Fractal-based edge extraction and its application to textured image retrieval

Pi, Ming Hong; Tong, Chong Sze; Basu, Anup

Published in:
Journal of Electronic Imaging

DOI:
10.1117/1.2176674

Published: 01/01/2006

Document Version:
Publisher's PDF, also known as Version of record

Journal of Electronic Imaging 15(1), 013012 (Jan–Mar 2006)

Fractal-based edge extraction and its application to textured image retrieval

M. Pi
C. S. Tong
Hong Kong Baptist University
Department of Mathematics
Kowloon, Hong Kong
China
E-mail: cstong@hkbu.edu.hk

A. Basu
University of Alberta
Department of Computing Science
Edmonton Alberta T6G 2E8
Canada

Abstract. *We discuss how an orthogonalized fractal decoding algorithm can be used to derive a fractal decomposition of the encoded image into a dc and a set of ac components that characterize the edge information in the image. The dc component can be characterized by its histogram. The ac components provide a model for the texture, and the distribution of each ac component can be well represented by a generalized Gaussian density (GGD), which can be efficiently characterized by two GGD parameters. Such characterization of the ac components is exploited in this work for texture image retrieval, and experimental results show that our proposed indexing technique based on our novel fractal dc and ac signature provides superior retrieval performance compared to current fractal indexing techniques. Moreover, the performance of our approach is comparable to the state of the art wavelet-based method at a fraction of the retrieval time.*
 [DOI: 10.1117/1.2176674]

1 Introduction

Images now play a crucial role in a variety of fields such as crime prevention, medicine, publishing, entertainment, education, and even scientific research. As the prevalence of digital images continues to grow, it is becoming more and more important to be able to find desired images from a large image database by automatically derived features, rather than by manually assigning indices to each image laboriously. This idea is popularly known as content-based image retrieval (CBIR),[1] and the term is widely used to describe the process of retrieving desired images from a large collection using primitive features such as color,[1] shape,[2] and texture,[3–10] etc., that are automatically extracted from the images. Various CBIR approaches have been surveyed in detail.[11–16]

The early approaches exploited primitive features mainly extracted from the pixel domain, and such features as color histograms have yielded very good retrieval results. As digital images become ever more prevalent, more and more effort is made to take into account the characteristics of the compression domain, and hence there are now interests in extracting primitive features from transform domains such as discrete cosine transform (DCT),[2] Gabor transform,[7] and wavelet transform (WT).[3–6,17]

1.1 *Review of Wavelet Signatures*

Wavelets analysis provides multiresolution and orientation representation of an image via subbands, and is now widely regarded as one of the most useful tools for image characterization. In the case of textured images, wavelet-based image retrieval is considered the state of the art approach.[4] We briefly review two of the most well-known wavelet-based signatures, namely, the energy signature[3] and the generalized Gaussian distribution (GGD) signature,[4] which is used as our benchmark.

After discrete wavelet transform with N_L levels, the original image is decomposed into $3N_L+1$ subbands (one low-pass subband and $3N_L$ high-pass subbands). Each subband histogram encodes the first-order statistical property of the subband wavelet coefficients. Although the histogram signature generally provides a good representation for texture images and hence leads to a good image retrieval rate, the dimension of the histogram signature is too large and more efficient characterization based on the histogram is often adopted instead.

1.1.1 *Extended energy signature*

The (normalized) energy of the subband B_n containing M_n coefficients $w_{n,m}$ (note that the 2-D wavelet coefficients are transferred into 1-D using the row scan order) is defined as

Paper 05016R received Jan. 29, 2005; revised manuscript received Jun. 27, 2005; accepted for publication Jul. 28, 2005; published online Mar. 3, 2006.
1017-9909/2006/15(1)/013012/9/$22.00 © 2006 SPIE and IS&T.

$$E_{B_n} = \frac{1}{M_n} \sum_{m=1}^{M_n} w_{n,m}^2. \quad (1)$$

The wavelet energy signature is the feature vector $(E_{B_1}, E_{B_2}, \ldots E_{B_{3N_L}})$, which represents the energy distribution along the high-pass subbands. The absolute mean of the subband B_n can be used as an alternative measure in conjunction with energy, and is defined as follows:

$$A_{B_n} = \frac{1}{M_n} \sum_{m=1}^{M_n} |w_{n,m}|. \quad (2)$$

In practice, the energy and absolute moment are often used in conjunction to enhance the retrieval rate, and we refer to the concatenated feature vector

$$(E_{B_1}, A_{B_1}, E_{B_2}, A_{B_2}, \cdots, E_{B_{3N_L}}, A_{B_{3N_L}}), \quad (3)$$

as the extended energy (EE) signature.

