Texture classification based on comparison of second-order statistics. I. Two-point probability density function estimation and distance measure

Alexei A. Goon

School of Optics and Center for Research and Education in Optics and Lasers, 4000 Central Florida Boulevard, University of Central Florida, Orlando, Florida 32816

Jannick P. Rolland

School of Optics and Center for Research and Education in Optics and Lasers , Department of Electrical Engineering and School of Computer Science, 4000 Central Florida Boulevard, University of Central Florida, Orlando, Florida 32816

Received September 1, 1998; revised manuscript received January 28, 1999; accepted February 23, 1999

The two-point probability density function (2P-PDF) gives a full description of the first- and second-order statistics of a random process. We propose a framework for texture classification based on a distance measure between 2P-PDF's after equalization of first-order statistics. This framework allows extraction of the structural information of the process independently of the dynamic range of the image. We present two methods for estimating the 2P-PDF of texture images, and we establish some criteria for efficient computation. The theoretical framework for noise-free texture images is validated with four texture ensembles. © 1999 Optical Society of America [S0740-3232(99)00807-8]

OCIS codes: 100.2960, 100.5010.

1. INTRODUCTION

Research in imaging aims to create better imaging systems and to develop methods of image processing and analysis that utilize the most important information in an image.[1] Information may take, for example, the form of defined targets embedded in complex backgrounds such as tumors in chest radiographs or mammograms or that of unique statistical textures such as liver scans or mammography textures.[2–7] We have recently focused our research on statistical texture synthesis and analysis because the mathematical modeling and characterization of such complex backgrounds will provide an important step forward in the quantitative assessment of image quality not only in medical imaging but also in material science and perhaps in the science of art.[8]

In this paper we present a method of texture classification based on an analysis and estimation of the full second-order statistics. Effective texture classification requires two problems to be solved: (1) defining a set of distinct classes and (2) finding a criterion for classification. The former is more challenging. Texture classes will be distinguished solely on the basis of their complete second-order statistics. Specifically, classification will be based on a proposed distance measure between the second-order probability density functions of the textures equalized in first-order statistics.[9,10] The second-order probability density function is referred to as the normalized two-point probability density function (2P-PDF).

Effectiveness in classification depends on how a chosen distance measure yields distinct classes. Figure 1 illustrates this problem. Suppose that a set of texture realizations can be classified as either class A or class B. Each realization in the set of textures has an associated 2P-PDF. We propose that texture images be considered elements in a metric space with a distance function d, defined below by Eq. (3), that measures the distance between their 2P-PDF's. In Fig. 1, for example, d_{AB} denotes the distance between two ensembles of texture images A and B. The simplest way to estimate the 2P-PDF is the relative-frequency method described in Subsection 3.B. We shall show, however, that the 2P-PDF can also be estimated with the maximum-entropy method described in Subsection 3.C. In that case a set of moments $\{\mu_{ij}\}$ determines the 2P-PDF. Given the 2P-PDF associated with each texture realization within a class ensemble (e.g., class A), it is possible to find the average set of moments $\langle\{\mu_{ij}\}\rangle_A$ for that class. In this paper angle brackets denote ensemble averaging. Consequently, the average 2P-PDF's $\rho^A(g_1, g_2, \Delta\mathbf{r})$ and $\rho^B(g_1, g_2, \Delta\mathbf{r})$ for the texture classes A and B are estimated with the averaged sets of moments for classes A and B, respectively, where $\Delta\mathbf{r}$ is the vector separation between two points of the random process. In this paper bold type denotes a vector. We then take the distance between two classes to be the distance between their average 2P-PDF's.

Next we need to define the spread of textures within a given class. Let us consider a texture class C and denote its average 2P-PDF as $\rho^C(g_1, g_2, \Delta\mathbf{r})$. There is a probability distribution function (i.e., the cumulative probability density function) associated with the distance between the averaged 2P-PDF and the 2P-PDF's of textures of class C. The radius of the class is defined as the distance at which the probability distribution function reaches a specific value, for instance, 0.99. We denote this radius as $r^C_{0.99}$, where the upper index indicates the class and the lower index indicates the value of the cumulative distribution. The larger the class radius, the broader the variety of textures included in that class. The value of the cumulative distribution, always less or equal to 1, can be chosen specifically for each problem. For a given probability distribution function the radius of the class can be related to the root-mean-square distance σ between the 2P-PDF's of textures with respect to the ensemble average 2P-PDF within that class. For instance, if the class has a uniform distribution of distances it can be shown that $r^C_{0.99} = 1.73\sigma$; for a Gaussian distribution $r^C_{0.99} = 2.32\sigma$, and for an exponential distribution $r^C_{0.99} = 3.26\sigma$. Therefore each texture class can now be defined by its average 2P-PDF and its radius, which as a pair form the basis for solving problem (1).

The criteria for discrimination, the second problem posed, is simple. For a given texture realization the 2P-PDF $\rho(g_1, g_2, \Delta\mathbf{r})$ is estimated. The distances d^A between $\rho(g_1, g_2, \Delta\mathbf{r})$ and the average 2P-PDF of class A, $\rho^A(g_1, g_2, \Delta\mathbf{r})$, and d^B between $\rho(g_1, g_2, \Delta\mathbf{r})$ and the average 2P-PDF of class B are estimated. Then the texture is assigned to belong to class A or B by, for instance, maximum-likelihood classification.[11,12] This approach can be generalized to an arbitrary number of classes.

The question is how close can classes be in texture space for texture realizations to be assigned correctly to their class ensemble? The efficacy in classifying a texture realization in class A or B, for example, is naturally a function of d_{AB}, $r^A_{0.99}$, and $r^B_{0.99}$ represented in Fig. 1. From a geometrical point of view, the efficacy requires distinct classes or

$$d_{AB} > r^A_{0.99} + r^B_{0.99}. \tag{1}$$

This condition implies that the spheres in texture space do not overlap, as illustrated in Fig. 1. This is naturally what we wish to obtain for a chosen distance and an associated spread measure. The textures cannot be discriminated by use of classes A and B if

$$r^A_{0.99} > d_{AB} + r^B_{0.99}. \tag{2}$$

This condition implies that sphere B is inside sphere A, and in this case the textures are fully undiscriminable. In these two extreme cases, the classification errors are minimal and maximal, respectively. In general, the classification error will be proportional to the classes' overlap.

In this paper we first review how we define the distance d between two texture realizations.[9–10] We then propose a method for the effective estimation of 2P-PDF's. Given an ensemble of texture images from a class, we then derive an expression for the value of σ within that class. Finally, we validate the expression found for σ and test the separability of four texture ensembles on the basis of the proposed distance measure d and the associated classes' radii.

