The semi-variogram and spectral distortion measures for image texture retrieval

Tuan Pham

Linköping University Post Print

N.B.: When citing this work, cite the original article.

http://dx.doi.org/10.1109/TIP.2016.2526902

©2016 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
http://ieeexplore.ieee.org/

Postprint available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-124603
The Semi-Variogram and Spectral Distortion Measures for Image Texture Retrieval

Tuan D. Pham, Senior Member, IEEE

Abstract—Semi-variogram estimators and distortion measures of signal spectra are utilized in this paper for image texture retrieval. On the use of the complete Brodatz database, most high retrieval rates are reportedly based on multiple features, and the combinations of multiple algorithms; while the classification using single features is still a challenge to the retrieval of diverse texture images. The semi-variogram, which is theoretically sound and the cornerstone of spatial statistics, has the characteristics shared between true randomness and complete determinism; and therefore can be used as a useful tool for both structural and statistical analysis of texture images. Meanwhile, spectral distortion measures derived from the theory of linear predictive coding provide a rigorously mathematical model for signal-based similarity matching, and have been proven useful for many practical pattern classification systems. Experimental results obtained from testing the proposed approach using the complete Brodatz database, and the UIUC texture database suggest the effectiveness of the proposed approach as a single-feature-based dissimilarity measure for real-time texture retrieval.

Index Terms—Texture analysis, image retrieval, geostatistics, semi-variogram, linear predictive coding, spectral distortion measures.

I. INTRODUCTION

Texture analysis is an important area of research in image processing and pattern recognition. Its applications are pervasively adopted in many areas of medicine [11]-[8], biology [9]-[13], and other interdisciplinary domains of science and engineering [14]-[21]. Texture is conventionally categorized into two major classes: regular and irregular [22]. However, many observations of real-life texture are often characterized with appearances defined on the continuous spectrum from regular (periodic) to irregular (random) patterns [23]. Until now, the notion of texture is rather subjective, because there exists no formal definition for texture [24].

Many techniques have been proposed for texture analysis, where a collection of important developments in this field are reviewed in [25]. Because of the imprecise properties of texture, the task of texture analysis is to mathematically acquire some measures that can distinguish or classify different types of texture. Broadly speaking, algorithms developed for texture analysis fall into two main approaches: structural and statistical [24]. The combination of both structural and statistical methods for texture analysis have also been developed. One of the most basic methods of the structural approach for texture analysis is by Fourier analysis that decomposes and groups the transformed image data into a set of measurements, which can be used for classification. Other commonly used transformed-based methods are the Gabor transform [26], wavelet transform [27], wavelet packet frame [28], and Gabor wavelet [29]. Meanwhile, the most primitive and widely adopted statistical approach for texture classification is the gray-level co-occurrence matrix (GLCM) [30]. Other structural, statistical, fuzzy and combined methods for texture analysis and classification include human visual perception [31], fractal analysis [32], shape grammars [22], Boolean models [22], mathematical morphology [33], and Markov random fields [34], [35], and fuzzy uncertainty texture spectrum [36].

Some of the most recently developed methods for texture analysis and classification include algorithms based on Markov random fields [37], structural characteristics [38], chi-squared transformation of local-binary-pattern feature [39], block truncation coding [40], texton encoding [41], adaptive median binary patterns [42], space-frequency co-occurrences [43], extension of local binary patterns with scale and orientation information [44], improved local binary pattern features [45], [46], and the combination of the Gabor and curvelet filters [47]. Regarding texture analysis methods applied to the database derived from the Brodatz album [48], while many methods have been applied to the small subsets of the Brodatz database[49], only a few attempts reported on texture classification with the complete Brodatz database [50]-[53], particularly for real-time image retrieval with large databases [50]. Although the Brodatz database is short of the presence of intra-class variation, it is the best well-known benchmark database for testing texture analysis methods due to its wide and vague presentation of texture, ranging from being perceptually similar to highly non-uniform, making it challenging to the task of texture classification [51], [49].