1.1.2 *Generalized Gaussian density signature*

Mallat[18] noticed that, for a variety of images, the distribution of subband wavelet coefficients appears symmetrical about zero and has a sharp peak at zero. Mallat proposed modeling "typical wavelet coefficients" of the subbands using the generalized Gaussian density (GGD) defined by:

$$p(x;\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp[-(|x|/\alpha)^{\beta}], \quad (4)$$

where $\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt, z > 0$, is the Gamma function; α corresponds to the width of the probability density function (standard deviation); and β corresponds to the shape of GGD. As β gets smaller, the GGD becomes sharper, and the height of the peak increases. These GGD parameters α and β can be estimated by moment estimation (ME)[19] or maximum likelihood estimation (MLE).[4] We refer to $\{\alpha_{B_n}, \beta_{B_n}\}_{n=1,2,\ldots,3N_L}$ as the GGD signature, where α_{B_n} and β_{B_n} denote the GGD parameters of the subband B_n.

The retrieval rate based on the L_1 distance of energy signature is generally lower than that of the GGD signature [used in conjunction with the Kullback-Leibler distance (KLD)][4] and hence we use the GGD signature as our benchmark.

1.2 *Review of Fractal-Based Signatures*

Since the 1980's, a new technique of fractal compression has been developed based on the encoding of self-similarity inherent within a digital image. Like vector quantization, fractal block coding[20,21] encodes an image in a block-wise manner. The encoding of each range block consists of finding the best affine map to minimize the distortion between a range block and a domain block by searching a global or partial domain block pool derived from the image itself. In the decoding, the image is reconstructed by iterating the affine maps on an arbitrary initial image. Convergence to the original image is assured by selecting contractive affine maps.[21]

It was recognized early on that the encoding of the self-similarity could be exploited for value-added image processing applications in addition to its potential for high compression. Pentland[8] was the first to propose a fractal signature extracted from the pixels for texture characterization. Zhang, Cheng, and Acharya[22] proposed a fractal indexing technique where the fractal codes are directly used as the image index [referred to as the fractal codes (FC) technique]. However, the FC technique has very high computational complexity, and the retrieval rate is rather low for a texture image database. Schouten and Zeeuw[9] extended the histogram-based technique to one of the affine map parameters, namely, the contrast scaling, and proposed to use the histogram of contrast scaling parameters as an image index [referred to as the histogram with quantized contrast scaling (HWQCS) technique]. This indexing technique is very fast, but unfortunately, its retrieval rate is not high enough.

Recently, Pi, Mandal, and Basu proposed to exploit statistical properties extracted from the fractal transform domain for texture image retrieval.[10] They proposed to employ the 2-D joint histogram of the means of the range blocks (range-block means) and contrast scaling parameters as an image index [referred to as the histogram of new fractal parameters (HNFP) technique]. Compared to previous fractal indexing techniques, the HNFP technique greatly improved on the retrieval rate. However, the storage requirement and computational complexity increased significantly.

The conventional affine parameters for the fractal encoding of digital images are the contrast scaling and luminance offset. However, it has been shown that these parameters are strongly correlated, and this correlation has limited the performance of image retrieval systems based on signatures derived from these fractal parameters. In this work, we discuss an orthogonalized fractal decoding algorithm based on an alternative set of affine parameterizations, in which the parameters are statistically independent, and show how it can be used to derive a fractal decomposition of the encoded image into a dc and a set of ac components. The dc component provides a coarse representation of the image, and it can be characterized by its histogram. The ac components correspond to high-frequency information, such as edges and texture. In this work, we show that the distribution of each ac component can be modeled by a general Gaussian density (GGD), which can be compactly characterized by two GGD parameters. We then combine the dc histogram and the ac GGD parameters as an image signature for a novel and effective fractal indexing technique. Experimental results show that the proposed indexing technique provides competitive retrieval performance compared to existing fractal indexing techniques.