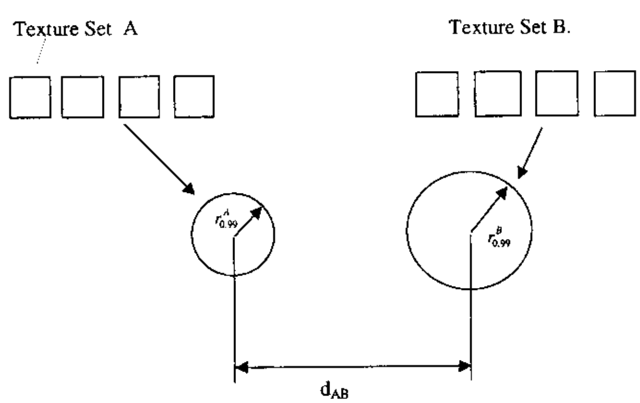

Fig. 1. Texture space.

2. DISTANCE MEASURE BETWEEN TWO PROBABILITY DENSITY FUNCTIONS

Given a random process, the first-order, or one-point, probability density function (1P-PDF) gives nonstructural information about the process (e.g., the mean value of the gray level on the number of pixels in a given range of gray levels). The second- and higher-order PDF's provide structural information about the process.[13] This can be understood by considering a representation of the joint probability distribution function of two gray levels, for example, $p(g_1, g_2)$ given as $p(g_1, g_2) = p(g_1|g_2)p(g_2)$, where $p(g_2)$ is the 1P-PDF and $p(g_1|g_2)$ is the conditional probability distribution. The conditional probability distribution is related to the structural information of the process.

In this paper we compute 2P-PDF's using two methods and apply the framework to texture images that we consider two-dimensional stationary random processes. While a method for the calculation of the distance between textures, based on the 2P-PDF,[9–10] was recently proposed, the crucial part of the algorithm is the estimation of the 2P-PDF for a given texture realization. The 2P-PDF $\rho(g_1, g_2, \mathbf{r}_1, \mathbf{r}_2)$ of a random process $g(\mathbf{r})$ is defined to be equal to the joint probability distribution of random variables $g(\mathbf{r}_1)$ and $g(\mathbf{r}_2)$. In the case of two-dimensional statistical textures, the random process $g(\mathbf{r})$ denotes the gray level at a point $\mathbf{r}$ in a texture image. For a stationary random process, $\rho(g_1, g_2, \mathbf{r}_1, \mathbf{r}_2)$ reduces to $\rho(g_1, g_2, \Delta\mathbf{r})$, where $\Delta\mathbf{r}$ is the vector separation between two points of the texture. Suppose we are given the texture images a and b and we let $\rho^a(g_1, g_2; \Delta\mathbf{r})$ and $\rho^b(g_1, g_2; \Delta\mathbf{r})$, respectively, be their estimated 2P-PDF's. The proposed distance measure for the stationary random processes is then given by the L_2 distance between probability density functions:

$$d = \left\{ \sum_{\Delta\mathbf{r}} \int\int [\rho^a(g_1, g_2, \Delta\mathbf{r}) - \rho^b(g_1, g_2, \Delta\mathbf{r})]^2 \mathrm{d}g_1 \mathrm{d}g_2 \right\}^{1/2}, \tag{3}$$

where the summation is over the finite set of vectors $\Delta\mathbf{r}$ and the double integration is over the range of gray scales. The set of vectors $\Delta\mathbf{r}$ can be considered finite for two reasons: (1) The texture images consist of discreet elements (i.e., pixels) whose size set a lower bound and (2) the 2P-PDF reduces to the product of 1P-PDF's for large separations $|\Delta\mathbf{r}|$. In this case, because the 1P-PDF's are the same for the textures, the corresponding terms in sum (3) vanish, which sets an upper bound. The summation over a range of distances makes it possible to avoid the problem of selecting an optimal parameter $\Delta\mathbf{r}$ for the computation of 2P-PDF's.[14,15] The adoption of this summation is critical to the method proposed because it allows the use of the structural information at multiple scales and orientations of the texture. Although the probability density function is always integrable, it may not always be square integrable. In our case, however, the methods of 2P-PDF estimation always gave finite values of the probability density functions, and because the image had a finite range of gray levels, 2P-PDF estimations were always square integrable. Although various types of distance function give similar results, further investigation is required to fully address this issue. To compute d accurately, the algorithm for estimating 2P-PDF's must provide the needed accuracy. The inaccuracy in estimating the 2P-PDF's would affect the distance between textures because it is equivalent to adding noise to the PDF's.

3. ESTIMATION OF PROBABILITY DENSITY FUNCTIONS

We consider two methods of estimating a PDF: (1) the relative-frequency[16] and (2) the maximum-entropy (ME) methods.[17] Both methods naturally consider an ensemble of realizations of a given random process. The relative-frequency method is equivalent to computing the relative frequencies of occurrence of a given parameter for the ensemble of realizations of the random process. For example, the relative-frequency method computes the occurrence of a gray-level value in the case of the 1P-PDF and of a pair of gray-level values for the 2P-PDF.

The ME method estimates a PDF as the maximum of the entropy functional

$$\int \rho(\mathbf{x}) \cdot \ln \rho(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{4}$$

with appropriate constraints in the form of moments given by

$$\mu_{\mathbf{i}} = \int \mathbf{x}^{\mathbf{i}} \cdot \rho(\mathbf{x}) \mathrm{d}\mathbf{x}. \tag{5}$$

In these expressions, $\mathbf{x}$ denotes a vector of random variables, $\mathrm{d}\mathbf{x} = \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_n$, $\rho(\mathbf{x})$ is the joint probability density function of these random variables, and $\mathbf{x}^{\mathbf{i}} = x_1{}^{i_1} x_2{}^{i_2} \ldots x_n{}^{i_n}$.

For the stationary texture images considered in this paper, $\rho(g_1, g_2, \Delta\mathbf{r})$ is obtained with the relative-frequency method by counting the number of occurrences of pairs of gray levels (g_1, g_2) separated by $\Delta\mathbf{r}$ in the ensemble of realizations of a given process. The ME method can be applied directly to the estimation of the 2P-PDF distribution of texture images if we denote $\mathbf{x}$ as (g_1, g_2) where g_1 and g_2 are the gray levels of pixels separated by vector $\Delta\mathbf{r}$. A question is how many samples should be considered to estimate a probability density function with known accuracy by use of the relative-frequency method? Similarly, how many moments should be considered in the ME method? We shall first address the latter question for the 1P-PDF and then extend the method to the 2P-PDF before revisiting both questions.

