Automatic Variogram Parameter Extraction for Textural Classification of the Panchromatic IKONOS Imagery

Qi Chen and Peng Gong

Abstract—**Range and sill are two important parameters of a variogram. Their extraction usually involves experimental fitting of variograms using models specified by the analyst and requires much use of trial and error. The objective of this paper is to design an algorithm for extracting the range and sill of a variogram automatically without fitting a model. Combined with the semivariance at the lag of one pixel (γ_1), the extracted range and sill were applied to the textural classification of a panchromatic IKONOS image over Xichang, Sichuan Province, China. Results show that any of these three parameters can lead to the increase of the classification accuracy. When all three parameters were used with the raw image data, the average kappa statistic for five window sizes increased from 0.24 to 0.76, indicating promise of the range and sill in texture classification.**

Index Terms—**IKONOS, range, sill, texture, variogram.**

I. INTRODUCTION

SINCE THEIR introduction in remote sensing [1], [2], geostatistical techniques have been widely applied in mineral mapping with hyperspectral Airborne Visible/Infrared Spectrometer (e.g., [3]), replacement of cloud pixels in Advanced Very High Resolution Radiometer (AVHRR) imagery using kriging method (e.g., [4]), noise estimation (e.g., [5]), design of optimal sampling strategies for remote sensing investigation (e.g., [6] and [7]), and evaluation of the effects of spatial resolution in certain applications (e.g., [8] and [9]). A general discussion of such applications can be found in [10]. Recently, much research has been undertaken to apply variogram in texture classification (e.g., [11]–[20]).

Texture, which is an important property existing in remote sensing imagery, refers to the variation of gray tones in a local neighborhood [21]. There are three categories of methods for textural classification: structural, spectral, and statistical methods [22]–[24], among which statistical methods are widely used. As a kind of second-order statistics, a variogram

captures the covariance structure of a certain spatial process [25]. Treating each pixel's digital number (DN) as a realization of the spatial process, an experimental variogram can be calculated in each pixel's neighborhood window. Either the semivariance at each lag or variogram model parameters can be used for textural classification. Correspondingly, algorithms that use variograms for textural classification can be divided into two categories [26], [27].

A. Using Semivariances Directly

A semivariogram textural classifier (STC) was proposed to perform image classification based on the semivariogram signature of remotely sensed data [12], [16], [17]. Carr and Miranda [12] compared a classical gray-level coccurrence method with a semivariogram using three kinds of optical images—Systeme Pour l'Observation de la Terre (SPOT) High Resolution Visible (HRV) near infrared; Landsat Thematic Mapper (TM), red; and the Indian Remote Sensing Linear Self-Scanning Sensor (LISS-II), green—and three kinds of microwave images—Magellan, Venus, S-band; SIR-C, X-band; and SIR-C, L-band—taken from national park, rainforest, and urban surface cover types. They found that the semivariogram texture measure provided higher classification accuracy for the microwave images but lower accuracy for the optical images than the cooccurrence method. The idea of using semivariance as texture measure is intuitive and can be easily realized. However, variogram parameters such as range, sill, and nugget, could not be obtained.

B. Using Derived Parameters From Experimental Variogram

To automate the extraction of variogram parameters for remote sensing image classification, Ramstein and Raffy [28] proposed an approximate algorithm to calculate sill and range (denoted as C and a, respectively, in their paper)

$$C = \lim_{h \to \infty} \gamma(h) = \lim_{h \to \infty} \frac{1}{2} \left[\text{mean}(L(x+h) - L(x))^2 \right]$$
$$= \underset{x \in v}{\text{mean}} \, L^2(x) \tag{1}$$

$$a = \frac{-1}{\log \left[1 - \gamma(1) / \underset{x \in v}{\text{mean}} \, L^2(x) \right]} \tag{2}$$

where $\gamma(h)$ is the semivariance at lag h, $L(x)$ is the gray-level value at pixel location x, and x varies in the pixel neighborhood window. Parameter a in (2) is derived based on the assumption

that the underlying variogram can be modeled using an exponential variogram

$$\gamma(h) = C\left[1 - \exp\left(\frac{-h}{a}\right)\right]. \tag{3}$$

There are several issues associated with this method: first, if the underlying variogram models are different (e.g., power, Gaussian models), the coefficients of the model fitted to the local variogram may be misleading [29]. Second, even if this assumption is relaxed, a real number a may not be obtainable in practice because there are many cases where $\gamma(1)$ is greater than $\text{mean}_{x \in v} L^2(x)$, leading to a meaningless value. This may happen under a variety of situations. For example, if the underlying variogram model is a pure nugget effect model, while there exist outliers, then the semivariance at lag 1 can be greater than the sill [e.g., Fig. 4(i)]. Moreover, $\text{mean}_{x \in v} L^2(x)$ can be used as the estimate of sill only under very ideal conditions; the window size must be large enough, while texture primitives in the neighborhood window are relatively homogenous. Unfortunately, as mentioned later in this analysis (Section III-D), this requirement is usually contradictory and cannot always be satisfied, possibly leading to C greater than $\gamma(1)$. Fig. 1 shows that in a paddy field there are many pixel windows from which a meaningful a could not be obtained.

Herzfeld [13] and Herzfeld and Higginson [14] developed another set of texture parameters from variograms for seafloor classification with remote sensing imagery. Their idea was to first smooth the experimental variogram using a Butterworth filter and then obtain a set of parameters from it, including *maximum variogram value*, the lag to the first minimum after the first maximum (*mindist*), significance of abyssal hills, which represent slope $p1$ and relative size $p2$, respectively. However, their parameter vector was designed specifically for seafloor classification and has not been tested over land. This method does not obtain the range, nugget, and sill of a variogram model either.

Common practice for extracting range, nugget, and sill is to fit an experimental variogram using a conditionally negative semidefinite (CNSD) model based on the weighted least square approximation method [29]. The model must be specified by the analyst, restricting the automation of parameter extraction.

Based on the above analysis, it is desirable to develop some algorithms that could: 1) automatically extract the important variogram parameters, i.e., range, sill, and nugget, and 2) be independent of the underlying variogram models. In this paper, we propose a new algorithm, dubbed SEVDV (i.e., smoothing the experimental variogram using difference of variance), to extract range and sill of a variogram automatically, and evaluate the utility of range and sill in the textural classification of panchromatic IKONOS imagery.