The remainder of the work is organized as follows. In Sec. 2, we briefly summarize the technique of fractal block coding using both the conventional affine parameters as well as an alternative orthogonalized affine parameterization. The idea of fractal decomposition of an image into dc and ac components is discussed based on a fractal decoding algorithm using the alternative affine parameterization. The section closes with a description of several of the current fractal indexing techniques. The concept of dc+ac signature is presented in Sec. 3, and experimental results are reported in Sec. 4, which is followed by our conclusions.

2 Review of Fractal Coding and Fractal Indexing

2.1 *Fractal Block Coding and Fractal Decomposition*

In fractal image coding, an original image is generally partitioned into nonoverlapping square range blocks of size $B \times B$. A domain block pool Ω is then obtained by sliding a window of size $2B \times 2B$, starting at the top left corner of the image, with search step-size δ, in both the horizontal and vertical directions. Hence, the domain block pool Ω includes

$$\left(\left\lfloor\frac{M-2B}{\delta}\right\rfloor+1\right) * \left(\left\lfloor\frac{N-2B}{\delta}\right\rfloor+1\right)$$

domain blocks for an $M \times N$ image (where $\lfloor x \rfloor$ denotes the integer part of x).

For each range block $R=\{r_{ij}\}$, we search the domain block pool Ω to find a domain block $D=\{d_{ij}\}$ and an affine transformation τ [i.e., $\tau(D)=s^*\sigma(D)+g$], such that $\tau(D)$ provides the best matching for R in the least squares sense. The parameters s and g are the scaling and luminance offset, respectively, and $\sigma(\cdot)$ is a decimation operator to contract the $2B \times 2B$ domain blocks into the same size as the $B \times B$ range blocks.

Although the scaling and luminance offset are the most obvious parameterization of the affine transform, Tong and Pi have shown that they are strongly correlated and advocated the use of an alternative, orthogonalized parameterization[23] in which the affine transformation parameter, luminance offset, is replaced by the mean of each range block. In terms of decoding, the two different parameterizations lead to different decoding algorithms, namely:

$$R_{\text{old}}^{(n)} = gU + s\sigma[D^{(n-1)}], \tag{5}$$

for the conventional parameterization, and

$$R_{\text{new}}^{(n)} = \bar{r}U + s\{\sigma[D^{(n-1)}] - \overline{\sigma[D^{(n-1)}]}U\}, \tag{6}$$

for the orthogonalized parameterization. Here, U denotes a $B \times B$ block with all entries equal to 1, $\overline{\sigma[D^{(n-1)}]}$ denotes the mean value of the decimated domain block $\sigma[D^{(n-1)}]$, and $\bar{r}$ denotes the mean value of the current range block R (referred to as the range block mean).

The superior convergence property of the orthogonalized decoding algorithm in Eq. (6) over the conventional decoding algorithm in Eq. (5) has been established by Øien and Lepsøy.[24] They also proved that their decoding algorithm converges to the attractor in a finite number of iterations. At the n'th iteration, a range block for the current decoded image $R_{\text{new}}^{(n)}$ consists of two parts: $\bar{r}U$ and $s[D^{(n-1)}-\overline{D^{(n-1)}}U]$. The first part is a constant block with the constant value set as the range block mean, and is equivalent to the dc component in JPEG decoding. This dc component is the range-average image, which is produced by replacing each pixel of the range block with the average of the range block. It is the optimal initial image for fractal decoding using both algorithms in Eq. (5) or Eq. (6).[24,25] The second part is referred to as the ac component of the decoded image.

