



HAL
open science

Texture Analysis and Classification With Linear Regression Model Based on Wavelet Transform

Zhi-Zhong Wang, Jun-Hai Yong

► **To cite this version:**

Zhi-Zhong Wang, Jun-Hai Yong. Texture Analysis and Classification With Linear Regression Model Based on Wavelet Transform. IEEE Transactions on Image Processing, Oct 2008, California, United States. inria-00517305

HAL Id: inria-00517305

<https://inria.hal.science/inria-00517305>

Submitted on 14 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Texture Analysis and Classification With Linear Regression Model Based on Wavelet Transform

Zhi-Zhong Wang and Jun-Hai Yong

Abstract—The wavelet transform as an important multiresolution analysis tool has already been commonly applied to texture analysis and classification. Nevertheless, it ignores the structural information while capturing the spectral information of the texture image at different scales. In this paper, we propose a texture analysis and classification approach with the linear regression model based on the wavelet transform. This method is motivated by the observation that there exists a distinctive correlation between the sample images, belonging to the same kind of texture, at different frequency regions obtained by 2-D wavelet packet transform. Experimentally, it was observed that this correlation varies from texture to texture. The linear regression model is employed to analyze this correlation and extract texture features that characterize the samples. Therefore, our method considers not only the frequency regions but also the correlation between these regions. In contrast, the pyramid-structured wavelet transform (PSWT) and the tree-structured wavelet transform (TSWT) do not consider the correlation between different frequency regions. Experiments show that our method significantly improves the texture classification rate in comparison with the multiresolution methods, including PSWT, TSWT, the Gabor transform, and some recently proposed methods derived from these.

Index Terms—Linear regression, texture analysis, texture classification, wavelet transform.

I. INTRODUCTION

TEXTURE provides essential information for many image classification tasks. Extensive research has been done on texture classification during the last three decades [1]–[4]. In the 1980s, most traditional approaches included gray level co-occurrence matrices (GLCM) [5], second-order statistic method [6], Gauss–Markov random field [7], and local linear transform [8], which are restricted to the analysis of spatial relations between neighborhood pixels in a small image region. As a consequence, their performance is the best for the class of so called microtextures [9]. With a study on human vision system, researchers begin to develop the multiresolution texture analysis

Manuscript received June 19, 2006; revised April 1, 2008. First published June 27, 2008; last published July 11, 2008 (projected). This work was supported in part by the Chinese 973 Program (2004CB719400) and in part by the National Science Foundation of China (60533070, 60625202). J.-H. Yong was supported in part by the project sponsored by the Foundation for the Author of National Excellent Doctoral Dissertation of China (200342), in part by the Fok Ying Tung Education Foundation (111070), and in part by the Program for New Century Excellent Talents in University (NCET-04-0088). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Srdjan Stankovic.

Z.-Z. Wang is with the Technique Research Department, China Center For Resource Satellite Data and Application, Yongfeng Industry Base, Haidian District, Beijing 100094, China (e-mail: wangzz04@126.com).

J.-H. Yong is with the School of Software, Tsinghua University, Beijing 100084, China (e-mail: yongjh@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TIP.2008.926150

models, such as the wavelet transform [9]–[13] and the Gabor transform [14]–[16]. Extensive research has demonstrated that these approaches based on the multiresolution analysis could achieve reasonably good performance [2], [4], so they have already been widely applied to texture analysis and classification. Recently, a large number of approaches have been proposed. In [17], a stochastic model called random set is employed to texture classification. In [18], the structural properties of the texture can be analyzed by a set of structuring windows defined by a neighborhood system. In [19], six logical operators from Boolean analysis are convolved with texture image to obtain texture characteristics. In [20], a texture image is transformed into a pattern map for characterizing texture information by a set of templates with principal component analysis (PCA). In [21], the Fisher criterion is employed to optimize FIR filters for texture feature extraction. In [22], the mutual information of different subbands after the multichannel decomposition is used for the sparse representation of texture image. Except for these above techniques, some papers focus on the development of combining the multiresolution analysis with the statistical methods, such as the spectral histogram [23] and GLCM [24], [25].

The most common multiresolution analysis approach is to transform a texture image into a local spatial/frequency representation by convolving this image with a bank of filters with some tuned parameters. This property coincides with the study that the human visual system can be modeled as a set of channels. This clearly motivates researchers to study how to extract more discriminable texture feature based on the multiresolution techniques. Currently, the wavelet transform and the Gabor transform are the most popular multiresolution methods. Compared to the wavelet transform, the Gabor transform needs to select the filter parameters according to different texture. There is a trade-off between redundancy and completeness in the design of the Gabor filters because of nonorthogonality. The Gabor transform is also limited to its filtering area. Consequently, we choose the wavelet transform to obtain the spectral information of the texture image. While the traditional multiresolution methods keep an eye on the spectral information of the texture image at different scales and use the statistics of the spectral information as the texture descriptor, they ignore the texture structural information. To make use of the texture primitives, GLCM as commonly popular structural descriptor is recently combined with the traditional multiresolution methods [24], [25]. Nevertheless, they simply bring two kinds of the texture features together so as to result in the united features with high dimension.

In this paper, it is found that there exists a distinctive correlation between some sample texture images, belonging to the same kind of texture, at different frequency regions obtained

by 2-D wavelet packet transform. Experimentally it is demonstrated that this correlation varies from texture to texture. A new texture classification method, in which the simple linear regression model is employed into analyzing this correlation, is presented. Experiments show that this method significantly improves the texture classification rate in comparison with the multiresolution methods, including PSWT, TSWT, the Gabor transform, and some recently proposed methods derived from these.

This paper is organized as follows. The brief review about 2-D wavelet transform and an application of the correlation between different frequency regions to texture analysis are described in Section II. Section III describes the texture classification algorithm with respect to the correlation. In Section IV, some experimental results are presented. Several traditional multiresolution techniques, such as PSWT, TSWT, and the Gabor transform, some recently proposed methods derived from these, and our method are compared. Finally, the conclusions are summarized in Section V.

