has a better match with the Kittler-Illingworth method (59 vs. 58) while the lacunarity-based method has a value not far from the one of Tsai (66 vs. 63). Results in Figure 4 shows more variation than in Figure 3 among the different methods for the image in the upper left corner of Figure 4. Lacunarity- and semivariance-based methods give the highest threshold values (135 and 132) followed next by the Kittler-Illingworth method (123). The Tsai result has the lowest one (102). It seems that, to reproduce the spatial variation of the original image, more emphasis should be given to the darkest component using the proposed approach (when compared with the others).

Conclusion

Two algorithms for image thresholding have been described. The rationale is to binarize an image in such a way that it best reproduces the spatial variation of the original image across several (spatial) scales, as measured by semivariance and lacunarity. A comparison with some other popular thresholding methods has been presented for one synthetic image and two real images. For the limited image sample analyzed, the results indicate that the scale-based methods match better, generally, with the result obtained with the Kittler-Illingworth method. The proposed approach probes several spatial scales and is therefore sensitive to noise effects that will manifest at the pixel scale level. Hence, in a two-component model where an image is composed of objects distributed over a noisy background, scale-based measures will include the characteristics of the noise.

References

Allain, C., and Cloitre, M. 1991. Characterizing the lacunarity of random and deterministic fractal sets, *Physical Review A 44*, Vol. 6, pp. 3552-3558.

Cohen, W. B., Spies, T. A., and Bradshaw, G. A. 1990. Semivariograms of digital imagery for analysis of conifer canopy structure, *Remote Sensing of Environment*, Vol. 34, pp. 167-178.

Kapur, J. N., Sahoo, P. K., and Wong, A. K. C. 1985. A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram, G*raphical Models and Image Processing*, Vol. 29, pp. 273-285.

Kittler, J., and Illingworth, J. 1986. Minimum Error Thresholding, *Pattern Recognition*, Vol. 19, pp. 41-47.

Marceau, D.J., and Hay, G.J. 1999. Remote sensing contributions to the scale issue. *Canadian Journal of Remote Sensing*, vol. 25, pp. 357-366.

Milne, B.T., and Cohen, W.B. 1999. Multiscale assessment of binary and continuous landcover variables for MODIS validation, mapping and modeling applications, *Remote Sensing of Environment*, Vol. 70, pp. 82-98.

Plotnick, R. E., Gardner, R. H., Hargrove, W. W., Prestegaard, K., and Perlmutter, M. 1996. Lacunarity analysis: a general technique for the analysis of spatial patterns, *Physical Review E. Statistical Physics, Plasmas, Fluids, & Related Interdisciplinary Topics*, vol. 53, pp. 5461-5468.
Sankur, B., and Sezgin, M., Image Thresholding Techniques: A Survey over Categories, submitted to *Pattern Recognition*, 2001.

Sezgin, M., and Sankur, B., Image Thresholding Techniques: Quantitative Performance Evaluation, (under review*) Pattern Recognition*, 2001.

Sun, G., and Ranson, K. J. 1998. Radar modeling of forest spatial structure, *International Journal of Remote Sensing*, Vol. 19, pp. 1769-1791.

Tsai, T. 1985. Moment-preserving thresholding: A new approach. *Computer vision graphics, Image Processing*, Vol. 29, pp. 377-393.

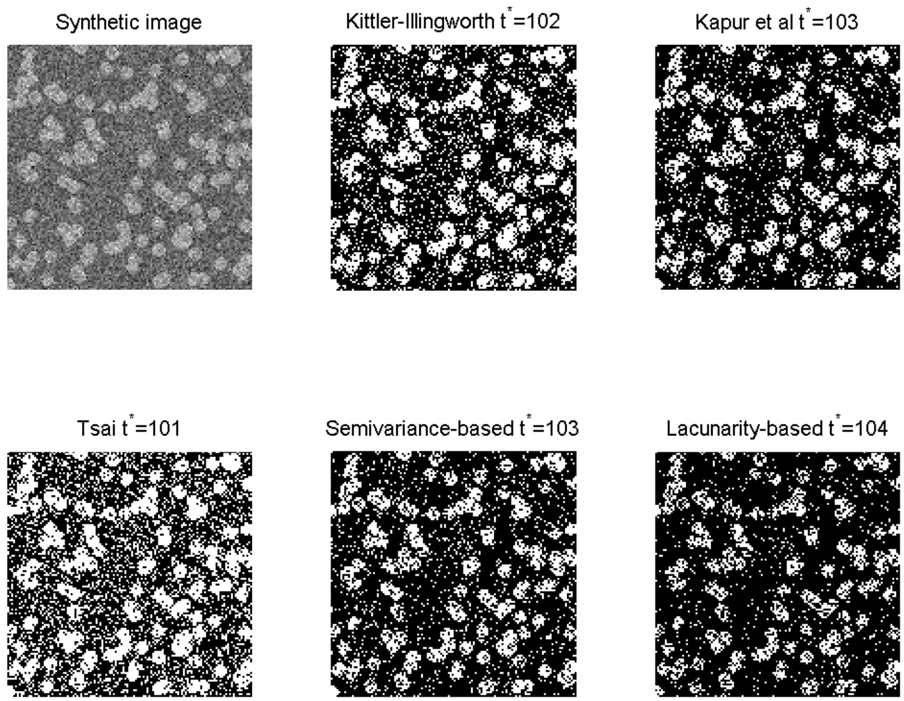

Figure 1. Thresholding results – Synthetic image.