II. Methods

A. Variogram

The gray-level value in an image can be treated as a *regionalized variable*, which is denoted as $\{L(x); x \in D\}$, a real-valued
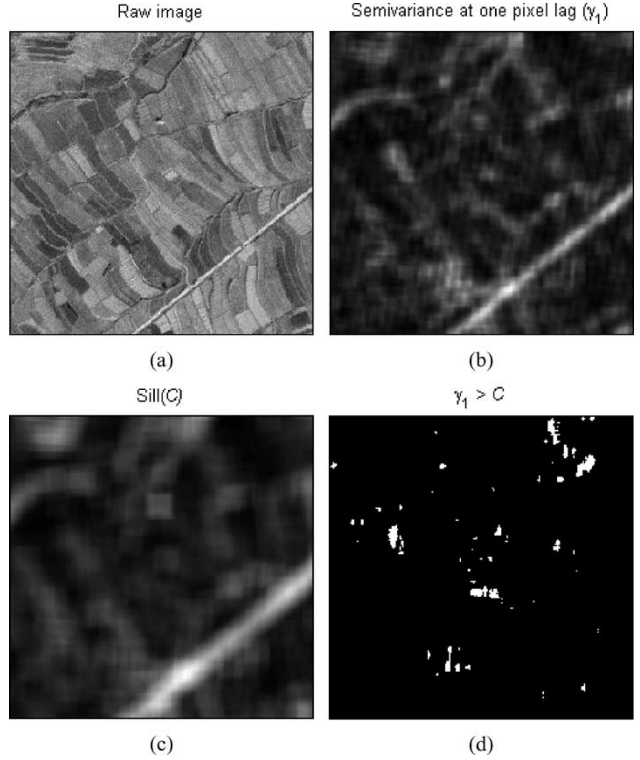

Fig. 1. Raw and derived images of a paddy field. (a) Raw image. (b) γ_1. (c) Estimate of sill C. (d) Binary image with white area representing pixels having γ_1 greater than C.

stochastic process defined over a domain D of R^2. The *intrinsic hypothesis* of the regionalized variables is

$$E(L(x+h) - L(x)) = 0 \tag{4}$$
$$\text{Var}(L(x+h) - L(x)) = 2\gamma(h), \qquad \text{where } x, x+h \in D. \tag{5}$$

$\gamma(h)$ is called the semivariogram or variogram. It can be estimated by [30]

$$\gamma^*(h) = \frac{1}{2N(h)} \sum [L(x) - L(x+h)]^2 \tag{6}$$

where $N(h)$ is the number of pairs of observations made at locations x and $(x+h)$, separated by a vector h.

The implication of (4) is that $L(x)$ is spatially independent, and its expectation does not vary in space. Equation (5) means that the difference of two regionalized variables is independent of specific spatial location and only related to the distance between them, which is a weak replacement of the *second-order stationarity*

$$C(h) = \text{cov}(L(x+h), L(x)), \qquad \text{where } x, x+h \in D. \tag{7}$$

When (7) holds, the intrinsic hypothesis and second-order stationarity can be related using

$$2\gamma(h) = 2(C(0) - C(h)). \tag{8}$$

Remotely sensed data over large areas often incorporates a nonstationary trend, resulting in violation of the assumptions. The trend should be removed before calculating the variogram.
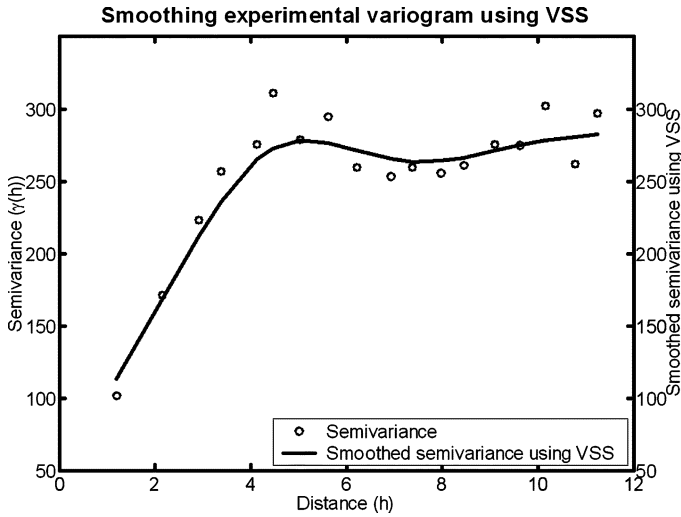
Fig. 2. Smoothing experimental variogram using VSS.

B. SEVDV Algorithm

The SEVDV algorithm for extracting variogram model follows: 1) calculating the experimental variogram of a pixel neighborhood; 2) smoothing the experimental variogram using a variable span smoother (SEV); and 3) calculating the lagwise difference of variance-to-mean ratio (DVmr) between the two sides of each lag, based upon which the range and sill are obtained. The last two steps will be explained in detail.

1) Smoothing the Experimental Variogram (SEV): To reduce the local variation the experimental variogram is smoothed first. Smoothing is a kind of nonparametric curve fitting technique. It is attractive because it is stable and able to fit nonlinear data locally. One possible way to do this is to fit the experimental variogram using a polynomial. However, the choice of terms in a polynomial is not always obvious, since the underlying variogram models could be quite different [31]. Other widely used smoothing techniques include local linear regression, cubic smoothing, kernel-type smoothing, and variable span smoother [32]. After many tests, the variable span smoother (VSS) proposed by Friedman [33] was chosen. In an ideal variogram, the curvature of semivariance near the range reaches maximum. A fixed span smoother will flatten the curve at the range and make it difficult to obtain the range. However, VSS can be optimal, since it shrinks its span near the range using the leave-one-out cross validation technique. Also, VSS is much faster than local weight linear regression, since it does not involve matrix inverse operation [31]. Fig. 2 shows an example of fitting an experimental variogram using VSS.

2) Rationale on Detecting Range and Sill Using DVmr: Denote the lag of a variogram as i, where, $i \in [1, n]$ with n being the maximum lag distance of the variogram. Also, denote the subset of semivariances from the smoothed experimental variogram with lags less than and equal to i as SEV_{Li}, and the subset with lags greater than i as SEV_{Ri}, i.e.,

$$\mathrm{SEV}_{Li} \equiv \{\gamma(h),\ h \le i\}$$
$$\mathrm{SEV}_{Ri} \equiv \{\gamma(h),\ h > i\}, \qquad \text{where } i \in [2,\ n-2].$$

(a)

(b)
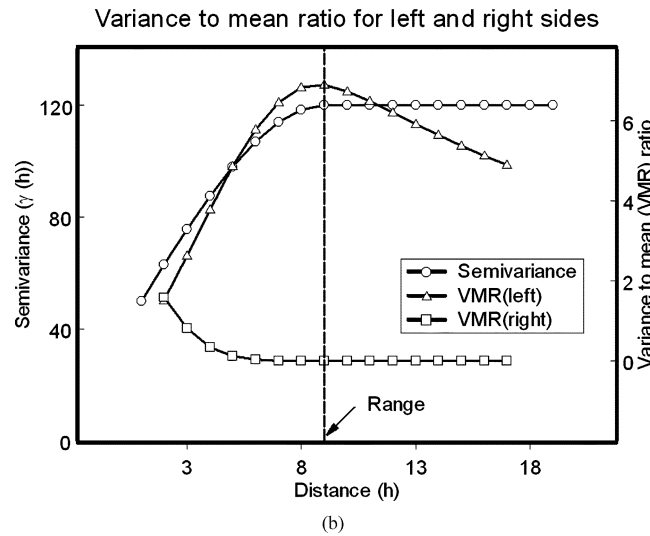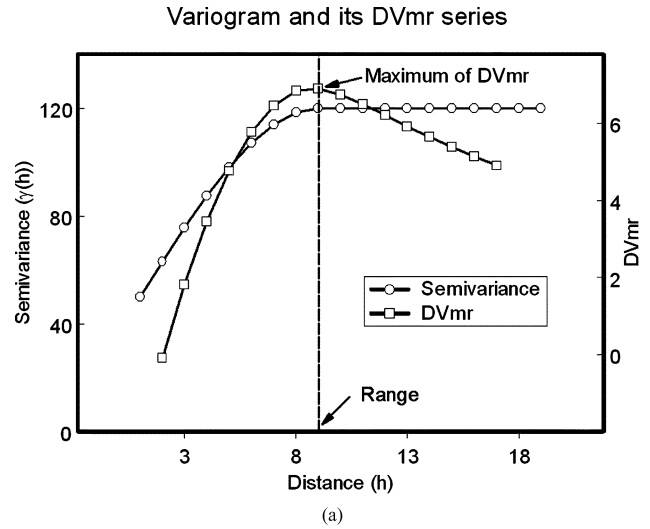
Fig. 3. (a) Variogram and its series of DVMr, and (b) the corresponding $\mathrm{VMR}_{\mathrm{left}}$ and $\mathrm{VMR}_{\mathrm{right}}$.