A. First-Order Probability Density Function Estimation with the Maximum-Entropy Method

The following problem needs to be solved: Find $\rho(g)$ such that

$$\int \rho(g) \ln \rho(g) \mathrm{d}g \rightarrow \max, \quad \int g^i \rho(g) \mathrm{d}g = \mu_i,$$
$$\mu_i = E\{g^i(r)\}, \qquad i = 0, N, \tag{6}$$

where $E\{\}$ denotes the averaging over location in a single realization. The larger the number of moments, the more difficult the problem becomes in a computational sense because a system of nonlinear equations needs to be solved. In addition, a function $\rho(x)$ can be overestimated. For instance, in the case of a Gaussian random process, the mean and the standard deviation constitute sufficient parameters for estimating the PDF. If we consider higher-order moments, computational errors lead to a decrease in accuracy in the estimation of the PDF. To find the optimal number of moments to estimate a PDF, the first k moments were computed from averaging over the locations of a random process realization and the PDF estimated. The next m moments were then calculated with estimated PDF $\rho(g)$ and compared with the moments obtained from averaging over the locations in the process by computing an error function given by

$$\varepsilon(k) = \sum_{i=k+1}^{m+k+1} \left(E\{g^i\} - \int g^i \rho(g) \mathrm{d}g \right)^2. \tag{7}$$

Figure 2 shows the dependence of ε with k for a realization of a particular texture. It is observed that $\varepsilon(k)$ has a sharp decline at $k = 4$. This finding was observed for all realizations of the random process considered (i.e., eight in this case). We then postulate that the opti-

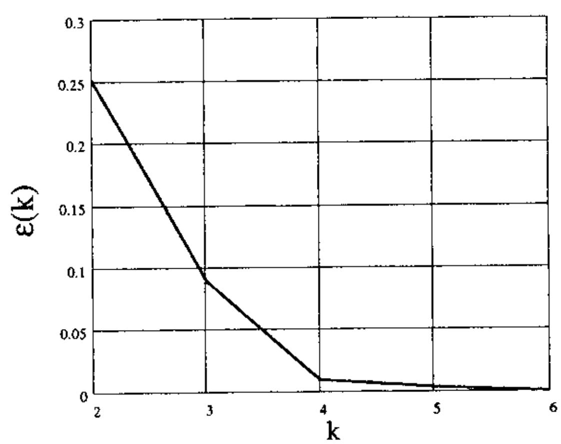

Fig. 2. $\varepsilon(k)$ defined by Eq. (7) as a function of the number of moments used in the estimation of the first-order probability distribution for one realization of a texture image.

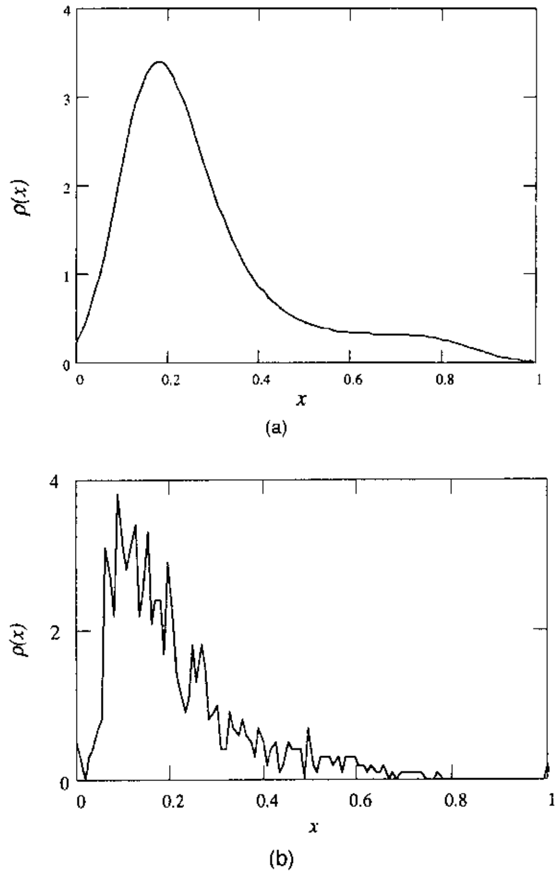

Fig. 3. (a) 1P-PDF computed with the ME method. Note that the x axis corresponds to gray levels normalized to the range 0–1. (b) 1P-PDF computed with the relative-frequency method. This is also known as the gray level histogram of a texture image.

mal number of moments to consider for this random process is five ($i = 0,...4$). It is important to note that the optimal number of moments for ME estimation depends on the process: The more complex the probability distribution function of the process, the more moments are likely to be needed. Figures 3(a) and 3(b) show the PDF calculated with the ME method on the basis of five moments and the relative-frequency method, respectively.

B. Second-Order Probability Density Function Estimation with the Relative-Frequency Method

The 2P-PDF is now estimated for a texture random process. In this estimation it is assumed that the texture is a stationary random process. The relative-frequency method was the first employed to calculate the number of occurrences of a given pair of a gray levels (g_1, g_2) for a given displacement vector $\mathbf{r}$ between two pixels.[16] This method is analogous to the relative-frequency method used for estimation of the 1P-PDF. The relative-frequency method, for a stationary random process, is illustrated in Fig. 4.

C. Second-Order Probability Density Function Computation with the Maximum-Entropy Method

The ME method is now considered with use of all moments of the form

$$\mu_{ij}(\Delta\mathbf{r}) = E\{g^i(\mathbf{r}) \cdot g^j(\mathbf{r} + \Delta\mathbf{r})\}, \qquad i, j = 0,...4, \quad (8)$$

where $g(\mathbf{r})$ is the value of the gray level at position $\mathbf{r}$ in the texture image. The range $i, j = 0,...4$ leading to 25 moments was chosen because five moments allowed us to estimate accurately the first-order probability distribution. Consequently, this range of i and j should allow us to estimate correctly the second-order probability distribution when the distance $|\Delta\mathbf{r}|$ is large and gray levels $g(\mathbf{r})$ and $g(\mathbf{r} + \Delta\mathbf{r})$ are independent. We assume that the algorithm also gives correct results for a smaller spacing $|\Delta\mathbf{r}|$.