II. TEXTURE ANALYSIS WITH LINEAR REGRESSION MODEL

A. Two-Dimensional Wavelet Packet Transform

The wavelet transform provides a precise and unifying framework for the analysis and characterization of a signal at different scales [26]. It is described as a multiresolution analysis tool for the finite energy function $f(x) \in L_2$ [27], [28]. It can be implemented efficiently with the pyramid-structured wavelet transform and the wavelet packet transform. The pyramid-structured wavelet performs further decomposition of a signal only in the low frequency regions. Adversely, the wavelet packet transform decomposes a signal in all low and high frequency regions. As the extension of the 1-D wavelet transform, the 2-D wavelet transform can be carried out by the tensor product of two 1-D wavelet base functions along the horizontal and vertical directions, and the corresponding filters can be expressed as $h_{LL}(k, l) = h(k)h(l)$, $h_{LH}(k, l) = h(k)h(l)$, $h_{HL}(k, l) = g(k)h(l)$, and $h_{HH}(k, l) = g(k)g(l)$. An image can be decomposed into four subimages by convolving the image with these filters. These four subimages characterize the frequency information of the image in the LL, LH, HL, and HH frequency regions, respectively. The 2-D PSWT can be constructed by the whole process of repeating decomposition in the LL regions, whereas the 2-D wavelet packet transform decomposes all frequency regions to achieve a full decomposition, as shown in Fig. 1. Therefore, the 2-D PSWT depicts the characteristics of the image in the LL regions and the 2-D wavelet packet transform describes the properties of the image in all regions.

Most of the research in the multiresolution analysis based on the wavelet domain focuses on directly extracting the energy values from the subimages and uses them to characterize the texture image. The energy distribution of a subimage can be calculated by one of the three commonly used functions: the magnitude $|\cdot|$, the squaring $(\cdot)^2$, and the rectified sigmoid $|\tanh(\alpha \cdot)|$ [2]. The magnitude and squaring functions are similar in the effect of the nonlinearity. The rectified sigmoid function requires an appropriate saturation parameter α . The mean and the standard deviation of the magnitude of the subimage coefficients

LL ₃	HL ₃	HL ₂	HL ₁
LH ₃	HH ₃		
LH ₂		HH ₂	
LH ₁			HH ₁

(a)

LL ₃	HL ₃						
LH ₃	HH ₃						
LL ₃	HL ₃						
LH ₃	HH ₃						
LL ₃	HL ₃						
LH ₃	HH ₃						
LL ₃	HL ₃						
LH ₃	HH ₃						

(b)

Fig. 1. (a) Three level 2-D PSWT decomposition of 128×128 image. (b) Three level 2-D wavelet packet decomposition of 128×128 image.

also can be calculated as texture feature in [14], [16]. In this paper, the mean of the magnitude of the subimage coefficients is used as its energy. That is, if the subimage is $x(m, n)$, with $1 \leq m \leq M$ and $1 \leq n \leq N$, its energy can be represented as

$$e = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x(i, j)| \quad (1)$$

where $x(i, j)$ is the pixel value of the subimage.

B. Analysis of Correlation Between Frequency Channels

The wavelet (packet) transform approximately decorrelates the image using the orthogonal bases and can be viewed as an approximation of the Karhunen–Loève transform (KLT)[29], [30]. In this paper, the correlation does not refer to the linear correlation between different frequency regions of a texture image after decomposition by 2-D wavelet transform, but instead indicates the spatial correlation between some sample texture images, belonging to the same kind of texture, at different frequency regions obtained by 2-D wavelet transform. Now we give the following description to illuminate this statement. Given that there are some sample texture images from the same kind of texture, these images should have the same spatial relation between neighborhood pixels as this texture. These images are all decomposed to obtain the same frequency regions by 2-D wavelet transform. The most common approach is to calculate all frequency regions' energy values of every image with the energy function and to characterize this texture

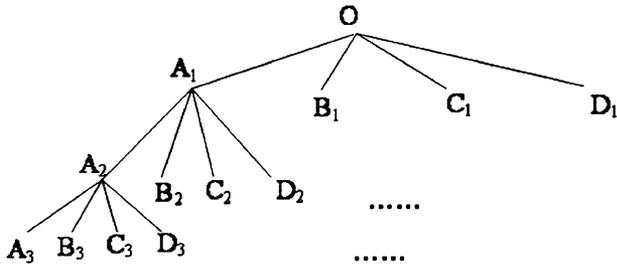


Fig. 2. Tree representation of Fig. 1(b).

by the statistics of these energy values. This approach ignores the spatial relation of these sample texture images. In this study, we capture this inherent texture property by learning a number of sample texture images. From a statistical perspective, a frequency region of a sample texture image can be viewed as a random variable and the energy values of this frequency region can be treated as the random values of this variable. Experimentally it is found that there exists a distinctive correlation between different variables (frequency regions).

As a powerful multiresolution analysis tool, the 2-D wavelet transform has proved useful in texture analysis, classification, and segmentation. The 2-D PSWT performs further decomposition of a texture image only in the low frequency regions. Consequently, it is not suitable for images whose dominant frequency information is located in the middle or high frequency regions. Although the 2-D wavelet packet transform characterizes the information in all frequency regions, this representation is redundant. On the other hand, in order to generate a sparse representation, the 2-D TSWT decomposes the image dynamically in the low and high frequency regions whose energies are higher than a predetermined threshold. However, it ignores the correlation between different frequency regions. Other approaches use information theoretic concepts of the entropy [31] and the mutual information [22], [29]. However, they only select the frequency region energy for sparse representation of the texture image and do not consider the correlation between different frequency regions which is a very distinctive characteristic of the texture image. In this subsection, we will elaborate on this correlation and present a preprocessing algorithm that determines the frequency channel pairs that have significant correlation.

Because the 2-D wavelet packet transform describes much more spectral information than the 2-D PSWT and is suitable for the texture modeled as the quasi-periodic signals [12], we use 2-D wavelet packet transform to obtain all frequency information of a texture image in this paper. The detail is described as follow. First, the original image is decomposed into four subimages, which can be viewed as the parent node and the four children nodes in a tree and named as O, A, B, C, and D, respectively, as shown in Fig. 2. They symbolize the original image, LL, LH, HL, and HH frequency regions. Second, we calculate the energy of these subimages by using the (1) and the tree forms four branches from the parent node. Third, we repeat the decomposition of these subimages and the tree branches at the power of four until satisfying the least size of the subimage. Practically, the least size of the subimage depends on the requirement, but it should not be less than 16×16 according to experience in

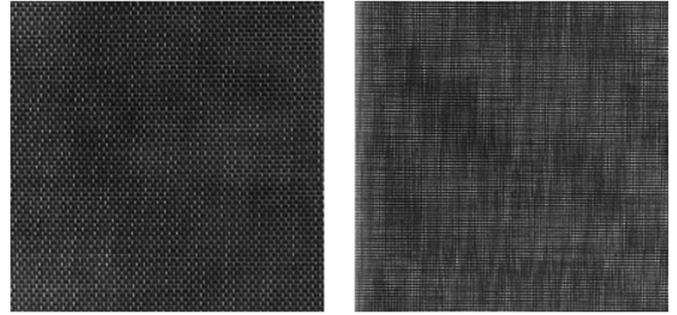


Fig. 3. Texture d6 and d21.

consideration of robustness. Finally, we get k frequency regions ($k = 4^n$, where n is the number of decomposition levels). Fig. 2 shows an example of Fig. 1(b). The frequency channel is defined as each branch of the tree. For example, the channel of the left branch in the tree is OAAA in Fig. 2. For all $k = 4^n$ channels, we record the energy of the subimage at the leaf of every branch as the energy of every channel and make these energy values to compose a vector of length k , called the channel-energy vector V . It represents k frequency channels of an image.