Clearly, SEV_{Li} includes i points, while SEV_{Ri} has $n - i$ points. The upper and lower limits of i ensure that there are at least two points in each subset. Varying i from 2 to $n - 2$, a total of $n - 3$ pairs of SEV_{Li} and SEV_{Ri} can be obtained.

Calculate the variance-to-mean ratio (VMR) for points in SEV_{Li} and SEV_{Ri}, respectively, and denote them as VMR_{Li} and VMR_{Ri}

$$\mathrm{VMR}_{Li} = \frac{\mathrm{VAR}(\mathrm{SEV}_{Li})}{\mathrm{mean}(\mathrm{SEV}_{Li})} \qquad (9)$$

$$\mathrm{VMR}_{Ri} = \frac{\mathrm{VAR}(\mathrm{SEV}_{Ri})}{\mathrm{mean}(\mathrm{SEV}_{Ri})}. \qquad (10)$$

Denote DVmr as the difference between SEV_L and SEV_R

$$\mathrm{DVmr}_i = \mathrm{VMR}_{Li} - \mathrm{VMR}_{Ri}, \qquad \text{where } i \in [2, n-2]. \quad (11)$$

Fig. 3(a) is an example of a DVmr data series for a spherical variogram. In this case, the lag distance where DVmr_j reaches maximum is exactly the range of the variogram.

To explain this, we need to recall what variance-to-mean attempts to quantify. VMR is an index widely used to quantify the
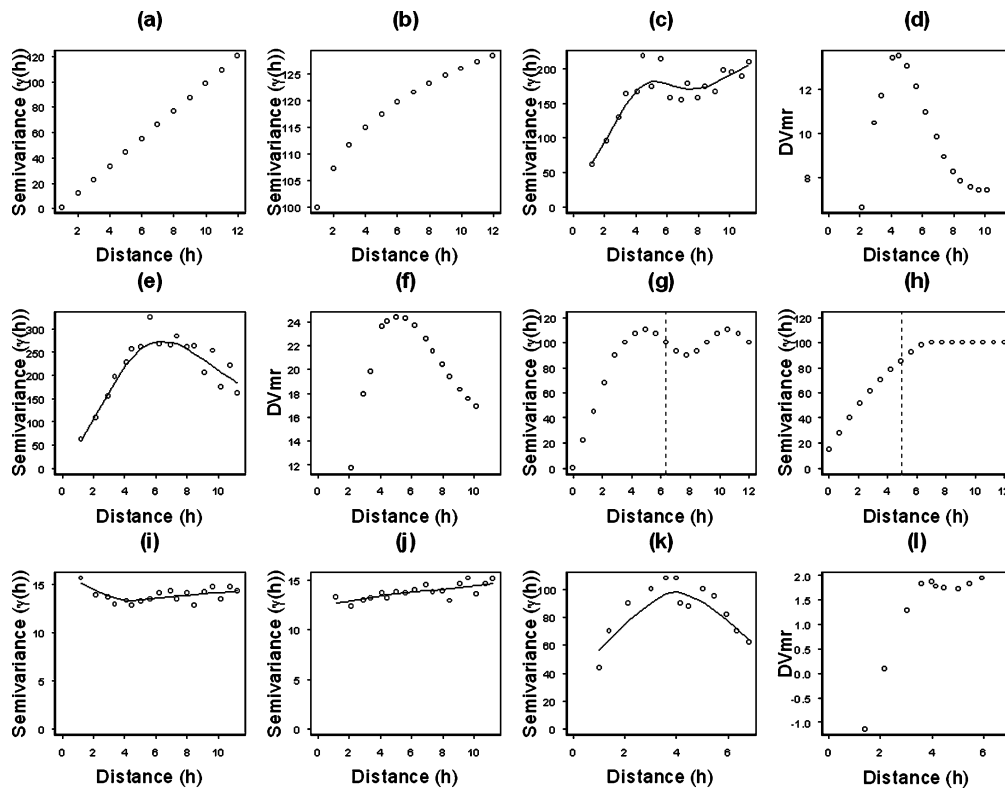
Fig. 4. Various variogram model cases. (a) Linear model. (b) Power model. (c) Experimental variogram with upward semivariances after range and its SEV (solid line). (d) Series of DVmr of (c). (e) Experimental variogram with downward semivariances after range and (solid line) its SEV. (f) Series of DVmr of (e). (g) Period model where the dashed line indicates the maximum lag distance. (h) Bounded variogram model where the dashed line indicates the maximum lag distance. (i) Experimental pure nugget effect model with SEV having a maximum value at the first lag. (j) Experimental pure nugget effect model with the variance-to-mean ratio of SEV being 0.023. (k) Experimental variogram with downward semivariances after range and (solid line) its SEV. (l) Series of DVmr of (k).

dispersion of point patterns in spatial statistics. Generally, VMR tends to be greater than one for clumped point pattern, equal to one for random pattern, and less than one for uniform pattern. Fig. 3(b) shows that when i increases from 2 to $n-2$, VMR_L first increases and reaches maximum around the range, then decreases. This is because the points become more and more dispersed as i increases before reaching the range, but when h approaches the range and thereafter, they become more clustered due to the inclusion of many points with similar values. However, VMR_R will decrease gradually and become stable after the range, since the points become more similar as the lag increases. Since VMR_L and VMR_R behave in opposite directions when the lag is less than the range, the difference between VMR_L and VMR_R (denoted as DVmr) will be exaggerated. As a result, DVmr tends to reach maximum at the lag around the range. This lag is used as an estimate of the range.

The range derived from this algorithm is not always exactly equal to the assumed range, depending on the specific type of variogram. This happens especially when the curvature around the range is very small. For example, in the cases of Gaussian or Exponential models, the DVmr usually reaches its maximum before the assumed range. Therefore, for variograms having a form similar to Gaussian and Exponential models, the range estimated with the proposed method can be regarded as an "effective range," which is defined as the distance at which the variogram reaches 95% of its maximum.