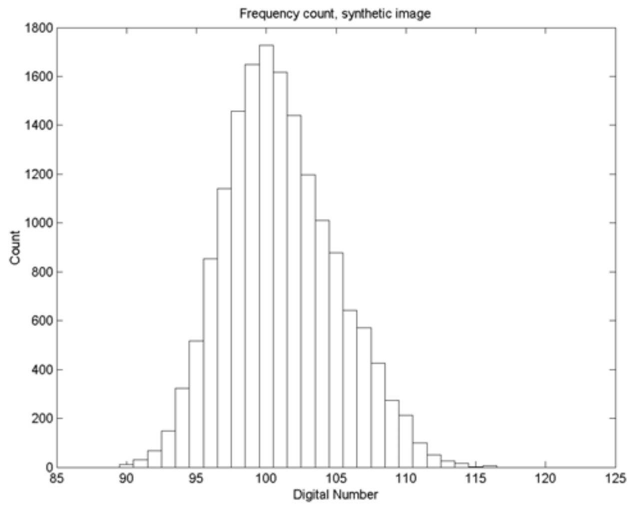

Figure 2. Frequency count – synthetic image.

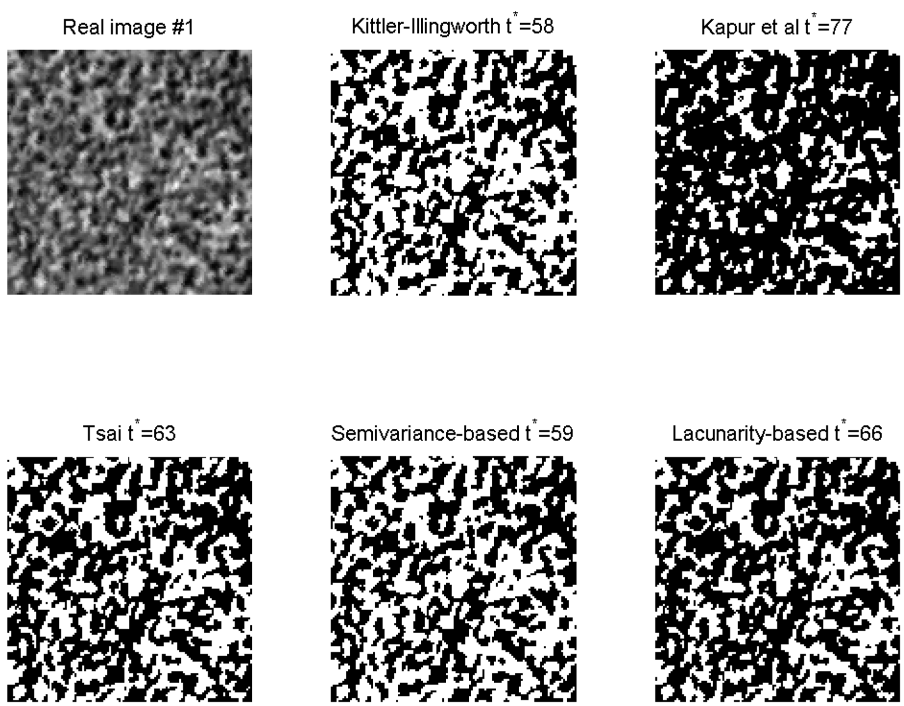

Figure 3. Thresholding results – Real image #1.

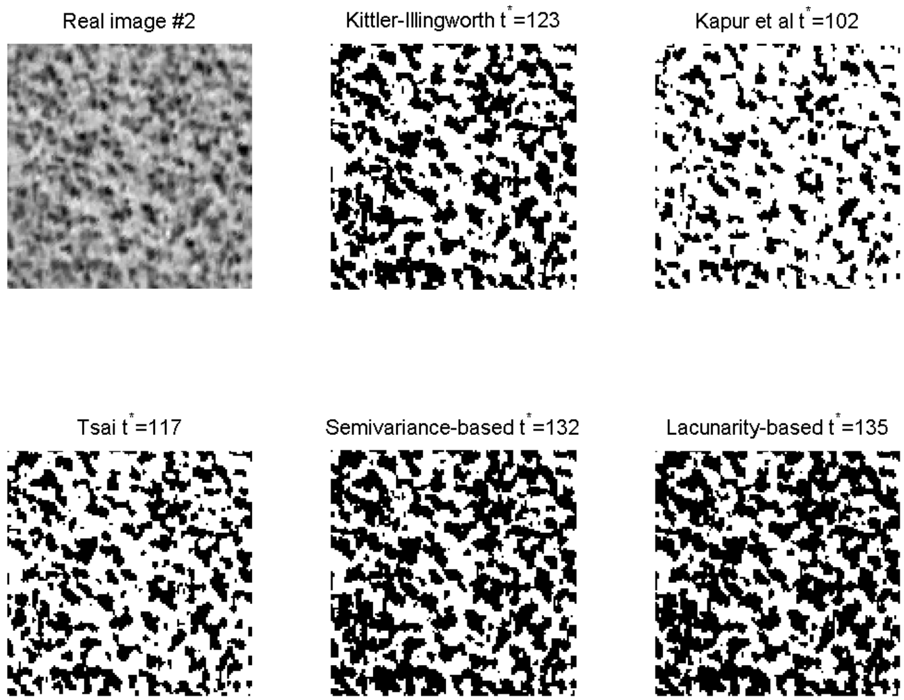

Figure 4. Thresholding results – Real image #2.