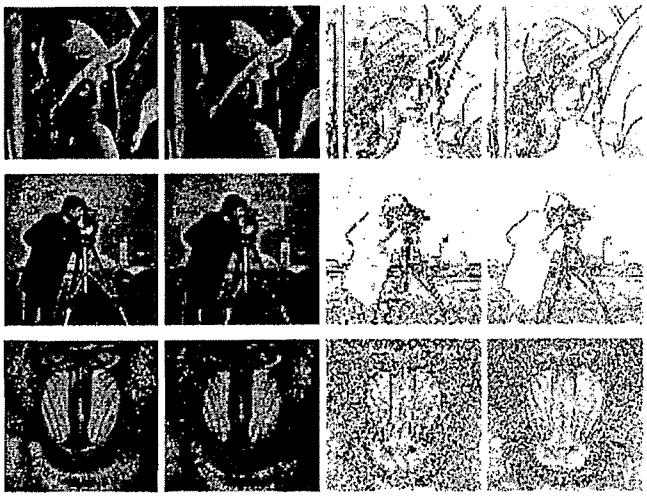


Fig. 1 Fractal decomposition for three 256×256 standard images. Column 1 is the decoding image; column 2 is the dc image (Ψ_{dc}); column 3 is the ac1 image (Ψ_{ac_1}); and column 4 is the ac2 image (Ψ_{ac_2}), $\Psi^{(2)}=\Psi_{\text{dc}}+\Psi_{\text{ac}_1}+\Psi_{\text{ac}_2}$.

For the special case when the search step size $\delta=B=2^b$, and the initial image is the dc component, then the decoding algorithm in Eq. (6) reaches its attractor in exactly b iterations. Each iteration produces a new layer of the ac component, and so the decoded image Ψ can be considered to be decomposed into the sum of the dc and ac components

$$\Psi^{(b)} = \Psi_{\text{dc}} + \Psi_{\text{ac}_1} + \Psi_{\text{ac}_2} + \cdots + \Psi_{\text{ac}_b}. \tag{7}$$

Figure 1 shows the fractal decomposition for three standard images where the images are encoded with $2^2 \times 2^2$ range blocks (i.e., $\delta=B=2^2$). The decoded images are decomposed into the dc component and two ac components, which describe high-frequency information, such as edges and textures, of the images. Note that the second ac component is finer than the first one, and so these ac components suggest that they may be considered as providing a multiresolution decomposition of the high-frequency content of the image.

2.2 *Quantization of Fractal Parameters*

So far, we have confined our discussion to the case of continuous parameterization. In practice, we need to quantize the parameters for useful compression ratios. The quantization of the conventional fractal parameters, namely, scaling and luminance offset, is very much a trial-and-error process due to the correlation between the two parameters. However, the alternative parameters, namely, scaling and the range block mean, are statistically independent, and their optimal quantization has been well established.[23] Thus the fractal encoding algorithm becomes the problem of minimizing the distortion:

$$\hat{E}(R,D) = \|R - \bar{r}_i U - s_j(D - \bar{D}U)\|^2 \tag{8}$$

over $D \in \Omega$ by working with a set of prequantized fractal parameters $\{\bar{r}_i\}$ and $\{s_j\}$ (where $0 < s_j \leq 1$). The fractal code of R is then calculated as

$$(\bar{r},s,x_D,y_D) = \arg\min_{D\in\Omega} \hat{E}(R,D),$$

where (x_D,y_D) are the coordinates of the top-left corner of the "best-matching" domain block. We do not describe our quantization procedure in detail here, except to point out that all the properties of the decoding algorithm in Eq. (6) and the corresponding fractal decomposition are preserved by our quantization.[26]

2.3 *Summary of Current Fractal Indexing Techniques*

In this work, we compare our approach to current fractal-based techniques. Zhang, Cheng, and Acharya[22] use the fractal codes (FC technique) directly as the image signature. This technique leads to rather long signatures and hence very long retrieval times, but nevertheless the retrieval rates remain low. Schouten and Zeeuw[9] proposed the HWQCS technique, where the histogram of weighted quadtree contrast scaling parameters is used as the image index. For single-level fractal coding, the index reduces to the histogram of the contrast scaling s, expressed as:

$$p_s(s_j) \quad (j = 1,2,\cdots,J).$$

Although, the HWQCS technique is fast, the retrieval rate is low.