When using ME, one must find the function $\rho(g_1, g_2, \Delta\mathbf{r})$ that maximizes the value of the functional

$$\int\int \rho(g_1, g_2, \Delta\mathbf{r}) \cdot \ln[\rho(g_1, g_2, \Delta\mathbf{r})] \mathrm{d}g_1 \mathrm{d}g_2 \rightarrow \max \quad (9)$$

and satisfies the following constraints:

$$\int\int g_1{}^i g_2{}^i \rho(g_1, g_2, \Delta\mathbf{r}) \mathrm{d}g_1 \mathrm{d}g_2 = \mu_{ij}(\Delta\mathbf{r}), \qquad i, j = 0, N. \quad (10)$$

The solution of this problem is given by[17]

$$\rho(g_1, g_2, \Delta\mathbf{r}) = \exp\left(\sum_{i,j=0}^{N} \beta_{ij} g_1{}^i g_2{}^i \right). \quad (11)$$

The coefficients β_{ij} can be determined by substituting the expression for $\rho(g_1, g_2, \Delta\mathbf{r})$ given by Eq. (11) into Eq. (10). Coefficients β_{ij} are functions of $\Delta\mathbf{r}$, and generally $\beta_{ij} \neq \beta_{ji}$. In the specific case of textures considered here, it is, however, possible to postulate that $\mu_{ij}(\Delta\mathbf{r}) = \mu_{ij}(-\Delta\mathbf{r}) = \mu_{ji}(\Delta\mathbf{r})$. The symmetry of the variables g_1 and g_2 in Eqs. (9)–(11) and the lack of a specific origin for the random process yield $\beta_{ij} = \beta_{ji}$. Results of the 2P-PDF estimations are shown in Fig. 5, where Figs. 5(a) and 5(b) show the 2P-PDF calculated with the ME method and the relative-frequency method, respectively. As expected, the ME method gives a smoother estimate.

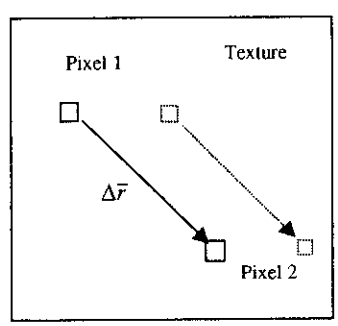

Fig. 4. Illustration of the computation of the 2P-PDF (co-occurrence matrix) with the relative-frequency method.

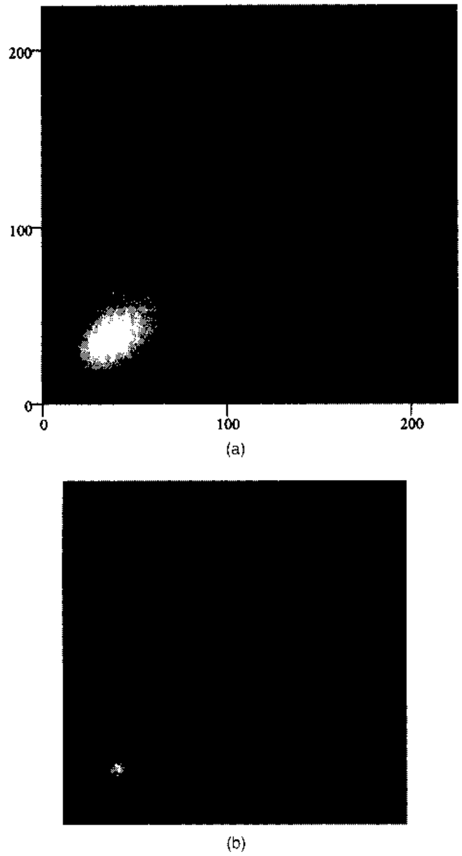

Fig. 5. (a) 2P-PDF calculated with the ME method. (b) 2P-PDF of a textured image calculated with the relative-frequency method.

4. ESTIMATION OF THE RADIUS OF A TEXTURE CLASS BASED ON TWO-POINT PROBABILITY DENSITY FUNCTION STATISTICS

Given that 2P-PDF's can now be estimated accurately and efficiently, we shall compute the spread σ in the distances between 2P-PDF's of textures belonging to a class and the ensemble-averaged 2P-PDF. Given an ensemble of d values between 2P-PDF's of textures and the ensemble-averaged 2P-PDF, σ is defined as

$$\sigma = \sqrt{\langle d^2 \rangle}. \tag{12}$$

Suppose that we are given a realization a of a particular texture ensemble. The distance measure between its 2P-PDF and the ensemble-averaged 2P-PDF is given by Eq. (3). Following the ME method, the 2P-PDF can be approximated with the needed accuracy by the function $\rho(g_1, g_2, \Delta\mathbf{r})$ given by Eq. (11). Let us fix vector $\Delta\mathbf{r}$ and denote $\rho(g_1, g_2, \Delta\mathbf{r})$ as $\rho(g_1, g_2)$. Let us suppose that the ME approximations of the ensemble-averaged distribution $\rho = \rho(g_1, g_2, \beta_{ij})$ and the distribution associated with texture a, $\rho^a = \rho^a(g_1, g_2, \beta'_{ij})$, are known. Given that texture a belongs to the ensemble, the difference between its 2P-PDF and the ensemble-averaged 2P-PDF is assumed to be small. Distribution ρ^a can then be decomposed into a Taylor expansion in the vicinity of ρ:

$$\rho^a(g_1, g_2, \beta'_{ij}) = \rho(g_1, g_2, \beta_{ij}) + \sum_{ij} \left[\frac{\partial \rho(g, g, \beta_{ij})}{\partial \beta_{ij}} (\Delta\beta_{ij}) \right], \tag{13}$$

where $\Delta\beta_{ij} = \beta'_{ij} - \beta_{ij}$. Substituting Eq. (13) into the expression of d^2 obtained from Eq. (3) for a given $\Delta\mathbf{r}$,

$$d^2 = \int\int \left\{ \sum_{ij} \left[\frac{\partial \rho(g_1, g_2, \beta_{ij})}{\partial \beta_{ij}} (\Delta\beta_{ij}) \right] \right\}^2 \mathrm{d}g_1 \mathrm{d}g_2 = \sum_{ij} \sum_{kl} \Delta\beta_{ij} \Delta\beta_{kl} A_{ijkl}, \tag{14}$$

where

$$A_{ijkl} = \int\int \left[\frac{\partial \rho(g_1, g_2, \beta_{ij})}{\partial \beta_{ij}} \frac{\partial \rho(g_1, g_2, \beta_{kl})}{\partial \beta_{kl}} \right] \mathrm{d}g_1 \mathrm{d}g_2, \quad i, j, k, l \leq N \tag{15}$$

and N is the number of moments used in the ME method.