After decomposing all j samples from the same texture with 2-D wavelet packet transform, we get j channel-energy vectors of length k and use them to form a matrix with j rows and k columns, called the channel-energy matrix $M_{j \times k}$. It represents k frequency channels of j sample images from this texture, where m_{ij} indicates the energy of the j th frequency channel of the i th sample, and then we figure out the covariance matrix $C_{k \times k}$ with k rows and k columns from $M_{j \times k}$, where c_{ij} is the correlation coefficient ρ between the i th frequency channel and the j th frequency channel. The frequency channel pair denotes two frequency channels with a certain correlation coefficient ρ . The correlation coefficient ρ describes how much correlation two frequency channels have. If ρ is smaller than the threshold, their correlation is considered not to be remarkable. The appropriate decision of the threshold will be discussed in the last two paragraphs of this subsection. Therefore, we only consider those channel pairs whose correlation ρ are enough large as the top channel pairs. It is assumed that different textures should have different top channel pairs. This assumption can be tested by the following experiment.

We acquire 81 sample images of size 128×128 with an overlap of 32 pixels between vertically and horizontally adjacent images from d6 texture image (woven aluminum wire) and d21 texture image (french canvas) of size 640×640 with 256 gray levels, respectively. They are very homogeneous as shown in the Fig. 3. Next, two 64×64 covariance matrices C_{d6} and C_{d21} can be obtained as what is said above. As ρ descends, the top ten channel pairs can be extracted from C_{d6} and C_{d21} , respectively. Table I shows the comparison between two groups of the top ten frequency channel pairs. It can be observed that the correlation between different frequency channels certainly exists and that the top ten channel pairs of d6 texture and d21 texture are different though two textures are very similar. This experiment implies that the correlation between frequency channels is a distinctive characteristic of the texture and can

TABLE I
TOP TEN CHANNEL PAIRS COMPARISON BETWEEN TEXTURE D6 (WOVEN ALUMINUM WIRE) AND TEXTURE D21 (FRENCH CANVAS)

No.	texture ID	channel pair		correlation (ρ)	texture ID	channel pair		correlation (ρ)
1	d6	OACC	OCCA	0.983	d21	OCAA	OCBB	0.938
2	d6	OAAD	OACB	0.983	d21	OCCA	ODDB	0.929
3	d6	OAAC	OACA	0.982	d21	OCCA	OCDB	0.924
4	d6	OAAC	OCCA	0.979	d21	OCAC	OCBD	0.916
5	d6	OACA	OACC	0.979	d21	OCAA	ODBB	0.914
6	d6	OACA	OCCA	0.973	d21	OCCA	OCCB	0.908
7	d6	OAAC	OACC	0.968	d21	OACA	OADB	0.897
8	d6	OACC	OCAC	0.965	d21	OCDB	ODDB	0.896
9	d6	OABC	OADA	0.963	d21	OCCC	OCDD	0.882
10	d6	OBBA	OBBA	0.96	d21	OCAC	OCAD	0.879

serve as a good candidate for texture representation to detect different texture.

The preprocessing algorithm is needed to output distinctive top channel pairs of a texture for texture classification. The algorithm is given as follows.

Algorithm 1: The Preprocessing Algorithm

[Input:] all j samples of a given texture

[Output:] the channel-pair list and the channel-energy matrix M

- 1) Decompose a sample of a given texture with 2-D wavelet packet transform into an output of k frequency channels.
 - 2) Calculate the energy of k channels with the (1), and get a channel-energy vector V of length k .
 - 3) Repeat the first and second steps for all j sample images of this texture and then take j channel-energy vectors to construct a $j \times k$ channel-energy matrix M .
 - 4) Figure out the $k \times k$ covariance matrix C from M .
 - 5) Select the top channel pairs with the correlation coefficient $\rho \geq T$ and order them into a list as ρ descends, and, finally, output this list and the channel-energy matrix M .
-

In the fifth step of *Algorithm 1*, the constant T is a controllable parameter which servers as a threshold for selecting the top channel pairs with enough correlation. According to the statistics [32], T is the critical value that separates the acceptance and rejection regions of the correlation. The correct decision of T is related to the significance level α and the sample number of texture. The threshold T can be computed from the following equation:

$$T_{\alpha}(n-2) = \sqrt{\frac{F_{\alpha}(1, n-2)}{(n-2) + F_{\alpha}(1, n-2)}} \quad (2)$$

where α is the significance level, n is the sample size, and $F(n_1, n_2)$ is F distribution with n_1 and n_2 degrees of freedom. The symbol α means the risk that the frequency pair with $\rho < T$ is excluded when it has high correlation. From the hypothesis testing perspective, this risk is also called the *Type I error*. It can

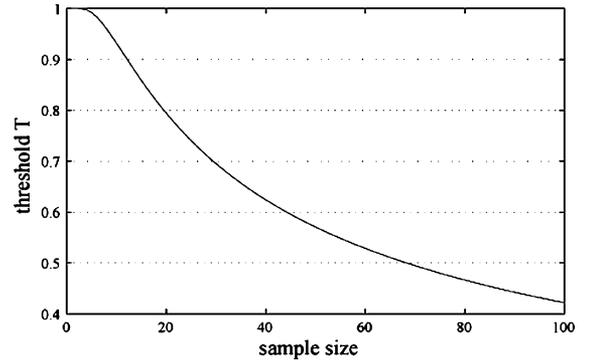


Fig. 4. Value of T as a function of the sample number with the significance level $\alpha = 0.00001$.

be lessened by diminishing the value of α . In addition, there is the *Type II error* which is error of including the frequency pair with low correlation though $\rho \geq T$. The goodness of the preprocessing algorithm greatly is measured by the size of two error rates.

Now we give a pair of hypotheses: the null hypothesis, denoted by H_0 , implying that all frequency channel pairs with $\rho < T$ do not have the remarkable correlation, and the alternative hypothesis, denoted by H_1 , being a contradiction of the null hypothesis. To avoid making the *Type I error* in H_0 , the value of T is not set too large. However, we probably commit the second error in this case. Compared to the *Type I error*, the *Type II error* is not unfortunately easy to be controlled and depends in a large part on an large sample size. Considering the computation cost caused by enlarging sample size, we must choose an appropriate sample size. With regard to the task of seeking the right top channel pairs as far as possible, the second error is more important. Consequently, the rule is provided that given an extremely small α , the value of T is selected by the use of the (2) with an economical sample size. From Fig. 4, it is shown that the value of T descends gradually with the rise of the sample size.