3) Estimating the Range and Sill of a Sample Variogram: The assumption of the above method is that the variogram model

reaches its sill asymptotically when h approaches the range, which can be guaranteed only for bounded variogram models (e.g., spherical, exponential, circular, pentaspherical, and Gaussian models). In practice, this assumption may be violated in the following cases.

Case 1) There is no meaningful range or sill for the unbounded variograms, e.g., linear model [Fig. 4(a)] or power model [Fig. 4(b)].

Case 2) There are upward or downward semivariances after range for bounded variograms [Fig. 4(c) and (e)]. Numerous trials show that our algorithm can still detect the range and sill effectively as long as the upward/downward trend is not very obvious [Fig. 4(d) and (f)]. Note that a periodic variogram model could appear to have decreasing semivariances after range due to small window size [Fig. 4(g)].

Case 3) Range cannot be found when the maximum lag distance is less than the range due to small window [Fig. 4(h)]. In this case, we cannot find the range because there are continuously increasing semivariances.

Case 4) No range and sill exist for the pure nugget effect model.

For the cases above, we designed a series of nodes and developed a set of decision rules at each node to obtain range and its corresponding sill. The maxima of SEV_i and DVmr_j are denoted as $\text{SEV}_{\max}$ and $\text{DVmr}_{\max}$, respectively.

Decision Node 1: $\mathrm{SEV}_{max} = \mathrm{SEV}_1$ or $\mathrm{var(SEV)/mean}$ $\mathrm{(SEV)} < \alpha$

This situation corresponds to Case 4), a pure nugget effect model, where SEV_{max} appears at the first lag due to noise [Fig. 4(i)] or variance-to-mean ratio of SEV is less than a certain threshold [Fig. 4(j)]. Note that the idea of the index of dispersion is used once again to quantify the uniformity of the semivariances of a pure nugget effect model. Threshold α is determined empirically, and here we used 0.1. For the pure nugget effect model, the range and sill are not meaningful. However, for the purpose of classification, we define the range as zero and the sill as the semivariance at the first lag of the smoothed experimental variogram

$$\widehat{r}\mathrm{ange} = 0 \qquad (12)$$
$$\widehat{s}\mathrm{ill} = \mathrm{SEV}_1. \qquad (13)$$

Decision Node 2: $\mathrm{SEV}_{max} \neq \mathrm{SEV}_1$, and $\mathrm{DVmr}_{max} \neq \mathrm{DVmr}_{n-2}$

This implies that DVmr reaches its maximum in the lag range of $[2, n-3]$. This occurs for most bounded variogram models, especially for variograms described in Case 2). Consequently, we can obtain the estimate of range directly through

$$\widehat{r}\mathrm{ange} = i_{(\mathrm{DVmr}_i=\mathrm{DVmr}_{max})} \qquad (14)$$
$$\widehat{s}\mathrm{ill} = \mathrm{SEV}_{\widehat{r}\mathrm{ange}}. \qquad (15)$$

Decision Node 3: $\mathrm{SEV}_{max} \neq \mathrm{SEV}_1$ or SEV_n, and $\mathrm{DVmr}_{max} = \mathrm{DVmr}_{n-2}$

In this case, the range cannot be estimated, since DVmr is always increasing. However, the SEV reaches maximum at a certain lag in the middle. This usually occurs when there are downward semivariances after the range [Fig. 4(k)], while the VMmr does not reach its maximum around the range [Fig. 4(l)]. The lag where SEV is maximum is used as the estimate of range

$$\widehat{r}\mathrm{ange} = i_{(\mathrm{SEV}_i=\mathrm{SEV}_{max})}. \qquad (16)$$

Similarly, the estimate of sill is the semivariance at the estimated range in the smoothed variogram.

Decision Node 4: $\mathrm{SEV}_{max} \neq \mathrm{SEV}_1$, but $\mathrm{SEV}_{max} = \mathrm{SEV}_n$, and $\mathrm{DVmr}_{max} = \mathrm{DVmr}_{n-2}$

This occurs in all unbounded variogram models or when the maximum lag distance is less than the true range, as for Cases 1) and 3). There is a monotonic increase in both SEV and DVmr. We arbitrarily use the maximum lag as the estimate of range

$$\widehat{r}\mathrm{ange} = n \qquad (17)$$
$$\widehat{s}\mathrm{ill} = \mathrm{SEV}_n. \qquad (18)$$

To compare the SEVDV method and the one that uses semivariances directly, the semivariance at one pixel lag distance (γ_1) is also calculated. The advantage of retrieving γ_1 also lies in that it can be treated as the surrogate of the nugget, which is another important parameter of a variogram.

III. EXPERIMENT AND ANALYSIS

A. Study Site and Data Description

Our study site is located in the valley of a mountainous area in Xichang City, Sichuan Province of Western China, with an elevation range of 1500–2500 m [34]. A panchromatic IKONOS image of the study site, with 1-m spatial resolution, was acquired in December 2000, covering an area of 137 km^2. The IKONOS image has been georeferenced to Universal Transverse Mercator projection based on the 1984 World Geodetic System. A total of five image chips were selected, each representing a specific land cover type of this area: forest, muddy land, paddy field, fish pond, and village [Fig. 5(a)]. Muddy land refers to terrace areas subject to frequent floods which render them barren and useless for agriculture. The land cover types are coded from 1–5. The size of each image chip is 256×256 pixels. However, to circumvent edge effects, the calculation is based on a larger image area depending on the window size. For example, if the neighborhood window is 21 pixels long on its side, the image chip used for calculation is 276×276 pixels.

Since per-pixel classification method is not efficient for spatially heterogeneous land cover and land use classes in high spatial resolution imagery [21], we hope to test the effectiveness of the spatial information based on variogram in improving classification accuracy.

B. Handling Nonstationarity

In practice the stationarity assumption is often violated and it is unrealistic to assume a constant mean. This is especially true for natural scenes. When there is nonstationarity in the mean, (4) can be modified as

$$E(L(x)) = \mathrm{d}(x) \qquad (19)$$

where $\mathrm{d}(\cdot)$ is called the drift. Theoretically, estimation of a drift and the variogram of residuals cannot proceed simultaneously. The common strategy is to remove the drift using either a polynominial of finite order or a median polish method (MP) [35], [36]. Chappell *et al.* [5] investigated the stationary assumptions at five sites across a range of ecological environments in Africa and found that a quadratic polynomial can remove the regional trend quite well. Here, a quadratic polynominial regression is used to fit the gray-level value in each neighborhood window

$$L(x) = S^T(x)\beta + \varepsilon(x) \qquad (20)$$

where $S(x)$ is a vector consisting of $(1, x_1, x_2, x_1^2, x_2^2, x_1x_2)$, with x_1, x_2 being the coordinates of the pixels. β is the vector of coefficients, and $\varepsilon(x)$ is a zero-mean random variable representing fluctuations from this trend. The residual variogram is calculated based on $\varepsilon(x)$. Fig. 5(b) shows the detrended mosaic image using (20).