This paper presents the combination of the semi-variogram, which is a function in spatial statistics or geostatistics [54], [55], and spectral distortion measures [56] for the retrieval of image texture . The semi-variogram for spectral image classification was investigated in order to compare its performance with the GLCM [57], in which based on training data, the classifier is a distance metric that calculates the sum of the absolute differences between the semi-variograms of two images to determine the class of the test images. Although the notion of the semi-variogram (or variogram) has been applied to texture analysis, mainly in the domain of remote sensing [58], [59]; the application of spectral distortion measures of semi-variograms to the classification of a large set of classes of texture presented in this paper has never been explored before. In fact, the usefulness of spectral distortion measures...
for signal recognition has been well recognized in speech processing research [56], and found in database searching of similar complex biological sequences [60]. The findings of this paper will be discussed to highlight the advantages of the spectral distortion measures of image semi-variograms as effective measures of a single feature for texture analysis, classification, and retrieval indexing.

The rest of this paper is organized as follows. Section II outlines the rationale for the combination of the semi-variogram with spectral distortion measures. Section III presents the semi-variogram for feature extraction of texture images. Section IV describes how the dissimilarity between two semi-variograms can be quantified using spectral distortion measures. Experimental design for testing the proposed approach, results and comparisons with other methods are given in Section V. Finally, Section VI is the conclusion of the research findings, including suggestions for further research in texture analysis.

II. REASONS FOR COMBINING THE SEMI-VARIOGRAHM AND SPECTRAL DISTORTION MEASURES

The semi-variogram, which is the fundamental concept of geostatistics, has been found a very useful tool for characterizing the spatial attributes of natural phenomena, known as regionalized variables. Such spatial natural attributes fit well into those of texture. Particularly, when deterministic models do not exist due to the complexity of the natural process, and knowledge about these attributes are partially gathered at a limited number of samples [61]. Furthermore, in a probabilistic framework, the semi-variogram does not require the knowledge of the mean of the random function for its estimation of the spatial autocorrelation. In addition, the semi-variogram relaxes a rigid assumption for its existence than the covariance. This can be illustrated by the discrete Wiener-Levy process, which is a mathematical description of the so-called Brownian motion—a random process with a semi-variogram and no covariance, implying that the semi-variogram can be defined in some cases, whereas the covariance function cannot be defined [62], [63]. In particular, the semi-variogram may increase with larger lags, corresponding to an infinite global variance [64]. Another advantage is that because the semi-variogram is formulated on the average of the squared differences of the variable, it enables the filtering of the influence of a spatially varying mean, and can be robust to noise.

Given advantages provided by the semi-variogram for dealing with the characterization of spatial attributes, the direct use of the experimental semi-variogram has limitations for the task of texture retrieval addressed in this study. In fact, it is a common practice in geostatistics that the experimental semi-variogram is to be fit into a theoretical semi-variogram, which can then be used to characterize the spatial continuity and structure of a spatial random field for the prediction of values at unsampled locations by kriging [61]. Two approaches are conventionally adopted to fit the experimental semi-variogram into a theoretical model. The first one is by manual fitting, in which a theoretical semi-variogram is selected based on the visual inspection of the experimental semi-variogram [65], [61]. The second approach is to automate the model fitting by using methods such as least squares or maximum likelihood [66]. The manual approach is obviously not suitable for automated texture classification. The accuracy of fitting methods such as maximum likelihood techniques relies on the normality assumption for the data distribution, whereas the generalized least squares method is computationally demanding, and all automated fitting models often encounter with poor fits or no fit at all [66], [67]. Thus, in the context of pattern recognition, there is a need to analyze and compare the dissimilarity of experimental semi-variograms of various image samples accurately and quickly for real-time application. One popular approach for analysis and matching of such kind of physical signals is the linear predictive coding (LPC) [68], [69] and its associated distortion measures [56]. The second-order stationary property of the semi-variogram makes it most suitable for the use of LPC that derives a small number of prediction parameters known as LPC coefficients to capture a compact and precise representation of spectral magnitude for the signals [69]. These LPC coefficients can be used as templates for pattern recognition. In addition, LPC has relatively simple computation and is of model-based analysis that provides more accurate spectral resolution than the non-parametric Fourier transform when the signal is stationary for a short time [69].