C. Linear Regression Model

In the above subsection, the correlation between different frequency regions has been validated as a sort of effective texture characteristic. In this subsection, we employ the simple linear regression model to analyze the correlation. Suppose that we have a set of the random data

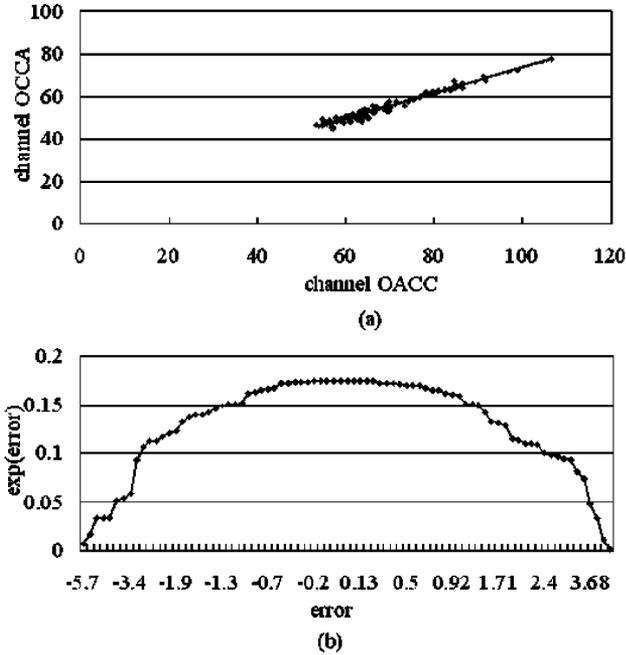


Fig. 5. (a) Channel energy distribution. (b) Residual distribution of the channel pair of OACC and OCCA with the correlation $\rho = 0.983$ belonging to texture d6.

$(x_1, y_1)^T, (x_2, y_2)^T, \dots, (x_n, y_n)^T$ for two numerical variables X and Y . Further, suppose that we regard X as a cause of Y . From the simple linear regression analysis [33], the distribution of the random data approximately appears a straight line in X, Y space when X and Y are perfectly related linearly. There is a linear function that captures the systematic relationship between two variables. This line function (also called the simple linear regression equation) can be given as follows:

$$\hat{y} = a \times x + b \quad (3)$$

where \hat{y} is called a fitted value of y .

We exploit the simple linear regression model to extract the texture features from the correlation in the frequency channel pairs. The channel-pair list includes all channel pairs with $\rho \geq T$. For two frequency channels of one channel pair in the list, we take out their energy values from the channel-energy matrix M and then consider these energy values as the random data $(x_1, y_1)^T, (x_2, y_2)^T, \dots, (x_n, y_n)^T$ for two variables X and Y (X and Y being two frequency channels of this channel pair). The distribution of these energy values should also similarly represent a straight line in X, Y space. The parameters a and b of the line can be figured out through the least square method. The expressions are written as follows:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (4)$$

$$b = \frac{1}{n} \sum_{i=1}^n y_i - \frac{a}{n} \sum_{i=1}^n x_i. \quad (5)$$

Fig. 5(a) shows the distribution of the energy of the channel pair of OACC and OCCA with the correlation $\rho = 0.983$ belonging to texture d6 shown in Table I. It appears approximately

as a black line whose parameters $a = 0.604$ and $b = 13.24$ by using the (4) and (5). Thus, the simple linear regression (3) describes the linear relationship of two frequency channels of each top channel pair. Obviously, when we take x of the random data with a high correlation into the regression (3) to obtain \hat{y} , there exists the residual between y and \hat{y} through the expression $|y - \hat{y}|$. According to the statistics, the exponential function values of such residuals exhibit a normal distribution curve. This phenomenon has also occurred in our experiment, as shown in Fig. 5(b).

The statistics implies that a normally distributed (or Gaussian) random variable X has probability density function $p(x) = (1)/(\sigma\sqrt{2\pi}) \exp[-((x - \mu)^2)/(2\sigma^2)]$ where the parameters (μ and σ) of the distribution are the mean and variance of X , respectively. When the variable X is in the area of $\mu \pm 3\sigma$, its probability is 99.72%. In the other words, the probability of a sample image satisfying the correlation of one top channel pair of a texture is highly 99.72% when its residual at this channel pair is in the area of $\mu \pm 3\sigma$. From the above discussion, it can be known that a texture always have many top channel pairs and the combination of these channel pairs is the characteristic of this texture. If a sample image meets the correlation of all top channel pairs of a texture, it can be inferred that it belongs to this texture. Therefore, the value of $\mu \pm 3\sigma$ can be considered as an appropriate way to assign an unknown texture image to a texture. The estimation of mean μ and variance σ can be carried out by the (6) and (7)

$$\mu = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (6)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - 2}}. \quad (7)$$

In general, the wavelet packet transform generates a multiresolution texture representation including all frequency channels of a texture image with the complete and orthogonal wavelet basis functions which have a “reasonably well controlled” [31] spatial/frequency localization property. Our method absorbs this advantage of the wavelet packet transform. It is worthwhile to comment on the difference between our method and other multiresolution-based texture analysis methods. PSWT loses the middle and high frequency information. TSWT does not take into account the correlation between different frequency channels. The Gabor transform uses a fixed number of filter masks with predetermined frequency and bandwidth parameters. In contrast, our method not only thinks about all frequency channels but also analyzes the correlation between them with the simple linear regression model. So, it can be viewed as a very useful multiresolution method.

III. TEXTURE CLASSIFICATION

A. Learning Phase

In Section II, *Algorithm 1* outputs the channel-pair list and the channel-energy matrix M of a texture. The learning algorithm follows directly from *Algorithm 1* and extracts the texture features. The detail is as follows.

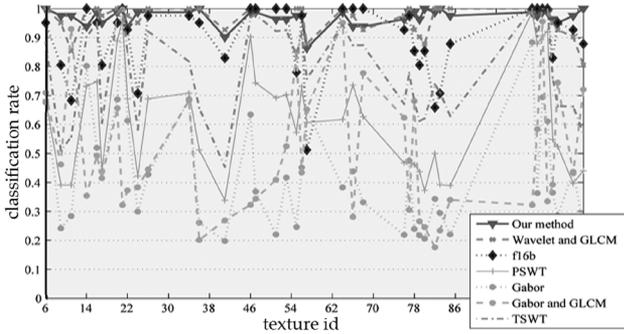


Fig. 6. Comparison of classification rate using different methods.