C. Robust and Omnidirectional Variogram

To reduce the effect of outliers, the "robust" variogram proposed by Cressie and Hawkins [36] is used instead of Matheron's "classical" method. The robust estimator called mean
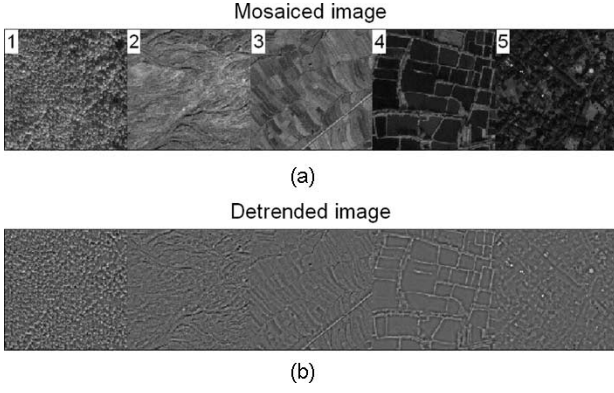
Mosaiced image

(a)

Detrended image

(b)

Fig. 5. (a) Image mosaic of five sampled land cover classes (1: forest, 2: muddy land, 3: paddy field, 4: fish pond, and 5: village) and (b) the detrended image using second-order trend surface regression.

square-root pair difference (SRPD) is calculated over a given lag by

$$\text{SRPD}(h) = \frac{1}{N_h} \sum |Z(x) - Z(x+h)|^{1/2}. \qquad (21)$$

Lark [15] compared the performance of SRPD and the corresponding semivariance, and his results show that SRPD gives better discrimination than semivariance for texture classification.

The omnidirectional instead of directional variogram is used, since in high spatial resolution imagery, the land cover unit (e.g., a house, pond, or paddy field) can be arranged in different orientations in space. And we care about the types of land cover rather than land covers with varying spatial orientations.

D. Choosing a Suitable Window Size

To make sure that the sample variogram converges, we need to increase the window size. However, if the window size is too large, homogeneity of texture can disappear and more computation will be involved [23], [28]. Window size is usually determined empirically, depending on the size of texture primitive in the image and the subject of interest. For example, Ramstein [28] showed that using a size of 11×11 pixels yielded favorable results when classifying land cover in many cities in France using a host of data types such as Landsat TM, AVHRR, and near-infrared photography. Chica-Olmo and Abarca-Hernandez [23] indicated that 7×7 was valid for most applications when discriminating three quaternary deposits using Landsat 5 TM imagery.

The SEVDV algorithm requires a larger window than usual, since we attempt to estimate the effective range. Consequently, window size should be greater than the ranges of most land covers. Experimentally, one must increase the window size to $3h_{\max}$ to obtain a statistically significant $\gamma(h)$ (for $h \in [h_{\min}, h_{\max}]$), [28]. Bearing this in mind, the window size should be at least three times the range. This requirement cannot always be satisfied in practice, since there are usually some land cover types with very large ranges. By experimenting, we found that the ranges of most land covers in our image were between four and seven pixels. Consequently, we selected window sizes of 13×13, 15×15, 17×17, 19×19, and 21×21 pixels.

The means of range for different window sizes

(a)

The means of sill for different window sizes

(b)

The means of γ_1 for different window sizes

(c)
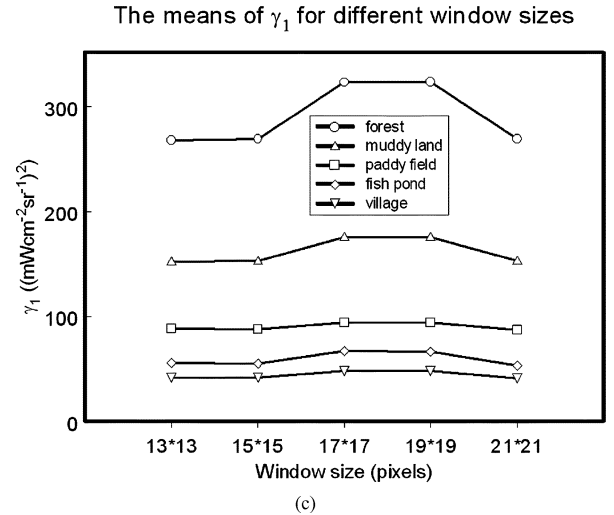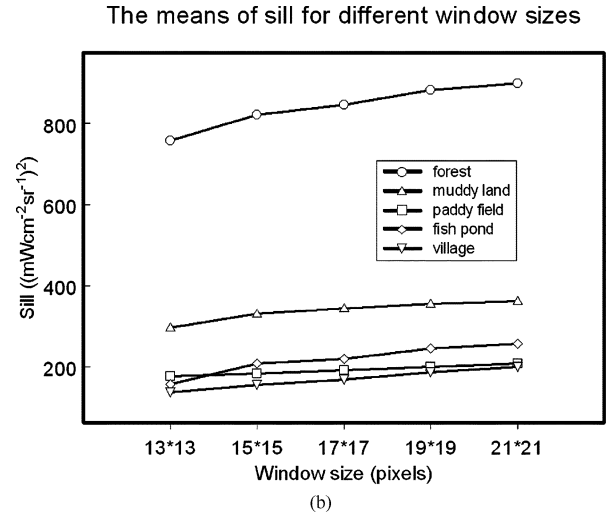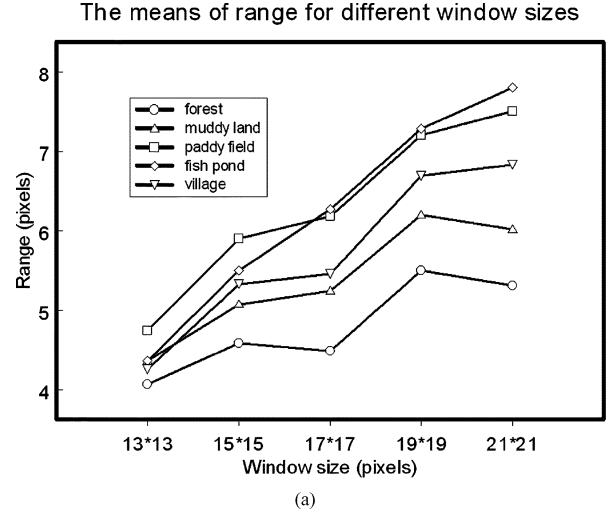
Fig. 6. Average of (a) range, (b) sill, and (c) γ_1 for five land cover classes with different window sizes 13×13, 15×15, 17×17, 19×19, and 21×21.