A spectral distortion is a measure of dissimilarity between two spectra of two digital signals, and such signals in this study are the semi-variograms of images. Spectral distortion measures can be formulated directly on LPC coefficients or LPC-derived cepstral coefficients [56]. As it well known in speech recognition, quantifying the dissimilarity between two patterns in terms of spectral distortion appears to be a very effective method for pattern comparison, both in terms of its mathematical tractability and its computational efficiency [56]. In particular, being analogous to the classification of speech patterns [56], since perceived texture differences can be modeled in terms of differences of spectral features, the measure of spectral distortion can be both mathematically and subjectively rational.

III. TEXTURE ANALYSIS WITH THE SEMI-VARIOGRAHM

Let $Z$, $x$, and $h$ be a random function, a spatial location, and a lag distance in the sampling space, respectively. The random function $Z(x)$ is assumed to be second-order stationary over the sampling domain, which implies $E[Z(x)] = m$, where $E[\cdot]$ denotes expected value, and $m$ is a constant mean. The variogram of the random function is defined as [61]

$$2\gamma(h) = \text{Var}[Z(x) - Z(x + h)],$$

(1)

where $\gamma(h)$ is called the semi-variogram of the random function.

Let $Z(x_i), i = 1, 2, \ldots, n$, be a sampling of size $n$, the unbiased estimator for the semi-variogram, which is called the experimental semi-variogram, of the random function is expressed as
whose elements are defined as distance \( h \). In this study, the semi-variogram of an image at a lag \( h \) is calculated by taking the sum of squared differences of the intensity values of pixel pairs separated by \( h \) in both rows and columns, then divided by the total number of the pixel pairs. An example for calculating the semi-variogram of an image of 5x5 pixels in typical directions is given in [57].

The estimator defined in Equation (2) is not robust with respect to outliers or severe skewness [70]. This drawback leads to the development of robust semi-variograms [71]. The unbiased estimator of the robust semi-variogram, denoted as \( \gamma_R(h) \), is defined as [72], [66]

\[
\hat{\gamma}_R(h) = \delta \left[ \frac{1}{N(h)} \sum_{i=1}^{N(h)} |Z(x_i) - Z(x_i + h)|^{1/2} \right]^4,
\]

where
\[
\delta = \frac{1}{2 \left( 0.457 + \frac{0.494}{N(h)} \right)}.
\]

IV. MATCHING SEMI-VARIOGRAMS WITH SPECTRAL DISTORTION MEASURES

The formulations of spectral distortion measures are based on the notion of linear predictive coding (LPC), which optimally transforms the spectral envelope of a digital signal into a compressed form (coding) based on a linear combination of previous samples [73]-[74]. The LPC model for the semi-variogram attempts to approximate \( \hat{\gamma}(h) \) as a linear combination of the past \( p \) semi-variances, such that

\[
\hat{\gamma}(h) = \sum_{i=1}^{p} a_i \hat{\gamma}(h - i)
\]

where \( a_i, i = 1, \ldots, p \) are the LPC coefficients, and can be determined as

\[
a = -R^{-1} r
\]

where \( a \) is a \( p \times 1 \) vector of the LPC coefficients, \( R \) is a \( p \times p \) autocorrelation matrix, and \( r \) is a \( p \times 1 \) autocorrelation vector whose elements are defined as

\[
r_i = \sum_{h=0}^{N} \hat{\gamma}(h)\hat{\gamma}(h + i), \; i = 1, \ldots, p.
\]