Algorithm 2: The Learning Algorithm

[Input:] all samples of all textures

[Output:] the feature lists of all textures

- 1) Given all samples obtained from the same texture, get the channel-pair list and the channel-energy matrix M of this texture by the preprocessing algorithm.
 - 2) With the channel-pair list and the channel-energy matrix M , figure out the parameters a, b, μ , and σ of each top channel pair by using the (4)–(7).
 - 3) Consider the parameters a, b, μ , and σ , two frequency channels and the correlation coefficient ρ of each top channel pair and the index of this texture as a feature of this texture. Put all such features into a list as ρ descends, called the feature list, and then insert this list into the database.
 - 4) Repeat three above steps for all textures.
-

Note that the feature lists of the textures are needed to store into the database in the third step. In the feature lists, every texture feature contains the parameters a, b, μ , and σ , two frequency channels and the correlation coefficient ρ of one channel pair and the index of a texture. The parameters a, b , and two frequency channels are used to compute the residual of an unknown texture image at a top channel pair of a texture, and the parameters μ and σ are used to get the threshold $\mu \pm 3\sigma$ in order to determine whether this image satisfies the correlation at this top channel pair of this texture as discussed in Section III.

B. Classification Phase

In the classification phase, we first perform a full decomposition to an unknown texture image with 2-D wavelet packet transform and obtain the channel-energy vector V of this image, and then we process this vector with the features in the feature lists from the database. The detail is as follows.

Algorithm 3: The Classification Algorithm

[Input:] an unknown texture image and the feature lists in the database

[Output:] the index of texture to which this unknown texture image is assigned

- 1) Decompose the unknown texture image with 2-D wavelet packet transform to obtain its channel-energy vector V .
 - 2) Order all textures in the database into a candidate list and perform the following iteration (set $k = 1$ at first).
 - a) Pick out the parameters a, b, μ , and σ of the k th channel pair of a texture in the candidate list from its feature list, and select the energy of two frequency channels identical to that of this top channel pair from the channel-energy vector V of this unknown texture image.
 - b) Take one energy of two channels as x_i into the (3) and get the residual by $|\hat{y}_i - y_i|$, where y_i is the other energy.
 - c) Remove the texture from the candidate list if the residual is larger than $\mu \pm 3\sigma$.
 - d) If there is only one texture is left in the candidate list, assign the unknown texture image to this texture. Otherwise, perform the next iteration by increasing the value of k by one.
-

In step 2d), we discard the very unlikely texture if the residual is larger than $\mu \pm 3\sigma$. In fact, it is found that $\mu \simeq 0$ in our experiment because the distribution of the residual is approximately the normal distribution. Since all features of every texture have already been ordered as the correlation coefficient ρ descends, the whole process of the iteration practices orderly according to the correlation of all top channel pairs of a texture. Although the feature number of a texture feature list may be large, the number of iteration is actually much small because our features are very distinctive and we use the leave-one-out method. This can be verified in the next section. Moreover, our method takes simply the threshold comparison in 1-D space in the classification phase owing to taking advantage of the texture inherent correlation characteristic in the learning phase.

IV. EXPERIMENTAL RESULT

In this section, we will verify the performance of the classification algorithm discussed in Section III. We used 40 textures, as shown in Fig. 9, obtained from the Brodatz's texture album [34]. Every original image is of size 640×640 pixels with 256 gray levels. 81 sample images of size 128×128 with an overlap of 32 pixels between vertically and horizontally adjacent images are extracted from each original image and used in the experiments, and the mean of every image is removed before the processing. These 3240 texture images are separated to two sets being used as the training set with 1600 images and the test set with 1640 images, respectively.

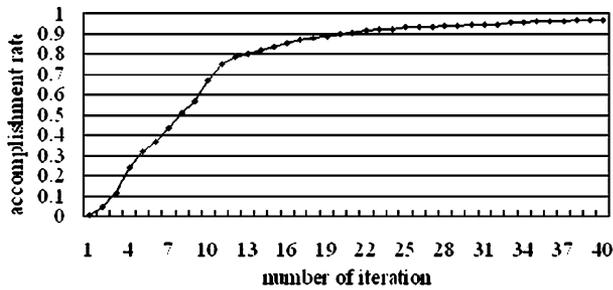


Fig. 7. Retrieval efficiency of our method.

We compare our method with other traditional multiresolution methods, such as the pyramid-structured wavelet transform (PSWT) [26], the tree-structured wavelet transform (TSWT) [12], and the Gabor transform [16]. In the experiments, the feature of PSWT is extracted from the energy of all four frequency regions in the third scale, three middle and high frequency regions in the second scale and three middle and high frequency regions in the first scale. In the classification algorithm of TSWT, its feature is constructed by the energy of all frequency regions in the third scale and its controllable parameters, like decomposition constant and comparison constant, are empirically set to 0.15 and 10, respectively. We implement the Gabor transform by convolving an image with a set of functions obtained by appropriate dilations and rotations of the Gabor mother function $g(x, y) = ((1)/(2\pi\sigma_x\sigma_y)) \exp\{-1/(2)[(x^2)/(\sigma_x^2) + (y^2)/(\sigma_y^2) + 2\pi jWx]\}$, where $W, \sigma_x,$ and σ_y are the predetermined parameters. There are totally 24 filters, so 24 subimages can be obtained in four scales and six orientations, and then we use the mean and standard deviation of the magnitude of the transform coefficients from 24 subimages to represent the feature vector whose length is 48. To make use of the texture primitives, many approaches have been recently proposed to combine the above traditional multiresolution methods with the statistics-based methods, like the fuse of the Gabor transform and GLCM [24] and the combination of the wavelet transform and GLCM [25]. We calculate three common descriptors specified in properties from GLCM, such as angular second moment descriptor for homogeneity, Contrast descriptor for local variation, and Correlation descriptor for linear dependency. Furthermore, Randen [2] provided a comparative study to numerous filtering methods used for texture classification and concluded that f16b filter bank have the overall best performance. F16b filter bank is quadrature mirror bank designed by Johnston [35] and was used for texture analysis by Randen [36]. Therefore, we will also evaluate them.