Pi, Mandal, and Basu[10] further improved on the HWQCS technique by incorporating the histogram of the range block mean $\bar{r}$. Their proposed HNFP technique uses the joint histogram of $\bar{r}$ and s as statistical indices, which are expressed as:

$$p_{\bar{r}}(\bar{r}_i) \quad (i = 1,\cdots,I) \text{ and}$$

$$\{q(\bar{r}_i,s_j)\} \quad (i = 1,\cdots,I;\ j = 1,\cdots,J).$$

The HNFP indexing technique greatly improves on the retrieval rate, compared with previous fractal-based retrieval methods. However, the computational complexity is high and the retrieval rate still needs to be enhanced for practical use.

In addition to the fractal-based techniques mentioned in this section, we also compare our approach to the current state of the art retrieval method based on wavelet transform.[4]

3 Proposed Indexing Technique

3.1 *Modeling dc and ac Coefficients*

Under the orthogonal decoding algorithm, the decoded image is decomposed into the sum of the dc component and several layers of the ac components. The dc component is a coarse approximation for the image and is characterized by its histogram (referred to as the dc signature). We use 16 bins for this dc signature.

In wavelet-based image retrieval techniques, the ac components (subbands) of the wavelet transform have been observed to be distributed as the generalized Gaussian density (GGD), and powerful image signatures have been derived from the GGD parameters for each such subband.[4,18] In a similar vein, we observe that for a variety of images, the distribution of each layer ac component in our fractal decomposition seems to be well modeled using the GGD distribution as defined in Eq. (4) in Sec. 1.1.2. We refer to the GGD parameters of the ac components as the ac signature.

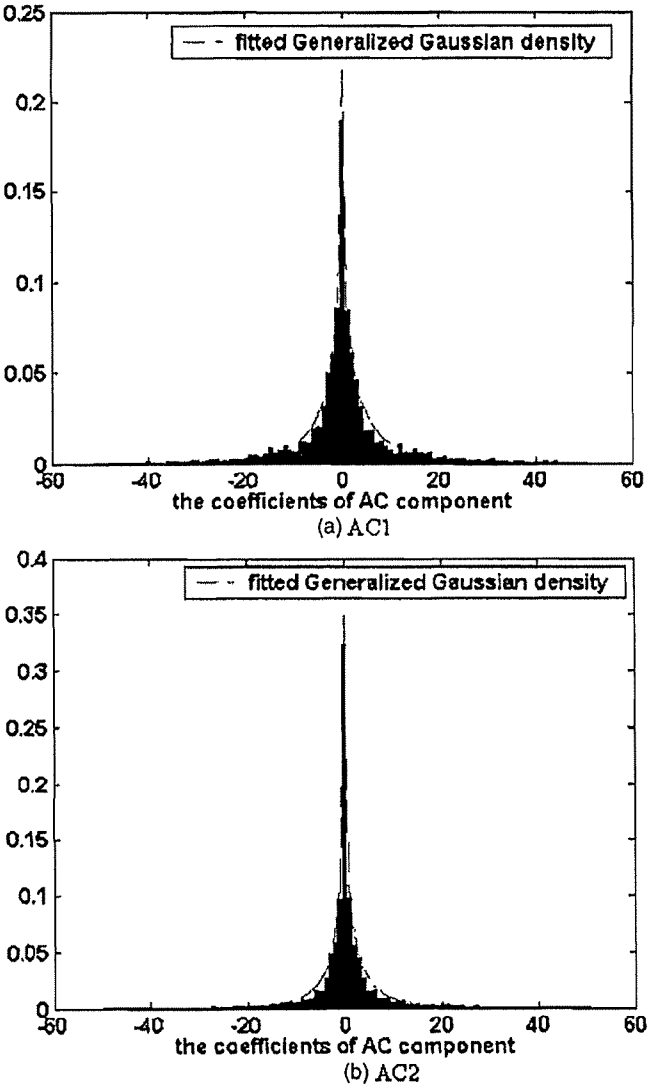


Fig. 2 Distribution of ac coefficients fitted with a GGD.