According to Eq. (11),

$$\frac{\partial \rho(g_1, g_2, \Delta\mathbf{r})}{\partial \beta_{ij}} = g_1{}^i g_2{}^i \rho(g_1, g_2, \Delta\mathbf{r}). \tag{16}$$

Substituting Eq. (16) into Eq. (15) yields

$$A_{ijkl} = \int\int g_1^{i+k} g_2^{j+l} \rho^2(g_1, g_2, \Delta\mathbf{r}) \mathrm{d}g_1 \mathrm{d}g_2 = \langle g^{i+k}(\mathbf{r}) g^{j+l}(\mathbf{r} + \Delta\mathbf{r}) \rho[g(\mathbf{r}), g(\mathbf{r} + \Delta\mathbf{r}), \Delta\mathbf{r}] \rangle. \tag{17}$$

Two expressions for the computation of coefficients A_{ijkl} are given by Eq. (17). The first expression necessitates computing the 2P-PDF first and then computing the value of the double integrals. The second expression allows one to estimate the 2P-PDF from available data and then find an ensemble average. The two approaches yield the same result if the 2P-PDF is estimated with the needed accuracy. Thus the elements of matrix A_{ijkl} can be computed with Eq. (17), and the distance between a pair of texture images can then be calculated with the simple expression for d given by Eq. (14).

The values of the random variables $\Delta\beta_{ij}$ are not easily accessible from the ensemble of textures. The values of the random deviations $\Delta\mu_{ij}(\Delta\mathbf{r})$ of moment $\mu_{ij}(\Delta\mathbf{r})$ from its average can be obtained instead. To obtain values of $\Delta\beta_{ij}$ from known values of $\Delta\mu_{ij}(\Delta\mathbf{r})$, the system of constraints given by Eq. (10) is decomposed into a Taylor expansion in the vicinity of the solution

$$\sum_{kl} \left\{ \frac{\partial \left[\int\int g_1{}^i g_2{}^j \rho(g_1, g_2, \Delta\mathbf{r}) \mathrm{d}g_1 \mathrm{d}g_2 \right]}{\partial \beta_{kl}} \cdot \Delta\beta_{kl} \right\} = \Delta\mu_{ij}(\Delta\mathbf{r}). \tag{18}$$

It is possible to differentiate with respect to variables β_{kl} under the sign of the integral in Eq. (18). With use of Eq. (11), Eq. (18) then yields

$$\frac{\partial\left[\int\int g_1{}^i g_2{}^j \rho(g_1, g_2, \Delta\mathbf{r})\mathrm{d}g_1\mathrm{d}g_2\right]}{\partial\beta_{kl}} = \int\int g_1^{i+k} g_2^{j+l}\rho(g_1, g_2, \Delta\mathbf{r})\mathrm{d}g_1\mathrm{d}g_2 = \mu_{i+k,j+l}(\Delta\mathbf{r}) = B_{ijkl}. \quad (19)$$

From Eq. (18) and Eq. (19) it follows that

$$\sum_{kl} B_{ijkl}\Delta\beta_{kl} = \Delta\mu_{ij}(\Delta\mathbf{r}). \quad (20)$$

Random variables $\Delta\beta_{ij}$ can then be expressed as

$$\Delta\beta_{kl} = \sum_{ij} (B^{-1})_{klij}\Delta\mu_{ij}(\Delta\mathbf{r}). \quad (21)$$

Substituting Eq. (21) into Eq. (14), it follows that

$$d^2 = \sum_{ij}\sum_{kl} [(B^{-1})^T A B^{-1}]_{ijkl}\Delta\mu_{ij}\Delta\mu_{kl}. \quad (22)$$

The square of the distance d^2 between any 2P-PDF of the texture of an ensemble and the ensemble-averaged 2P-PDF is a random variable and can thus be computed with Eq. (22) for every texture in the ensemble. Equation (22) represents d^2 as a quadratic form of the random variables $\Delta\mu_{ij}$, and according to Eq. (12) the expectation value of d^2 is required for computing σ. Let us denote C as the correlation matrix of $\Delta\mu_{ij}$:

$$C_{ijkl} = E\langle\Delta\mu_{ij}\Delta\mu_{kl}\rangle. \quad (23)$$

σ is then expressed as

$$\sigma^2 = E\langle d^2\rangle = \sum_{ij}\sum_{kl} [(B^{-1})^T A B^{-1}]_{ijkl} C_{ijkl}. \quad (24)$$

To simplify, the lexicographic representation of fourth-rank tensors A, B, and C would be used:

$$A_{ijkl} \to A_{mn}, \quad B_{ijkl} \to B_{mn}, \quad C_{ijkl} \to C_{mn}; \quad i,j,k,l \leq N; \quad m,n \leq N^2. \quad (25)$$

The expectation value of the square of the distance within the texture ensemble for a given separation $\Delta\mathbf{r}$ is then given by[18]

$$E\langle d^2\rangle = \mathrm{tr}[(\mathbf{B}^{-1})^T \quad \mathbf{A}^T \quad \mathbf{B}^{-1} \quad \mathbf{C}] = \sigma^2(\Delta\mathbf{r}). \quad (26)$$

Matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are functions of separation $\Delta\mathbf{r}$. Consequently, the distance within the ensemble is given by[18]

$$\sigma = \left[\sum_{\Delta\mathbf{r}} \sigma^2(\Delta\mathbf{r})\right]^{1/2} = \left(\sum_{\Delta\mathbf{r}} \mathrm{tr}\{[\mathbf{B}^{-1}(\Delta\mathbf{r})]^T[\mathbf{A}(\Delta\mathbf{r})]^T\mathbf{B}^{-1}(\Delta\mathbf{r})\mathbf{C}(\Delta\mathbf{r})\}\right)^{1/2}. \quad (27)$$

The lexicographic representation is used here to overcome difficulties in finding an invariant form of Eq. (24) that employs fourth-rank tensors. But, on the other hand, use of fourth-rank tensors instead of lexicographic representation permits investigation of possible symmetries of the texture classification problem and needs to be further investigated.