E. Analyzing Variogram Parameters

Range: Fig. 6(a) shows that the range of each land cover depends on the window size. Preliminary analysis reveals: 1) that in most cases, the range increases with the increment of

TABLE I
MEAN VALUES OF RANGE, SILL, AND SEMIVARIANCE AT LAG 1 FOR FIVE WINDOW SIZES

Range

Land cover	13×13	15×15	17×17	19×19	21×21
Forest	4.1	4.6	4.5	5.5	5.3
Muddy land	4.4	5.1	5.2	6.2	6.0
Paddy field	4.7	5.9	6.2	7.2	7.5
Fish pond	4.4	5.5	6.3	7.3	7.8
Village	4.3	5.3	5.5	6.7	6.8
Average	4.4	5.3	5.5	6.6	6.7

Sill

Land cover	13×13	15×15	17×17	19×19	21×21
Forest	756.6	821.5	846.7	882.5	897.8
Muddy land	296.1	330.6	344.0	355.3	362.3
Paddy field	175.6	182.8	191.1	198.7	207.4
Fish pond	157.2	208.1	219.4	245.3	256.1
Village	137.2	154.6	167.5	186.4	199.9
Average	304.5	339.5	353.8	373.7	384.7

γ_1

Land cover	13×13	15×15	17×17	19×19	21×21
Forest	268.2	269.4	323.4	323.9	269.7
Muddy land	152.1	152.9	175.7	175.8	152.9
Paddy field	88.4	88.0	94.0	93.9	87.2
Fish pond	55.6	55.2	66.7	66.0	53.2
Village	41.6	41.6	48.1	48.1	41.2
Average	121.2	121.4	141.6	141.5	120.8

window size. It can be found that the ranges at window sizes 15×15, 17×17, 19×19, and 21×21 increase by 20.9%, 26.5%, 50.6%, and 53.1%, respectively, when compared with the range at 13×13 (Table I). This might be because the variogram can capture texture with larger range when the window size is larger. 2) The mean ranges of the five kinds of land cover types are similar when the window size is 13×13. Nevertheless, separability among classes increases gradually as window size increases. This implies that increasing window size can help to discriminate texture from range.

The spatial pattern of ranges [Fig. 7(a) and (b)] shows that the range in a specific land cover is more homogeneous when window size is larger. It seems that estimates of range become more stable as window size increases. Large range occurs in heterogeneous areas (e.g., the forest with large gaps) or transition zones between two distinct cover types. Bright areas appear along the road in a paddy field and along the boundaries of fish ponds. The reason is that the trend in the two different types has not been completely removed by a second-order polynomial surface model. Consequently, the semivariances increase steadily and the estimate for range is large.

Sill: Similar to range, the mean sill for each land cover type increases with the window size [Fig. 6(b)], resulting in a 33.5% increment when varying window size from 13×13 to 21×21 (Table I). It is because more variability is usually induced when

the neighboring window is larger. The forest has much larger sills than other cover types due to the strong contrast of gray level between the peaks of trees and the shade sides of crowns. When observing the spatial arrangement of sill in the same land cover [Fig. 7(c) and (d)], the image tends to be smoother for larger window size. From sill, the paddy field and village are difficult to be discriminated.

γ_1: Unlike range and sill, γ_1 does not vary a lot (with only -0.1% and -1.1% fluctuation between window sizes 15×15 and 21×21) [Fig. 6(c)]. As mentioned earlier γ_1 is a surrogate of nugget. Nugget is the unstructured variance in the variogram, which arises partly from measurement error, or from spatially dependent variation that occurs over distances much smaller than the image pixel [5]. Therefore, γ_1 should not be affected much by window size. Like sill, larger window size will smooth the results and make γ_1 more homogeneous for each class [Fig. 7(e) and (f)].

F. Classification Accuracy

Analyses indicate that the distributions of sill and γ_1 are positively skewed for most classes except the forest class, which violates the normality assumption for class probability distribution of each ground cover class if maximum-likelihood classifier is used. To correct this, a logarithmic transformation (with a base 10) is applied to sill and γ_1.
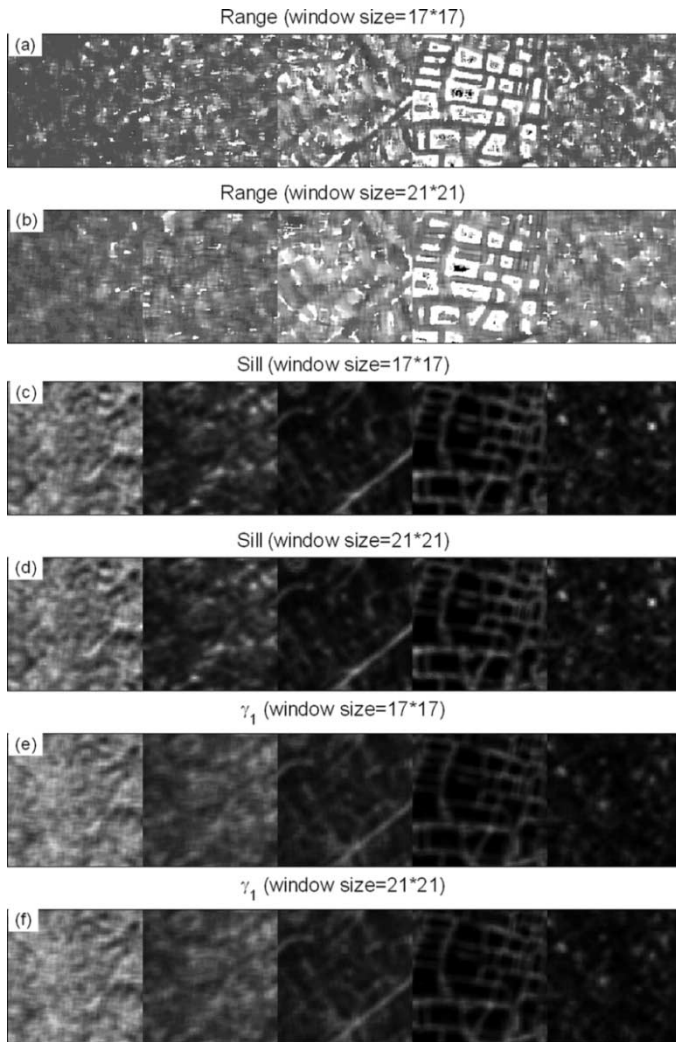
Range (window size=17*17)

Range (window size=21*21)

Sill (window size=17*17)

Sill (window size=21*21)

γ_1 (window size=17*17)

γ_1 (window size=21*21)

Fig. 7. Images of (a), (b) range, (c), (d) sill, and (e), (f) γ_1 at two different window sizes 17×17 and 21×21.

TABLE II
KAPPA STATISTIC OF CLASSIFICATION RESULTS. THE UNDERLINE INDICATES THE MAXIMUM KAPPA STATISTIC AMONG THE FIVE WINDOW SIZES