The average retrieval rate defined in our method is different from that of other methods because our classifier is threshold-based comparison in one dimension. All query samples are processed by our method and respectively assigned to the corresponding texture. We count the samples of this texture, which are assigned to the right texture, and get the average percentage number as the average retrieval rate of this texture. In view of other methods, the distance $d(i, j)$, where i is the query sample and j is a texture from the database, is firstly computed with the feature vectors discussed above. The distances are then sorted

in increasing order and the closet set of patterns are then retrieved. In the ideal case all the top 41 retrievals are from the same texture. The sample retrieval rate is defined as the average percentage number of samples belonging to the same texture as the query sample in the top 41 matches. The average retrieval rate of a texture is the average number of all sample retrieval rates of this texture. Several different distance similarity measures, like Euclidean, Bayesian, Mahalanobis, can be explored to obtain different results for other methods. The different distances are applied to different methods in our experiment on behalf of their reasonably good performances.

Table II shows the retrieval accuracy of these different multiresolution methods for 40 textures. Fig. 6 gives a graph illustrating the retrieval performance as a function of texture ID. Noted that the Gabor transform gets the worst result in exception of some textures, such as D14, D21, and D101. Because Gabor filters use a fixed number of filter masks with predefined frequencies and bandwidths, these parameters probably are not adaptive for all textures, and it may not have sufficient resolution in the frequency regions for some textures. The fuse of Gabor and GLCM behaves slightly better than Gabor, 48.995%. Although GLCM can help Gabor greatly enhance the achievement in some textures, like D20, D34, D102, the performance of this approach is still limited to the Gabor transform. PSWT behaves a bit better than two above methods, but it suffers from the loss of spectral resolution in high frequency regions. TSWT is not very disappointing and slightly exceeds our method for some textures, like D101, D103 and D104. We are very amazed with the outstanding results of the combination of the wavelet transform with GLCM and f16b filter bank. Their correct rates come to highly 96.707% and 90.061%, respectively. Especially, after the wavelet transform and GLCM are used together they exhibit the classification rate greatly close to our method. It is shown that the wavelet transform is superior to the Gabor transform in the way capturing the frequency information of the texture image and the texture spatial relation is effective characteristic. However, the average accuracy of our method is still first class, highly 97.151%, and much greater than other methods, 79.166% (TSWT), 61.588% (PSWT), and 43.429% (the Gabor transform), respectively. In summary:

- TSWT keeps much more information of the spectral resolution in high frequency regions, so it can get better performance than the Gabor transform and PSWT. It is worthwhile to notice that f16b filter bank shows excellent performance compared to TSWT, PSWT, and Gabor. This filter bank is, like the class of the wavelet filter, a very broad class of filters whose parameters are much more suitable for textures in our database. However, our method considers not only all the frequency information of the texture but also the correlation between them. Therefore, it can achieve the best result in comparison with TSWT, PSWT, Gabor, and f16b.
- The way of combining the multiresolution methods, like the wavelet transform and the Gabor transform, and the statistic-based methods, like GLCM, can achieve the superior performance in comparison with the individual multiresolution methods because they take the texture primitives into account. In particular, the combination of the

TABLE II
COMPARISON OF THE EXPERIMENTAL RESULTS USING DIFFERENT METHODS

ID	Our method	F16b	Wavelet and GLCM	TSWT	PSWT	Gabor	Gabor and GLCM
D6	100%	95.122%	100%	88.225%	68.596%	67.639%	70.941%
D9	97.531%	80.488%	95.122%	50.941%	39.136%	24.151%	24.583%
D11	97.531%	68.298%	85.366%	55.756%	39.074%	28.395 %	30.54%
D14	93.827%	100%	100%	90.154%	73.318%	80.201%	74.429%
D16	98.765%	95.122%	100%	95.278%	74.799%	51.914%	46.173%
D17	95.062%	80.488%	97.651%	60.509%	44.398%	43.719%	38.333 %
D20	98.765%	95.122%	100%	98.241%	88.812%	68.673%	95.139 %
D21	100%	100%	100%	100%	96.343%	99.753%	92.963%
D22	93.827%	92.683%	97.561%	84.969%	68.889%	61.312%	58.349 %
D24	98.765%	70.732%	95.122%	58.241%	42.207%	29.907%	33.981%
D26	98.765%	97.561%	100%	92.114%	68.858%	42.762%	68.102%
D34	98.765%	97.561%	100%	81.728%	70.833%	68.688%	85.201 %
D36	100%	95.122%	100%	65.679%	51.25%	26.096%	23.38 %
D41	90.123%	82.927%	92.683%	45.324%	33.827%	19.738%	17.593%
D46	98.765%	100%	100%	96.96%	90.278%	63.457%	72.052%
D47	98.765%	100%	100%	97.886%	74.367%	36.852%	65.602%
D51	96.296%	100%	100%	92.207%	69.228%	22.022%	33.472 %
D53	96.296%	100%	100%	92.577%	70.247%	41.682%	49.259 %
D55	97.531%	78.049%	100%	83.704%	57.238%	24.63%	32.269%
D56	97.531%	97.561%	100%	91.574%	73.133%	45.139%	43.41%
D57	86.42%	51.22%	87.802%	75.725%	60.694%	65.216%	65.17%
D64	98.765%	100%	100%	94.383%	61.713%	38.287%	43.148%
D66	93.827%	100%	97.561%	87.315%	73.58%	43.796%	39.259%
D68	93.827%	100%	92.683%	87.361%	62.685%	33.148%	26.62%
D76	96.296%	92.683%	97.561%	67.022%	46.713%	21.867%	41.466%
D77	98.765%	97.561%	100%	77.824%	47.824%	47.515%	35.401%
D78	98.765%	85.366%	92.683%	67.963%	46.142%	23.935%	32.176%
D79	96.296%	80.488%	90.244%	61.188%	43.688%	21.713%	28.056%
D80	100%	85.366%	87.805%	62.114%	37.253%	20.617%	20.123%
D82	98.765%	65.854%	100%	73.904%	50%	34.306%	52.469%
D83	100%	70.732%	100%	71.019%	39.182%	29.414%	44.506%
D85	97.531%	87.805%	100%	62.901%	38.92%	22.052%	26.929%
D101	98.765%	100%	100%	100%	38.904%	88.364%	96.571%
D102	97.531%	100%	100%	90.478%	87.809%	36.296%	77.716%
D103	98.765%	100%	100%	99.907%	90.571%	69.398%	62.299%
D104	98.765%	100%	100%	99.846%	92.114%	61.142%	68.318%
D105	98.296%	82.927%	95.122%	76.049%	54.815%	36.451%	37.284%
D106	95.062%	95.122%	92.683%	66.343%	52.346%	28.488%	34.336%
D109	97.531%	92.683%	90.244%	66.235%	39.676%	43.657%	40.957%
D111	100%	87.805%	80.488%	56.991%	44.059%	24.753%	32.222%
average	97.151%	90.061%	96.707%	79.166%	61.588%	43.429%	48.995%

wavelet transform and GLCM excels our method in many textures though its average correct rate is still slightly lower than mine. This dramatic trait can be illuminated that the wavelet transform captures the frequency information of the texture image at different scales and GLCM characterizes the local spatial properties of the texture. However, this approach leads to increase the feature space dimension and brings in the computation complexity in classification scheme, even the curse of dimensionality. Feature reduction like principle component analysis (PCA) alleviates the curse to some extent at the cost of affecting the performance [24].