It is possible to write an explicit expression for the probability distribution function of d^2 within an ensemble of texture realizations. Again, let us fix vector $\Delta\mathbf{r}$. In this case, the probability distribution function can be written as[18]

$$P\{d^2 \leq D^2\} = \int_S \rho(\mu_{01}, \mu_{02}, \ldots, \mu_{nm})\mathrm{d}\mu_{01}\mathrm{d}\mu_{02}\ldots\mathrm{d}\mu_{nn}, \quad (28)$$

where $\rho(\mu_{01}, \mu_{02}, \ldots, \mu_{nn})$ is the PDF of the random variables μ_{ij} and S is the region of integration for points satisfying the condition

$$\sum_{ij}\sum_{kl} [(B^{-1})^T A B^{-1}]_{ijkl}\Delta\mu_{ij}\Delta\mu_{kl} \leq D^2. \quad (29)$$

This set of points represents an n^2-dimensional ellipse. This distribution can be simplified for computational reasons if we use the fact that the quadratic forms associated with tensors $(B^{-1})^T A^T B^{-1}$ and C can be reduced simultaneously to the sum of squares because at least one is positive definite. The knowledge of the distribution function of d can be useful in texture classification problems based on maximum-likelihood estimation.

Let us find a relationship between the root-mean-square distance σ' between 2P-PDF's of textures and the root-mean-square distance σ between 2P-PDF's of textures and ensemble-averaged 2P-PDF. This relationship will be used for justification of Eq. (26). Again, let us consider a particular separation $\Delta\mathbf{r}$. It is convenient to consider the PDF of moments of the texture ensemble,

$$\rho(\mu_{00}, \mu_{01}, \ldots, \mu_{nn}) = \rho(\{\mu\}), \quad (30)$$

where $\{\mu_{ij}\}$ denotes $\mu_{00}, \mu_{01}, \ldots, \mu_{nn}$, the set of moments. The mean square distance σ^2 between two arbitrary textures a and b in the ensemble can be found as

$$(\sigma')^2 = \sum_{ij}\sum_{kl} [(B^{-1})^T A B^{-1}]_{ijkl} \times \int\int(\mu_{ij}^a - \mu_{ij}^b)(\mu_{kl}^a - \mu_{kl}^b) \times \rho(\{\mu\}_a)\rho(\{\mu\}_b)\mathrm{d}\{\mu\}_a\mathrm{d}\{\mu\}_b, \quad (31)$$

where $\mathrm{d}\{\mu\}_a$ denotes the product $\mathrm{d}\mu_{00}^a\mathrm{d}\mu_{01}^a\ldots\mathrm{d}\mu_{nn}^a$ of differentials. The mean square distance between 2P-PDF's of textures and ensemble-averaged 2P-PDF can be found as

$$\sigma^2 = \sum_{ij}\sum_{kl} [(B^{-1})^T A B^{-1}]_{ijkl}\int(\mu_{ij} - \langle\mu_{ij}\rangle) \times (\mu_{kl} - \langle\mu_{kl}\rangle)\rho(\{\mu\})\mathrm{d}\{\mu\}. \quad (32)$$

Because textures a and b are independent, the integral in Eq. (31) can be evaluated as

$$\int\int(\mu_{ij}^a - \mu_{ij}^b)(\mu_{kl}^a - \mu_{kl}^b)\rho(\{\mu\}_a)\rho(\{\mu\}_b)\mathrm{d}\{\mu\}_a\mathrm{d}\{\mu\}_b$$
$$= \int\int(\mu_{ij}^a\mu_{kl}^a + \mu_{ij}^b\mu_{kl}^b - \mu_{ij}^a\mu_{kl}^b - \mu_{ij}^b\mu_{kl}^a)$$
$$\times\ \rho(\{\mu\}_a)\rho(\{\mu\}_b)\mathrm{d}\{\mu\}_a\mathrm{d}\{\mu\}_b$$
$$= 2(\langle\mu_{ij}\mu_{kl}\rangle - \langle\mu_{ij}\rangle\langle\mu_{kl}\rangle). \qquad (33)$$

The integral in Eq. (32) can be evaluated as

$$\int(\mu_{ij} - \langle\mu_{ij}\rangle)(\mu_{kl} - \langle\mu_{kl}\rangle)\rho(\{\mu\})\mathrm{d}\{\mu\}$$
$$= \int(\mu_{ij}\mu_{kl} + \langle\mu_{ij}\rangle\langle\mu_{kl}\rangle - \mu_{ij}\langle\mu_{kl}\rangle - \mu_{kl}\langle\mu_{ij}\rangle)$$
$$\times\ \rho(\{\mu\})\mathrm{d}\{\mu\} = \langle\mu_{ij}\mu_{kl}\rangle - \langle\mu_{ij}\rangle\langle\mu_{kl}\rangle. \qquad (34)$$

Combining Eqs. (34) and (32) as well as Eqs. (33) and (31) yields the following relation:

$$\sigma' = \sqrt{2}\sigma. \qquad (35)$$

Consequently,

$$\sigma' = \sqrt{2}\left[\sum_{\Delta\mathbf{r}}\sigma^2(\Delta\mathbf{r})\right]^{1/2}$$
$$= \sqrt{2}\left(\sum_{\Delta\mathbf{r}}\mathrm{tr}\{[B^{-1}(\Delta\mathbf{r})]^T[A(\Delta\mathbf{r})]^T B^{-1}(\Delta\mathbf{r})C(\Delta\mathbf{r})\}\right)^{1/2}. \qquad (36)$$

5. APPLICATION TO TEXTURE DISCRIMINATION

The problem of classification can now be more precisely formulated. Suppose we are given two classes A and B and some texture images to classify. The average 2P-PDF's for classes A and B and the distance distribution functions given by Eq. (28) can be estimated. Given a particular texture image X to be classified and the averaged 2P-PDF distributions of classes A and B, the distances

$$d_1 = d(X, A),$$
$$d_2 = d(X, B),$$
$$d_{AB} = d(A, B) \qquad (37)$$

are computed. Parameters d_1 and d_2 are the parameters used in the texture classification. These parameters are not independent: First, they are positive, and second, they satisfy the conditions for a distance measure, given by triangle inequalities

$$d_1 + d_2 \geqslant d_{AB},$$
$$d_1 + d_{AB} \geqslant d_2,$$
$$d_2 + d_{AB} \geqslant d_1. \qquad (38)$$

These inequalities restrict the values of d_1 and d_2. It is possible to use a geometrical representation of this problem. One can assign a pair (d_1, d_2) to a point in the plane, and the triangle inequalities will form a region of possible positions for these points. Figure 6(a) illustrates this region. Lines 1, 2, and 3 are the boundaries imposed on the classification by the system of inequalities given by relations (38). Permitted (d_1, d_2) pairs are located within the dashed area.