Let \( S(\omega) \) and \( S'(\omega) \) be the spectral density functions of a \( p \)-th order all-pole model of the semi-variograms \( \hat{\gamma}(h) \) and \( \hat{\gamma}'(h) \), respectively, where \( \omega \) is normalized frequency ranging from \(-\pi\) to \(\pi\). The spectral density \( S(\omega) \) is defined as [56]

\[
S(\omega) = \frac{\sigma^2}{|A|^2},
\]

where \( \sigma^2 = a^T Ra \), and \( A = 1 + a_1 e^{-i\omega} + \ldots + a_p e^{-ip\omega} \).

The log spectral distance between two spectra \( S(\omega) \) and \( S'(\omega) \) on a frequency scale, denoted as \( d_{LS}(S, S') \) is defined as [56], [75]

\[
d_{LS}(S, S') = \log S(\omega) - \log S'(\omega).
\]

The log-likelihood-ratio (LLR) distortion between \( S(\omega) \) and \( S'(\omega) \), denoted as \( d_{LLR}(S, S') \), is defined as [75]

\[
d_{LLR}(S, S') = \frac{a^T Ra'}{a^T Ra},
\]

which is also called the Itakura distortion measure [76], and \( a' \) is the vector of the LPC coefficients of \( S' \).

The likelihood-ratio distortion measure between \( S(\omega) \) and \( S'(\omega) \), denoted as \( d_{LR}(S, S') \), is defined as [75]

\[
d_{LR}(S, S') = \frac{a^T Ra'}{a^T Ra} - 1.
\]

The weighted-cepstral distortion measure between \( S(\omega) \) and \( S'(\omega) \), denoted as \( d_{WC}(S, S') \), is defined as [56]

\[
d_{WC}(S, S') = \sum_{i=1}^{L} w_i(c_i - c'_i)^2,
\]

where \( w(n) \) is any suitable lifter function, such as the following used in this study [56]:

\[
w_i = \left\{ \begin{array}{ll}
1 + g \sin \left( \frac{\pi n}{N} \right) & : i = 1, \ldots, L \\
0 & : i \leq 0, i > L,
\end{array} \right.
\]

in which \( g \) is usually chosen as \( L/2 \), \( L \) is the truncated term and taken as \( p - 1 \) in this study, where \( p \) is the LPC number of poles, and the cepstral coefficients \( c_i \) is derived with the following recursion [56]:

\[
c_i = -a_i - \frac{1}{i} \sum_{k=1}^{i-1} k \; c_k \; a_{i-k}, \; i > 0,
\]

where \( a_i, i = 1, \ldots, p \), are the LPC coefficients, \( a_0 = 1 \), \( a_k = 1 \) for \( k > p \), \( c_0 = \log \sigma^2 \), and \( c_{-1} = c_i \).

V. EXPERIMENTAL DESIGN AND RESULTS

A. Retrieval Design

The proposed approach was tested using the whole original Brodatz database, which consists of 112 classes of natural texture images, with one image for each class [48]. This image database is available for public access as documented in [77]. Although the Brodatz database does not demonstrate a high intra-class variation, it is widely recognized as a challenging benchmark for evaluating texture analysis methods, particularly the classification of the whole database, which consists of a large number of classes (112 textures) [49], [51]. Each image of the Brodatz database is of size 640 \times 640 pixels. Because of only one image representing each class, each image was split into 9 non-overlapping sub-images (except 5 columns and 5 rows of pixels for the corresponding sub-images in the last row-wise and column-wise partitions), yielding each sub-image of size 215 \times 215 pixels. Thus, the dataset results in 1008 generated sub-images for 112 classes, 9 samples for each
The sub-images will be subsequently referred to as images.