Bands \ Window size	13×13	15×15	17×17	19×19	21×21	Average
1) DN	0.24	0.24	0.24	0.24	0.24	0.24
2) DN, γ_1	0.58	0.60	0.66	_0.68_	0.65	0.63
3) DN, range	0.34	0.41	0.41	0.42	_0.45_	0.41
4) DN, sill	0.56	0.57	_0.59_	_0.59_	_0.59_	0.58
5) DN, γ_1, range	0.65	0.67	0.71	_0.72_	_0.72_	0.69
6) DN, γ_1, sill	0.66	0.70	0.70	_0.73_	0.69	0.70
7) DN, range, sill	0.60	_0.70_	0.66	0.66	0.69	0.66
8) DN, γ_1, range, sill	0.68	0.75	0.77	0.79	_0.82_	0.76
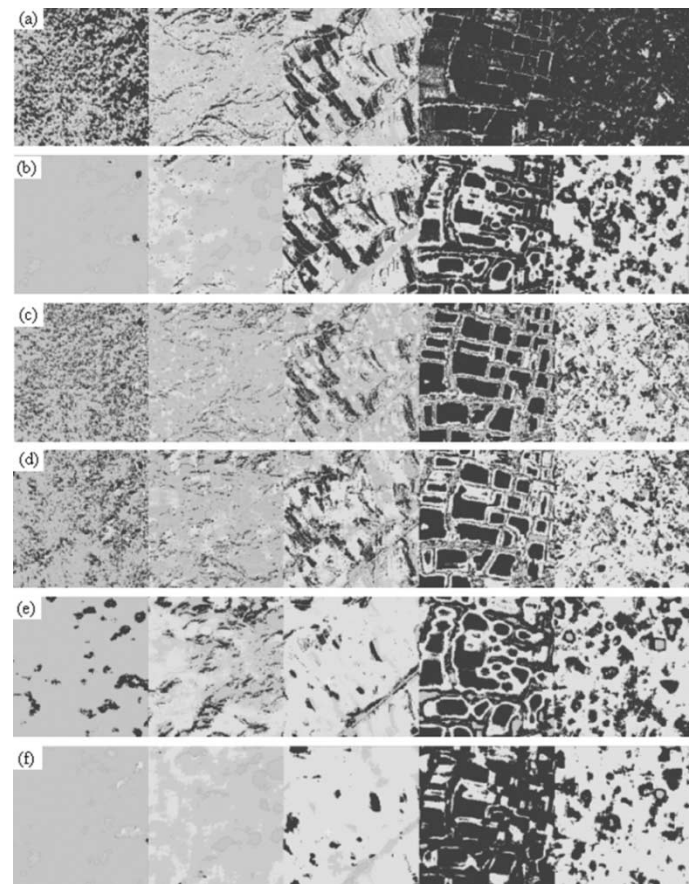
Fig. 8. Classification results using (a) DN, (b) DN and γ_1 (window size = 21×21 pixels), (c) DN and range (window size = 13×13 pixels), (d) DN and range (window size = 21×21 pixels), (e) DN and sill (window size = 21×21 pixels), and (f) DN, γ_1 and sill (window size = 21×21 pixels).

To examine the performance of classification using the variogram parameters, the classification results are compared for different combinations of digitial number (DN) and variogram parameters for five window sizes including: 1) DN, 2) DN, γ_1, 3) DN, range, 4) DN, sill, 5) DN, γ_1, range, 6) DN, γ_1, sill, 7) DN, range, sill, and 8) DN, γ_1, range, and sill (Table II). The Kappa statistic is used to assess classification accuracy against known validation data.

All of the parameters can increase the classification accuracy. Considering each parameter individually, we found that γ_1 increased the kappa statistic most significantly, with an average increase of 0.39, compared with the one using DN only. As the window size increases, the classification accuracy from DN and γ_1 exhibits a little fluctuation and reaches maximum when window size is 19. This phenomenon is consistent with the previous observation that each class's γ_1 is most separate when the window size is 17 or 19 (Fig. 6). Sill could also considerably enhance the accuracy, but its effects were not as strong as those of γ_1. For sill, the kappa statistic does not change when the window size increases from 17 to 21. Among the three parameters, range contributes least to classification accuracy improvement. Nevertheless, range is very effective for discriminating fish pond from

village in this image [Fig. 8(c)]. And its discrimination ability increases steadily when the window size increases from 13 to 21. In Fig. 8(c), the forest, muddy land, and paddy field cannot be easily distinguished from each other using range. However, when the window size increases to 21 [Fig. 8(d)], they are easier to be separated. This implies that the discriminating ability of range for texture classification is mainly limited by the window size. It is possible that the kappa statistic will increase further when the window size enlarges.

If any two of the three model parameters are combined with DN, the combination of γ_1 and range has relatively high classification accuracy. This is interesting because either γ_1 or sill has higher classification accuracy than range individually and they are expected to be the best group. This is due to the high correlation between γ_1 and sill, similar to the information redundancy among the semivariances of different lags. As expected, when all parameters are included, best classification accuracy can be obtained for all window sizes.

It can be found that much classification error occurs along the transition zone between two subclasses of land cover. There are several implications for "subclass." First, despite of efforts to choose the image chips with homogeneous texture, the homogeneity cannot always be guaranteed. For example, there is a village road (a subclass) in the paddy field chip, which is misclassified as muddy land in Fig. 8(f). Second, the concept of texture depends upon the scale on which we investigate. For example, the water and surrounding trails in the fish pond image chip interwind regularly. The water and trials can be thought of as one texture primitive or as separate texture primitives on a finer scale. Error usually occurs where two subclasses adjoin. This issue may be alleviated by increasing the size of the neighborhood window.

IV. Conclusion

The variogram model parameters, specifically range and sill, can be used to identify and quantify the spatial characteristics of land surface features and, therefore, can provide an effective alternative to remote sensing texture classification. In this paper, the SEVDV algorithm is proposed to extract range and sill for bounded variograms. For the purpose of textural classification, we design a set of decision rules to obtain the range and sill for all kinds of variograms. The case study shows that these decision rules work well. This is because for remote sensing imagery, there are few unbounded variograms. If the window size is increased to be large enough, in most cases we can find a specific bounded variogram model to capture the large-scale spatial variability.

Our analysis indicates that range is a very useful parameter for texture classification, which has not been automatically extracted before. However, the maximum accuracy from range can be obtained only when window size is relatively large, which requires more computation. We expect the rapid enhancement of computer hardware performance can alleviate such a requirement in the future. A further research focus is on reducing such a computational requirement.

We expect that the proposed method can be widely applied in geostatistical application besides remote sensing texture classification. The standard method of extracting range and sill involves intense interaction between the geostatistical analyst and computer software. Our method is able to automatically extract these two important parameters and, therefore, can be used to alleviate the workload of analysts and reduce the subjectivity of their calculations.

References

[1] P. J. Curran, "The semi-variogram in remote sensing: An introduction," *Remote Sens. Environ.*, vol. 24, pp. 493–507, 1988.

[2] C. E. Woodcock, A. H. Strahler, and D. L. B. Jupp, "The use of variograms in remote sensing I: Scene models and simulated images," *Remote Sens. Environ.*, vol. 25, pp. 323–348, 1988.

[3] F. Van Der Meer, "Classification of remotely-sensed imagery using an indicator kriging approach: Application to the problem of calcite-dolomite mineral mapping," *Int. J. Remote Sens.*, vol. 17, pp. 1233–1249, 1996.

[4] E. A. Addink and A. Stein, "A comparison of conventional and geostatistical methods to replace clouded pixels in NOAA-AVHRR images," *Int. J. Remote Sens.*, vol. 20, pp. 961–977, 1999.

[5] A. Chappell, J. W. Seqquist, and L. Eklundh, "Improving the estimation of noise from NOAA AVHRR NDVI for Africa using geostatistics," *Int. J. Remote Sens.*, vol. 22, pp. 1067–1080, 2001.