Furthermore, let us assume that the distance distributions given by Eq. (28) within each ensemble are either estimated or known. We shall treat the case of a uniform distance distribution as an example. We further assume, without loss of generality for the framework, that the radius of each ensemble, denoted $r_1 = r_a{}^A = r_1{}^B$, is the same for both textures. Let us consider r to be the distance between texture images and the averaged 2P-PDF within a class. The PDF of r is then given by

$$p(r) = \begin{cases} 1/r, & r \leqslant r_1 \\ 0, & r > r_1 \end{cases}. \qquad (39)$$

Figure 7 illustrates an error rate of the classification procedure as a function of r_1/d_{AB}, assuming that the probability distribution function is given by Eq. (39) and $r_1{}^A = r_1{}^B$. To compute the error, we assume that the error

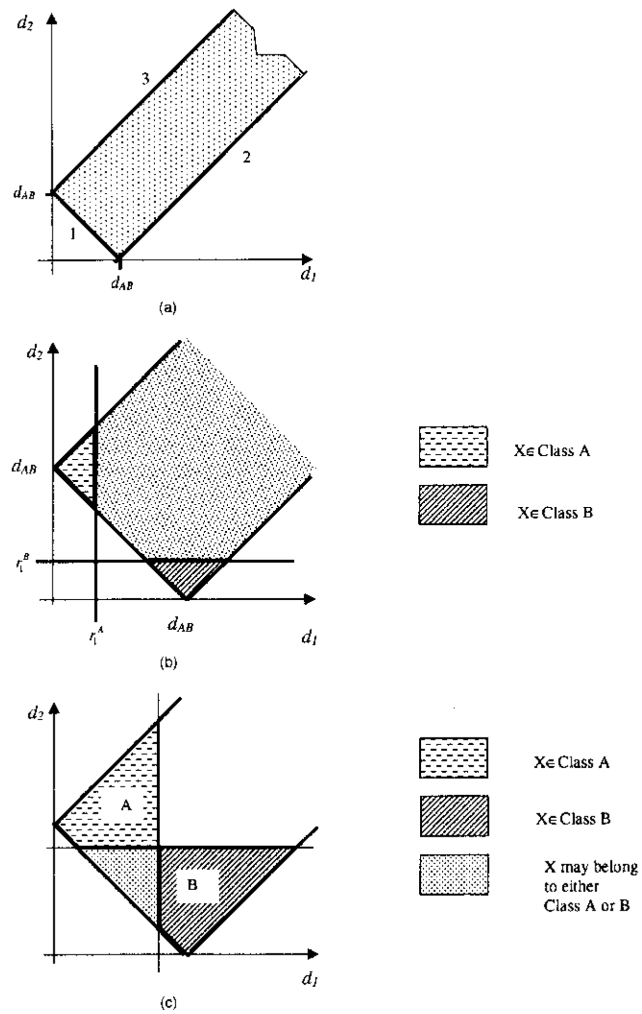

Fig. 6. (a) Space of the discrimination parameters. Points within boundaries 1, 2 and 3 satisfy Eq. (38). (b) Shaded regions show permitted points in space (d_1, d_2). Classes A and B are fully discriminable. $d_{AB} > r_1^A + r_1^B$. (c) Classes A and B have common points and are not fully discriminable. $d_{AB} < r_1^A + r_1^B$.

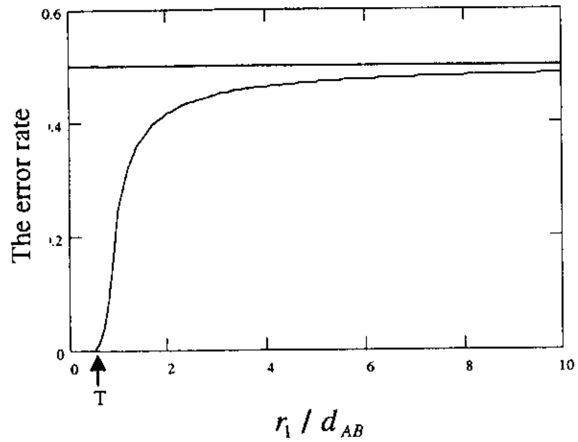

Fig. 7. Error rate of the classifications between two ensembles with uniform distance distribution and the same radii. Point $T(r_1/d_{AB} = 0.5)$ corresponds to the moment when Eq. (1) becomes an equality.

rate is proportional to the area of the overlapping region. Figures 6(b) and 6(c) show configurations of the plane (d_1, d_2) corresponding to various error rates of the classification procedure. Figure 6(b) illustrates that, given the selected values of $r_1{}^A$ and $r_1{}^B$, classes A and B are fully discriminable. Figure 6(c) shows an example of a choice of $r_1{}^A$ and $r_1{}^B$ that yields overlap in the two classes and thus a nonzero error rate.

6. RESULTS AND DISCUSSION

A. Computation of σ' for Four Texture Ensembles—Validation of Equation (36)

We compare the root-mean-square distance within an ensemble of textures calculated by using Eq. (36) with that obtained by averaging the distances between textures within an ensemble of texture realizations. Four texture ensembles were chosen: granite 1, granite 2, grass, and residue, shown in Fig. 8 before equalization of their first-order statistics. The residue image was obtained from a mammographic image.[6] The first-order statistics (i.e., the histogram) of granite 1, granite 2, and grass were equalized to that of the residue image, which was Gaussian. Each ensemble was composed of twelve 256×256 texture images of 256 gray levels. Results of the computation are shown in Table 1. The slope of a regression line through the data is equal to 1.067, validating the expression found by Eq. (36).

B. Application of the Proposed Method to Texture Classification

The same texture ensembles were chosen to test the proposed texture classification method. For classification efficacy the texture classes must have variations between classes greater than the variations within classes.

Table 2 represents the distances calculated with Eq. (3). The ME method was used to compute the 2P-PDF. The pixels' separation were chosen to be $(\Delta \mathbf{r})_x = \{-2, -1, 0, 1, 2\}, (\Delta \mathbf{r})_y = \{-2, -1, 0, 1, 2\}$. To find the 2P-PDF for each pixel separation, twenty-five (5×5) moments should be used as explained in Section 3. We limited ourselves to nine (3×3) moments because of computational speed. The data on the diagonal represent the variation of the distance parameter within a class. The distance between classes is shown in the off-diagonal cells. These results show that the variation of the distance between classes is greater than the variation of the distance within each class.