To obtain results comparable with those given by other existing methods, the testing of the proposed approach was carried as follows. A retrieval was performed for each of 1008 images of the database. For each input query image \( q \), and given a specified number of searched images, the retrieval rate is calculated as

\[
R_q(|I|) = \frac{|A| - |B|}{|A|} \times 100, \quad q = 1, \ldots, Q, \tag{15}
\]

where \( R_q(|I|) \) is the percentage rate of correct retrieval for query image \( q \), \( I \) is the set of retrieved images that are closest to query image \( q \), with a specified cardinality \( |I| \), \( Q \) is the total number of images contained in the database, \( A \) is the set of images that are of the same class of the input image, and \( B \subseteq I \) is the set of correctly retrieved images, \( |A| \) and \( |B| \) are the cardinalities of \( A \) and \( B \), respectively.

For the testing of this generated Brodatz image database, \( Q=1008 \), and \( |A| = 8 \). The total average retrieval rate, denoted as \( R(|I|) \), is therefore

\[
R(|I|) = \frac{1}{Q} \sum_{q=1}^{Q} R_q(|I|). \tag{16}
\]

The experimental and robust semi-variograms are considered isotropic when they depend only on the lag \( h \). The semi-variogram is anisotropic when it varies in different directions. Thus, the semi-variogram should be constructed for different orientations given the configuration of the data. Having mentioned earlier, in this study, \( h \) is taken in both horizontal and vertical directions of each image as these two orientations have been reported to work reasonably well for image analysis [6], [78], [79]. The semi-variograms were standardized using the z-score to enhance the comparison of the spectral distortion measures. Thus, a standardized vector of \( \tilde{\gamma}(h) \), \( h = 1, \ldots, 30 \), was used to calculate the corresponding vector of LPC coefficients with \( p = 20 \).

### B. Retrieval Results

Table I shows the retrieval rates using the complete Brodatz database obtained from the four spectral distortion measures \( d_{LS}, d_{LLR}, d_{LR}, d_{WC} \), and the Euclidean distance of the experimental \( \gamma(h) \) and robust \( \tilde{\gamma}(h) \) semi-variograms, with \( |I| = 8, 10, 20, 30, 40, \) and 50 images. In this experiment, the use of the spectral distortion measures of the experimental and robust semi-variograms produced the same results. Table I also shows the retrieval rates using the Euclidean distance measure of the experimental \( \gamma(h) \), and robust \( \tilde{\gamma}(h) \) semi-variograms. By definition, the semi-variogram is a function of a variable \( h \), called the lag. A semi-variogram that increases in dissimilarity with distance over short lags and then levels off is called transitive semi-variogram. The lag at which the experimental semi-variogram approximately reaches a constant value is called the range, and the value of the experimental semi-variogram at or beyond the range is referred to as the sill. For the type of data in this study, the robust semi-variogram reduces the skewness of the semi-variograms of the texture images by scaling the corresponding sills to certain values, while maintaining the shapes of the experimental semi-variograms (see Figure 1, which shows texture D5, its experimental and robust semi-variograms). Both semi-variances produced by the two semi-variograms therefore would not make marked differences in the LPC analysis and the LPC-based distortion measures. This can be a reason that the spectral distortion measures of either the experimental or robust semi-variogram of the texture images result in the same retrieval rates.

The retrieval of 8 similar images to each query image of the total database performed by the log-likelihood-ratio distortion measure \( d_{LLR} \) of the experimental semi-variograms, which achieves the highest retrieval rates, resulted in 12 images (D20, D21, D23, D26, D40, D54, D56, D65, D79, D86, D94, and D95) that have the average retrieval rates of 100% (total success), and 1 image (D51) that has the average retrieval rates of 0% (total failure).