[6] R. Webster, P. J. Curran, and J. W. Munden, "Spatial correlation in reflected radiation from the ground and its implications for sampling and mapping by ground-based radiometry," *Remote Sens. Environ.*, vol. 29, pp. 67–68, 1989.

[7] P. M. Atkinson, "Optimal ground-based sampling for remote sensing investigations: Estimating the regional mean," *Int. J. Remote Sens.*, vol. 12, pp. 559–567, 1991.

[8] D. L. B. Jupp, A. H. Strahler, and C. E. Woodcock, "Autocorrelation and regularization in digital images I. Basic theory," *IEEE Trans. Geosci. Remote Sensing*, vol. 26, pp. 463–473, July 1988.

[9] P. M. Atkinson and P. J. Curran, "Defining an optimal size of support for remote sensing investigations," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 1–9, Jan. 1995.

[10] P. J. Curran and P. M. Atkinson, "Geostatistics and remote sensing," *Progr. Phys. Geography*, vol. 22, pp. 61–78, 1998.

[11] J. R. Carr, "Spectral and textural classification of single and multiple band digital images," *Comput. Geosci.*, vol. 22, pp. 849–865, 1996.

[12] J. R. Carr and F. P. Miranda, "The semivariogram in comparison to the co-occurrence matrix for classification of image texture," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1945–1952, Nov. 1998.

[13] U. C. Herzfeld, "A method for seafloor classification using directional variogram, demonstrated for data from the western flank of the Mid-Atlantic Ridge," *Math. Geol.*, vol. 25, pp. 901–924, 1993.

[14] U. C. Herzfeld and C. A. Higginson, "Automated geostatistical seafloor classification-principles, parameters, feature vectors, and discrimination criteria," *Comput. Geosci.*, vol. 22, pp. 35–52, 1996.

[15] R. M. Lark, "Geostatistical description of texture on an aerial photograph for discriminating classes of land cover," *Int. J. Remote Sens.*, vol. 17, pp. 2115–2133, 1996.

[16] F. P. Miranda, J. A. Macdonald, and J. R. Carr, "Application of the semivariogram textural classifier (STC) for vegetation discrimination using SIR-B data of Borneo," *Int. J. Remote Sens.*, vol. 13, pp. 2349–2354, 1992.

[17] F. P. Miranda, L. E. N. Fonseca, J. R. Carr, and J. V. Taranik, "Analysis of JERS-1(FUYO-1) SAR data for vegetation discrimination in Northwestern Brazil using the semivariogram textural classifier (STC)," *Int. J. Remote Sens.*, vol. 17, pp. 3523–3529, 1996.

[18] S. W. Myint, "Fractal approaches in texture analysis and classification of remotely sensed data: Comparisons with spatial autocorrelation techniques and simple descriptive statistics," *Int. J. Remote Sens.*, vol. 24, pp. 1925–1947, 2003.

[19] M. Philippe, "Comparing texture analysis methods through classification," *Photogram. Eng. Remote Sens.*, vol. 69, pp. 357–368, 2003.

[20] J. Le'vesque and D. J. King, "Spatial analysis of radiometric fractions from high-resolution multispectral imagery for modeling individual tree crown and forest canopy structure and health," *Remote Sens. Environ.*, vol. 84, pp. 589–602, 2003.

[21] P. Gong and P. J. Howarth, "Frequency-based contextual classification and gray-level vector reduction for land-use identification," *Photogramm. Eng. Remote Sens.*, vol. 58, pp. 423–437, 1992.

[22] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, pp. 786–804, 1979.

[23] M. Chica-Olmo and F. Abarca-Hernandez, "Computing geostatistical image texture for remotely sensed data classification," *Comput. Geosci.*, vol. 26, pp. 373–383, 2000.

[24] T. Randen and J. H. Husoy, "Filtering for texture classification: A comparative study," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, pp. 291–310, Apr. 1999.

[25] T. C. Bailey and A. C. Gatrell, *Interactive Spatial Data Analysis*. New York: Wiley, 1995.

[26] S. Berberoglu, C. D. Lloyd, P. M. Atkinson, and P. J. Curran, "The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean," *Comput. Geosci.*, vol. 26, pp. 385–396, 2000.

[27] P. M. Atkinson and P. Lewis, "Geostatistical classification for remote sensing: An introduction," *Comput. Geosci.*, vol. 26, pp. 361–371, 2000.

[28] G. Ramstein and M. Raffy, "Analysis of the structure of radiometric remotely-sensed images," *Int. J. Remote Sens.*, vol. 10, pp. 1049–1073, 1989.

[29] P. M. Atkinson and D. R. Emery, "Exploring the relation between spatial structure and wavelength: Implication for sampling reflectance in the field," *Int. J. Remote Sens.*, vol. 20, pp. 2663–2678, 1999.

[30] A. G. Journel and G. H. Huijbregts, *Mining Geostatistics*. New York: Academic, 1978.

[31] Insightful Corp., *S-PLUS 6 for Windows Guide to Statistics*. Seattle, WA: Insightful Corp., 2001, vol. 1.

[32] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. New York: Springer-Verlag, 2001.

[33] J. H. Friedman, "A variable span smoother," Dept. Statistics, Lab. Comput. Stat., Stanford Univ., Stanford, CA, Tech. Rep. 5, 1984.

[34] B. Xu, P. Gong, E. Seto, and R. Spear, "Comparison of gray level reduction schemes with a revised texture spectrum method for land-use classification using IKONOS imagery," *Photogramm. Eng. Remote Sens.*, vol. 69, pp. 529–536, 2003.

[35] N. Cressie and G. F. Glonek, "Median based covariogram estimators reduce bias," *Stat. Prob. Lett.*, vol. 2, pp. 299–304, 1984.

[36] N. Cressie and D. M. Hawkins, "Robust estimation of the variogram," *Math. Geol.*, vol. 12, pp. 115–125, 1980.

Qi Chen received the B.S. degree in geography and M.S. degree in remote sensing and GIS from Nanjing University, Nanjing, China, in 1998 and 2001, respectively. He is currently pursuing the Ph.D. degree at the University of California, Berkeley.

His current research interests include remote sensing image classification, biophysical and biochemical parameters extraction from remotely sensed image, terrestrial ecosystem modeling, and global climate change.

Peng Gong received the B.S. and M.S. degrees from Nanjing University, Nanjing, China, in 1984 and 1986, respectively, and the Ph.D. degree from the University of Waterloo, Waterloo, ON, Canada, in 1990.

He is currently a Professor in the Department of Environmental Science, Policy, and Management, University of California, Berkeley, and Co-Director of the Center for Assessment and Monitoring of Forest and Environmental Resources (CAMFER), University of California, Berkeley. He was with York University, Toronto, ON, and the University of Calgary, Calgary, AB, Canada, before joining the University of California, Berkeley, in 1994. His research interests include photoecometrics, global change monitoring, and the role of technology in society. He is an author/coauthor of over 200 papers and five books. He serves as the Director of the International Institute for Earth System Science, Nanjing University, Editor-in-Chief of *Geographic Information Sciences*, and Editor for the *International Journal of Remote Sensing*.