- Our method is to take advantage of the texture inherent correlation characteristic in the learning phase. Thus, our classification algorithm simply takes the threshold comparison

in one dimension space in order to avoid the difficulty in choosing a distance function that is suitable for the distributions of the multidimensional texture features [37]. Furthermore, it is obvious that the simple threshold comparison in one dimension space greatly diminishes the great computation complexity in comparison with the distance measure similarities in high-dimensional space.

- Although the texture features with correlation characteristic may be very large, it is not necessary to compare the sample with all feature of a texture. Because we develop a progressive classification algorithm (leave-one-out) for applying distinctive features to reduce the number of iteration. It is shown in Fig. 7 that more than 90% of the query samples accomplish the retrieval at the 21st iteration in the classification phase.

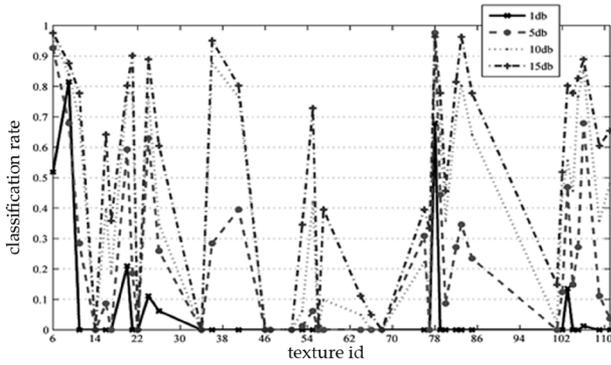


Fig. 8. Correct classification rate with gaussian noisy data.

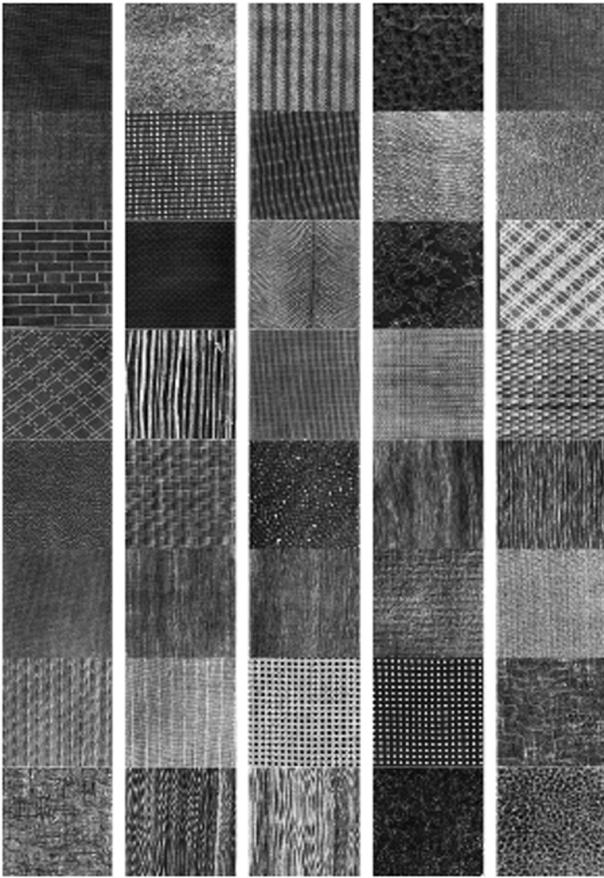


Fig. 9. Forty textures used in the experiments.

The sensitivity of our method to noisy data is tested by adding Gaussian noise to the sample images before the classification. The result is shown in Fig. 8. Although our method still perform well in some textures, like D6, D36, D78, and D83, the classification rate is extremely affected by the noise. It progressively sinks to 52.315%, 41.049%, 22.253%, and 6.36% when the SNR is 15, 10, 5, and 1 dB, respectively. Note that the performance of our method is deteriorating with the decreasing SNR. To explain this phenomenon, we recall that the statistical method plays an important role in the learning phase of our classification algorithm and the noise with a smaller SNR extremely influence the pixel gray value in the texture image. Another reason is that the

energy of Gaussian noise spreading over the entire spectrum affects the frequency channels with small energy. Consequently, they result in the sensitivity of the features to the noise. Furthermore, the threshold rule in the classification phase has also been greatly affected by the noise.

V. CONCLUSION

In this paper, a new approach to texture analysis and classification with the simple linear regression model based on the wavelet transform is presented and its good performance on the classification accuracy is demonstrated in the experiments. Although the traditional multiresolution methods, like PSWT, TSWT, the Gabor transform and f16b filter bank, are suitable for some textures, our method is natural and effective for much more textures. While the fuse of the Gabor transform and GLCM is confined to the Gabor filter, the way of combining the wavelet transform with GLCM obtains the excellent performance very close to our method. However, our method surpasses them in the classification scheme because it adopts the simple threshold comparison and the leave-one-out algorithm. It is worthwhile to point out several distinctive characteristics of this new method. To begin with, our method provides all frequent channels of the quasi-periodicity texture in comparison with PSWT, Gabor, and f16b. So, it is able to characterize much more spectral information of the texture at different multiresolution. Next, the correlation of different frequent channels can be applied in this method through the simple linear regression model. In contrast, TSWT does not consider the correlation between different frequent regions though it also keeps more frequent information and employs the energy criterion to get over the overcomplete decomposition. Therefore, this texture intrinsic characteristic helps our method go beyond TSWT. Furthermore, most of the research in the multiresolution analysis directly computes the energy values from the subimages and extracts the features to characterize the texture image at the multidimension space, and, yet, our method employs the correlation between different frequency regions to construct the texture feature. So, it employs the threshold comparison in one dimension space rather than some distance measures in the multidimension space. Therefore, it is very easy and fast to examine the change of different frequent channels for texture image.

Our current work has focused so far on the simple linear regression model which is used to employ the linear correlation. More thorough theoretical analysis of the linear correlation is expected in the future. In addition, the application of this method to texture segmentation is under our current investigation. Due to our method being quite sensitive to noise shown from Section IV, it is the urgent requirement of achieving its noisy invariance.

ACKNOWLEDGMENT

The authors would like to thank the editor and the anonymous reviewers for their contributing comments, as well as Prof. M. S. Landy, New York University, and T. Randen, Schlumberger Geco-Prakla (Norway), for providing the software package.