Although the framework needs to be applied to a larger set of texture ensembles, including texture ensembles with subtle texture differences, these results provide some early experimental support for the mathematical model that we proposed to effectively classify textures on the basis of the full normalized 2P-PDF. Future work may also generalize the proposed framework to include higher-order statistics and test their benefit in solving the problem of texture classification. When the complete normalized second-order statistics are considered as pro-

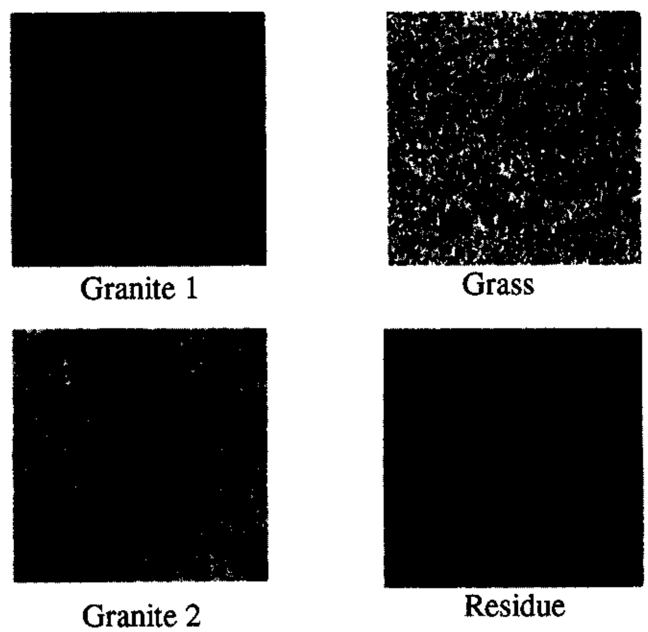

Fig. 8. Some realizations of the four texture ensembles used in the computations before equalization of their first-order statistics.

Table 1. Values of σ' Computed with Either Eq. (36) or Direct Averaging

Ensemble	Eq. (36)	Direct Averaging
Granite 1	0.088	0.123
Granite 2	0.094	0.110
Grass	2.882	3.085
Residue	0.067	0.068

Table 2. Values of Interdistances and Intradistances for Four Textures Classes

Class	Granite 1	Granite 2	Grass	Residue
Granite 1	0.123	1.144	5.025	1.521
Granite 2	1.144	0.110	5.216	1.721
Grass	5.025	5.216	3.085	4.179
Residue	1.521	1.721	4.179	0.068

posed in this paper rather than as a subset of features extracted from nonnormalized second-order statistics, a question for investigation is whether higher-order statistics are required.

7. CONCLUSION

We presented a method of texture classification based on a distance measure between normalized 2P-PDF's. The method proposed here differs from previous approaches in that it considers the complete second-order statistics instead of a subset of features from the second-order statistics. The mathematical framework for computation of this distance as well as the radii of texture classes was presented. Early experimental results were presented in support of the mathematical framework presented.

ACKNOWLEDGMENTS

We thank Harry Barrett and Eric Clarkson of the University of Arizona for stimulating discussions at early stages of this research about the 2P-PDF. This research was supported in part by the University of Central Florida small grant program, the Florida Hospital Gala Endowed Program for Oncologic Research, and the I4-corridor Award program.

The authors can be reached at the addresses on the title page. A. Goon can be reached by telephone at 407-823-6922, by fax at 407-823-6880, or by e-mail at goon@creol.ucf.edu. J. Rolland can be reached by telephone at 407-823-6870, by fax at 407-823-6880, or by e-mail at rolland@creol.ucf.edu.

REFERENCES

1. H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers, "Model observers for assessment of image quality," Proc. Natl. Acad. Sci. USA **90**, 9758–9765 (1993).
2. J. P. Rolland and H. H. Barrett, "Model of background uncertainty and observers detection performance," J. Opt. Soc. Am. A **9**, 649–658 (1992).
3. G. Revesz, H. L. Kundel, and M. A. Graber, "The influence of structured noise on detection of radiological abnormalities," Invest. Radiol. **9**, 479–486 (1974).
4. F. O. Bochud, F. R. Verdun, C. Hessler, and J. F. Valley, "Detectability on radiological images: the influence of anatomical noise," in *Medical Imaging 1995: Image Perception*, H. L. Kundel, ed., Proc. SPIE **2436**, 156–165 (1995).
5. J. W. Byng, M. J. Yaffe, G. A. Lockwood, L. E. Little, D. L. Tritchler, and N. F. Boyd, "Automated analysis of mammographic densities and breast carcinoma risk," Cancer (N.Y.) **80**, 66–74 (1997).
6. J. P. Rolland and R. Strickland, "An approach to the synthesis of biological tissue," Opt. Express **1**, 414–423 (1997).
7. E. B. Gargill, "A mathematical liver model and its application to system optimization and texture analysis," Ph.D dissertation (University of Arizona, Tucson, Ariz. 1989).
8. S. H. L. Hylen, "Image modifiers for use in photography," U.S. patent 5,649,259 (15 July 1997).
9. J. P. Rolland, A. A. Goon, and L. Yu, "Synthesis of textured complex backgrounds," Opt. Eng. **37**, 2055–2063 (1998).
10. J. P. Rolland, A. A. Goon, E. Clarkson, and L. Yu, "Synthesis of biomedical tissue," in *Medical Imaging 1998: Image Perception*, H. L. Kundel, ed., Proc. SPIE **3340**, 85–90 (1998).
11. R. Gonzalez and R. Woods, *Digital Image Processing* (Addison-Wesley, Reading, Mass., 1992), Chap. 8.
12. K. R. Castleman, *Digital Image Processing* (Prentice-Hall, Englewood Cliffs, N.J., 1979), Chap. 16.
13. M. G. A. Thomson and D. H. Foster, "Role of second- and third-order statistics in the discriminability of natural images," J. Opt. Soc. Am. A **14**, 2081–2090 (1997).
14. A. Tremeau, J. Bousigue, and B. Laget, "Co-occurrence shape descriptors applied to texture classification and segmentation," in *Machine Vision Applications in Industrial Inspection IV*, A. R. Rao, N. Chang, eds., Proc. SPIE **2665**, 135–147 (1996).
15. R. S. Mia, M. H. Loew, K. A. Wear, R. F. Wagner, and B. S. Garra, "Quantitative ultrasound tissue characterization using texture and cepstral features," in *Medical Imaging 1998: Image Processing*, K. M. Hanson, ed., Proc. SPIE **3338**, 211–219 (1998).
16. R. M. Haralick, "Statistical and structural approaches to texture," Proc. IEEE **67**(5), 786–804 (1979).
17. A. Papoulis, *Probability, Random Variables, and Statistical Processes* (McGraw-Hill, New York, 1991).
18. A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables* (Marcel Dekker, New York, 1992), Chap. 3.2.