The time required for calculating \( \tilde{\gamma}(h) \), with 30 lags, for an image of 215 × 215 pixels is 0.070 seconds on a laptop computer ProBook 6570b, Core i7 (CPU@2.90GHz), Windows 8. Using the same computer, the total retrieval times for all 1008 images as well as each image taken for the Euclidean distance of the experimental semi-variogram \( (d_{LS}(h)) \), and the four spectral distortion measures \( (d_{LS}, d_{LLR}, d_{LR}, d_{WC}) \) are given in Table II. For both times required for retrieving all 8 most similar images and the most similar image among the total of 1007 images, the log-spectral distortion took the longest time (847 seconds for 8 images and 806 seconds for a single image), while the Euclidean distance required the least amount of time (93 seconds for 8 images and 82 seconds for a single image).

Moreover, a question of interest in assessing the performance of the proposed approach is as follows. Assuming that there are \( n \) images with \( m \) different types of textures, with \( n > m \). Would the proposed approach be able to find all different types of textures \( m \) in \( n \) images? A typical answer to this question can be sought by carrying out the retrieval task on a dataset consisting of 10 textures (\( m = 10 \)) of the Brodatz database: D4, D9, D19, D21, D24, D28, D29, D36, D37, and D38. Having mentioned earlier, each original Brodatz image was split into 9 non-overlapping sub-images, making \( n = 90 \) (9 samples for each of the 10 textures). The confusion matrix for the texture retrieval, using \( D_{LLR} \), is given in Table IV, which shows the abilities of the proposed approach in identifying all 10 different types of texture, either completely (D19, D21, D28, and D29) or partially (D4, D9, D24, D36, D37, and D38). Although the shapes of the experimental semi-variograms of D4, D9, D24, D28 and D29 look similar; the sills and ranges of these experimental functions are different and exhibit visual appearances of the spherical (theoretical) semi-variogram function [61], allowing the extraction of their LPC coefficients more discriminative to be implemented with the spectral distortion measure for pattern matching. As a result of the discriminative power, the retrieval rates of these textures are among the highest: 100% for D28 and D29, and 93% for D4 and D9. D38 has the lowest retrieval rate (36%).
which was mostly confused with D28, D9, and D19. D37 is the texture that has the second lowest retrieval rate (69%). The misclassified samples of D37 were assigned to D28. The most likely reason for resulting in low retrieval rates for D38 and D37 is due the wide bandwidth of their semi-variograms as shown in Figure 2.

C. Retrieval Comparisons

To compare the performance of the proposed spectral distortion measures with other methods tested against a smaller set of images of the Brodatz database, the log-spectral ($d_{LS}$), log-likelihood-ratio ($d_{LLR}$), and likelihood-ratio ($d_{LR}$) were tested on the same set of 40 images studied in [80]. The indices of these 40 images are: D6, D9, D11, D14, D16, D17, D20-D22, D24, D26, D34, D36, D41, D46, D47, D51, D53, D55-D57, D64, D66, D68, D76-D80, D82-D84, D101-D106.
Fig. 2. Subimages of 10 Brodatz textures (left) and their experimental semi-variograms (right): from left to right–1st row: D4 and D9; 2nd row: D19 and D21; 3rd row: D24 and D28; 4th row: D29 and D36; and 5th row: D37 and D38.

TABLE IV
CONFUSION MATRIX OF RETRIEVAL RATES (%) FOR 10 BRODATZ TEXTURES USING $D_{L,L,R}$.

<table>
<thead>
<tr>
<th></th>
<th>D19</th>
<th>D21</th>
<th>D24</th>
<th>D28</th>
<th>D29</th>
<th>D36</th>
<th>D37</th>
<th>D38</th>
<th>D4</th>
<th>D9</th>
</tr>
</thead>
<tbody>
<tr>
<td>D19</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D21</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D24</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>D28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D36</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>84</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D37</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>69</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D38</td>
<td>14</td>
<td>0</td>
<td>6</td>
<td>19</td>
<td>8</td>
<td>0</td>
<td>36</td>
<td>0</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>D4</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D9</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>93</td>
</tr>
</tbody>
</table>


Figure 4. Comparisons of decrease in accuracy with increasing level of white Gaussian noise using UIUC texture database.