REFERENCES

- [1] Y. Rui, T. S. Huang, and S. F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *J. Vis. Commun. Image Represent.*, vol. 10, pp. 39–62, 1999.
- [2] R. Randen and J. H. Husøy, "Filtering for texture classification: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 4, pp. 291–310, Apr. 1999.
- [3] J. Zhang and T. Tan, "Brief review of invariant texture analysis methods," *Pattern Recognit.*, vol. 35, pp. 735–747, 2002.
- [4] C.-C. Chen and C.-C. Chen, "Filtering methods for texture discrimination," *Pattern Recognit. Lett.*, vol. 20, pp. 783–790, 1999.
- [5] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-6, no. 6, pp. 610–621, Nov. 1973.
- [6] P. C. Chen and T. Pavlidis, "Segmentation by texture using correlation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 5, no. 1, pp. 64–69, Jan. 1983.
- [7] R. L. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of images," *IEEE Trans. Inf. Theory*, vol. 29, no. 1, pp. 60–72, Jan. 1983.
- [8] M. Unser, "Local linear transforms for texture measurements," *Signal Process.*, vol. 11, pp. 61–79, 1986.
- [9] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Process.*, vol. 4, no. 11, pp. 1549–1560, Nov. 1995.
- [10] T. Chang and C.-C. J. Kuo, "A wavelet transform approach to texture analysis," in *Proc. IEEE ICASSP*, Mar. 1992, vol. 4, no. 23–26, pp. 661–664.
- [11] T. Chang and C.-C. J. Kuo, "Tree-structured wavelet transform for textured image segmentation," *Proc. SPIE*, vol. 1770, pp. 394–405, 1992.
- [12] T. Chang and C.-C. J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Process.*, vol. 2, no. 4, pp. 429–441, Oct. 1993.
- [13] G. H. Wu, Y. J. Zhang, and X. G. Lin, "Wavelet transform-based texture classification with feature weighting," *Proc. IEEE Int. Conf. Image Processing*, vol. 4, no. 10, pp. 435–439, Oct. 1999.
- [14] W. Y. Ma and B. S. Manjunath, "A comparison of wavelet transform features for texture image annotation," *Proc. IEEE Int. Conf. Image Processing*, vol. 2, no. 23–26, pp. 256–259, Oct. 1995.
- [15] W. Y. Ma and B. S. Manjunath, "Texture features and learning similarity," *Proc. IEEE CVPR*, no. 18–20, pp. 425–430, Jun. 1996.
- [16] B. S. Manjunath and W. Y. Ma, "Texture feature for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.
- [17] I. Epifanio and G. Ayala, "A random set view of texture classification," *IEEE Trans. Image Process.*, vol. 11, no. 8, pp. 859–867, Aug. 2002.
- [18] A. Çarkacıoğlu and F. Yarman-Vural, "SASI: A generic texture descriptor for image retrieval," *Pattern Recognit.*, vol. 36, pp. 2615–2633, 2003.
- [19] V. Manian, R. Vásquez, and P. Katiyar, "Texture classification using logical operation," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1693–1703, Oct. 2000.
- [20] X.-Y. Zeng, Y.-W. Chen, Z. Nakao, and H. Lu, "Texture representation based on pattern map," *Signal Process.*, vol. 84, pp. 589–599, 2004.
- [21] T. Randen and J. H. Husøy, "Texture segmentation using filters with optimized energy separation," *IEEE Trans. Image Process.*, vol. 8, no. 4, pp. 571–582, Apr. 1999.
- [22] K. Huang and S. Aviyente, "Information-theoretic wavelet packet sub-band selection for texture classification," *Signal Process.*, vol. 86, pp. 1410–1420, 2006.
- [23] X. Liu and D. Wang, "Texture classification using spectral histograms," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 661–670, Jun. 2003.
- [24] D. A. Clausi and H. Deng, "Desing-based texture feature fusion using gabor filters and co-occurrence probabilities," *IEEE Trans. Image Process.*, vol. 14, no. 7, pp. 925–936, Jul. 2005.
- [25] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Process.*, vol. 8, no. 4, pp. 592–598, Apr. 1999.
- [26] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. Beijing, China: China Machine Press, 2003, pp. 220–374.
- [27] I. Daubechies, "Philadelphia, pennsylvania: Society for industrial and applied mathematics," *Tex Lectures on Wavelets*, pp. 129–163, 1992.
- [28] C. K. Chui, *An Introduction to Wavelets*. San Diego, CA: Academic, 1992, pp. 119–176.
- [29] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Trans. Image Process.*, vol. 10, no. 11, pp. 1647–1658, Nov. 2001.
- [30] E. P. Simoncelli, "Modeling the joint statistics of image in the wavelet domain," *Proc. SPIE*, vol. 3813, pp. 188–195, Jul. 1999.
- [31] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inf. Theory*, vol. 38, no. 3, pp. 713–718, Mar. 1992.
- [32] W. Mendenhall, R. J. Beaver, and B. M. Beaver, *Introduction to Probability and Statistics*, 11th ed. Beijing, China: China Machine Press, 2005, pp. 320–354.
- [33] M. H. Kutner, C. J. Nachtsheim, and J. Neter, *Applied Linear Regression Models*, 4th ed. Beijing, China: Higher Education Press, 2005, pp. 2–212.
- [34] P. Brodatz, *Textures: A Photographic Album for Artists & Designers*. New York: Dover, 1966.
- [35] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," *Proc. IEEE ICASSP*, vol. 5, pp. 291–294, Apr. 1980.
- [36] T. Randen and J. H. Husøy, "Multichannel filtering for image texture segmentation," *Opt. Eng.*, vol. 33, pp. 2617–2625, Aug. 1994.
- [37] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001, pp. 161–201.



Zhi-Zhong Wang received the B.S. degree from the Hunan University, Changsha, China, in 2003, and the M.S. degree in software engineering from the Tsinghua University, Beijing, China, in 2007.

He is currently a faculty member in the Technique Research Department, China Center For Resource Satellite Data and Application. His current research interests include image processing, analysis, and retrieval.



Jun-Hai Yong received the B.S. and Ph.D. degrees in computer science from the Tsinghua University, China, in 1996 and 2001, respectively.

He has been a Professor with the School of Software, Tsinghua University, Beijing, China, since 2006. He held a visiting researcher position in the Department of Computer Science, Hong Kong University of Science and Technology, in 2000. He was a Postdoctoral Fellow with the Department of Computer Science, University of Kentucky, Lexington, from 2000 to 2002. His research interests

include computer-aided design, computer graphics, computer animation, and software engineering.