D109, and D111. The results obtained from the 3 spectral distortion measures of the experimental semi-variograms, with the number of lag = 35 lags and $p = 25$, are given in Table III.

Furthermore, the comparisons of single-feature-based retrieval rates obtained from the proposed approach (without training data) and other three methods (with training data), all tested against the complete Brodatz database, with the same experimental setups, are given in Table I. The other three methods are the multiscale simultaneous autoregressive (SAR) method with level 1 [50], tree-structured wavelet transform (TWT) [81] as reported in [52], and Tamura features (coarseness, contrast, directionality) [82] as also reported in [52]. These features have still been used and cited for comparing the performances of features for image retrieval and annotation [83]-[85]. Both log-likelihood-ratio and likelihood-ratio distortion measures consistently achieved the highest retrieval rates. The results obtained from the log-spectral distortion measure is slightly lower than the log-likelihood-ratio and likelihood-ratio distortion measures. The retrieval rates obtained from the Euclidean distance measure of the experimental and robust semi-variograms are similar and higher than the weighted-cepstral distortion measure. The SAR with a single level and TWT performed equivalently. The Tamura features yielded the lowest retrieval rates across different numbers of retrieved images.

Regarding the results of the test using the set of 40 images of the Brodatz database as shown in Table III, the log-spectral, log-likelihood-ratio and likelihood-ratio distortion measures (without training data) outperformed other single-feature-based methods reported in [80]; tree-structured wavelet transform (TSWT) [81], pyramid-structured wavelet transform (PSWT) [86], Gabor transform [87], and the combination of the Gabor filters and probabilities computed from the angular second moment, contrast, and correlation of the gray-level co-occurrence matrix (GLCM) [88]; three of which used training data for the retrieval task. In this experiment, the log-spectral distortion measure achieved the highest results, following by the other two spectral distortion measures.

The results of some other methods, which combined several algorithms or and multiple features for the retrieval using the whole Brodatz database, are also mentioned here to relatively highlight the performance of the spectral distortion measures of the image semi-variograms. The results reported in [51] were obtained from an approach requiring the use of two region detectors for the extraction of affine regions of the images, shape normalization involving two rotation-invariant descriptors [89], [90], cluster analysis, and the earth mover’s distance [91]. In [51], the retrieval of 8 similar images using the combination of the Harris descriptor (H), Laplacian descriptor (L), together with the ellipse channels (E): (H+L) E; and (H+L) (S+R+E), where S and R stand for the SIFT descriptor [90] and RIFT descriptor [51], respectively; are 44.59% and 76.26%, respectively. The performance of the random features extracted from sparse representation and compressive sensing [49], carried out on the complete Brodatz database, is slightly lower than that of the sparse texture using local affine regions [51].

The principal component analysis and SAR model [50], where the latter is equivalent to the SAR model developed in [92], with the Mahalanobis distance were applied to the retrieval of the whole Brodatz database [50], in which the retrieval rates of 8 nearest patterns obtained from the use of 99 feature variances and 99 eigen-filter-based features, using a training set of 100 images, are about 46%, and 50%, respectively. The results obtained from the log-spectral, log-likelihood-ratio and likelihood-ratio distortion measures of the experimental semi-variograms as a single feature, without training data, are competitive among these two multiple-features-based retrieval methods.

Having compared the performance of the proposed method, namely SV to stand for the semi-variogram based likelihood-ratio distortion measure, with other published results using single-feature-based methods, given that those obtained from the SV are without training data; the proposed method was also compared with other texture analysis methods in terms of robustness to noise. The other methods include the local binary pattern (LBP), robust (Weber) local image descriptor (WLD), and scale invariant feature transform (SIFT), all reported in [93], in which only 32 Brodatz textures and training data were used; while the proposed method was applied to retrieving the whole database with 112 textures and without training data. The retrieval accuracy is evaluated in the presence of various levels of white Gaussian noise added to the images. The decrease in accuracy of the SV, LBP, WLD, and SIFT versus increasing white Gaussian noise levels are shown in Figure 3. The noise level is the logarithm of the inverse of the signal-to-noise ratio of the power of the reference and degraded images. The proposed method (SV) consistently performs better than the other three methods, particularly at higher noise levels. For the noise level of 0.2, the retrieval accuracies of the LBP, WLD, and SIFT are dropped to between 60%-75%, whereas the accuracy of the SV is dropped to below 40%.

To compare further the performance of the proposed method
in the presence of noise, the texture retrieval was carried out using the UIUC texture database [51], which can be downloaded at http://www-cvr.ai.uiuc.edu/ponce_grp/data/. This texture database has 25 texture classes, there are 40 samples for each class, making up the total of 1000 images, and the size of each image is of $640 \times 480$ pixels. The results obtained from the proposed method are compared with the LBP, completed modeling of the LBP (CLBP), and SIFT reported in [94], in terms of accuracy with the presence of increasing levels of white Gaussian noise. Figure 4 shows the plot of the decrease in accuracy of the SV, LBP, CLBP, and SIFT versus increasing white Gaussian noise levels. The consistently high level of performance of the SV is once again demonstrated in this experiment. For the noise level of 0.2, the retrieval accuracies of the LBP, CLBP, and SIFT are dropped to about 55%, whereas the accuracy of the SV is decreased to about 35%.

Table V shows the comparison of computational times for extracting various features for a single image with different sizes, using the LBP, fuzzy local binary pattern (FLBP), local ternary pattern (LTP) (computational times for LBP, FLBP, and LTP are reported in [95], Table 3, using Intel Dual Core T2410 with CPU@2GHz), and SV methods (using Core i7, CPU@2.90GHz). These results highly demonstrate that the time taken for the extraction of texture feature provided by the SV is much shorter than the other three feature extraction methods.

VI. CONCLUSION

The combination of the semi-variogram developed in geostatistics, and the spectral distortion measures developed in speech recognition are very promising for texture analysis as well as the retrieval of a large number of image texture classes. The computer implementation of the proposed approach is practically simple and its computational speed is cost-effective, suitable for real-time database retrieval (Matlab codes for the methods studied in this paper are available at https://sites.google.com/site/professortuanpham/codes).

Having illustrated the effectiveness of the proposed method, there are limitations to the discriminative power of the semi-variogram and distortion measures, which can be overcome by other methods. For example, in the test of 10 Brodatz images, the retrieval rate obtained from the proposed method for D38 is low (36%). As another example, for testing the whole Brodatz database, the retrieval rates for D48 and D87 provided by the proposed method are 64% and 90%, respectively; while those for D48 and D87 by the approach reported in [51] are 97% and 100%, respectively. Therefore, it would be useful to combine texture features of complementary strengths to improve the retrieval.

Having mentioned earlier, textural structures are classified according to their surface characters [23]: regular, near-regular, irregular, and stochastic. This kind of classification is inherently subject to inconsistent interpretation. Therefore appropriate theoretical models of the SV can be used to fit to different experimental SVs to categorize texture in an objective way. While there are many suitable functions defined for the theoretical SV, typical theoretical models of the SV are the spherical, exponential, Gaussian, power, and sine hole effect [61], which are shown in Figure 5 and worth exploring for an objective quantification of texture perception.

On the use of spectral distortion measures for computing pairwise SV dissimilarity, these measures are built on LPC that has some certain advantages in providing a good approximation of short time series such as the semi-variograms: being an analytically tractable model, and being simple and straightforward in both software and hardware implementation. Finally, quantifying the difference between SV functions of an image pair in terms of average spectral distortion appears to be a good tool for matching patterns, both in terms of its mathematical tractability and its computational efficiency.

ACKNOWLEDGMENT

A major part of this work was carried out while the author was with the Medical Image Processing Lab at the University of Aizu, Japan.

REFERENCES