Figure 2 shows the distributions and GGD fitting of Ψ_{ac_1} and Ψ_{ac_2} for the 256×256 Lena image, and demonstrates that the distribution of the ac component can be accurately fitted by GGD. As a result, the ac component can be compactly characterized by the GGD parameters (α and β).

3.2 *Similarity Measurement*

Having defined our dc and ac signatures, we now proceed to discuss how these signatures are to be compared. We employ different measures of similarity between dc and ac signatures. For the dc signature, which can be viewed as a form of grayscale histogram technique, the L_1-norm is the natural similarity measurement:

$$H(Q,C)=\sum_{i=1}^{I}|h_i^{(Q)}-h_i^{(C)}|,$$

where $\{h_i^{(Q)}\}$ and $\{h_i^{(C)}\}$ is the normalized dc component's histograms of the query image (Q) and candidate image (C), and I is number of bins. Of course, L_2-norm can also be used, but in practice, the results are similar to the L_1-norm, and since the L_2-norm can be sensitive to outliers, we do not consider this norm in this work.

For the ac signature, we may also use the L_1-norm to define a similarity measurement as follows:

$$G(Q,C)=\sum_{k=1}^{b}|\alpha_1^{(k)}-\alpha_2^{(k)}|+\sum_{k=1}^{b}|\beta_1^{(k)}-\beta_2^{(k)}|,$$

where $\{\alpha_1^{(k)},\beta_1^{(k)}\}$ and $\{\alpha_2^{(k)},\beta_2^{(k)}\}$ are the GGD parameters of the k'th ac components of the query image (Q) and candidate image (C), respectively. However, the GGD distributions are quite sensitive to the change of its parameters, i.e., subtle deviations of the GGD parameters can lead to very different shapes of the GGD distributions. Thus, a more natural similarity measurement should be based on the similarity of the actual distributions rather than the corresponding parameters. The similarity between distributions can be described by the minimal Kullback-Leibler distance (KLD), and it has been shown that for the wavelet-based retrieval technique, the KLD is a much more appropriate measure than L_1-norm: it significantly improves the retrieval rate.[4] The Kullback-Leibler distance between two GGD distributions is defined as follows:

$$\begin{aligned}K[\{\alpha_1^{(k)},\beta_1^{(k)}\},\{\alpha_2^{(k)},\beta_2^{(k)}\}]&=\log\left\{\frac{\beta_1^{(k)}\alpha_2^{(k)}\Gamma[1/\beta_2^{(k)}]}{\beta_2^{(k)}\alpha_1^{(k)}\Gamma[1/\beta_1^{(k)}]}\right\}\\&+\left[\frac{\alpha_1^{(k)}}{\alpha_2^{(k)}}\right]^{\beta_2^{(k)}}\frac{\Gamma\{[\beta_2^{(k)}+1]/\beta_1^{(k)}\}}{\Gamma[1/\beta_1^{(k)}]}\\&-\frac{1}{\beta_1^{(k)}},\end{aligned}\tag{9}$$

and the minimal KLD (MKLD) is

$$\begin{aligned}&MK[\{\alpha_1^{(k)},\beta_1^{(k)}\},\{\alpha_2^{(k)},\beta_2^{(k)}\}]\\&\quad=\min\{K[\{\alpha_1^{(k)},\beta_1^{(k)}\},\{\alpha_2^{(k)},\beta_2^{(k)}\}],K[\{\alpha_2^{(k)},\alpha_2^{(k)}\},\{\alpha_1^{(k)},\beta_1^{(k)}\}]\}.\end{aligned}$$

The KLD between two ac signatures is the sum of MKLDs over all the ac components.

$$MK(Q,C)=\sum_{k=1}^{b}MK[\{\alpha_1^{(k)},\beta_1^{(k)}\},\{\alpha_2^{(k)},\beta_2^{(k)}\}].$$

We have thus defined two different distances to measure the similarity between the query image (Q) and candidate image (C):

$$D_1(Q,C)=\gamma_1 G(Q,C)+H(Q,C),\tag{10}$$

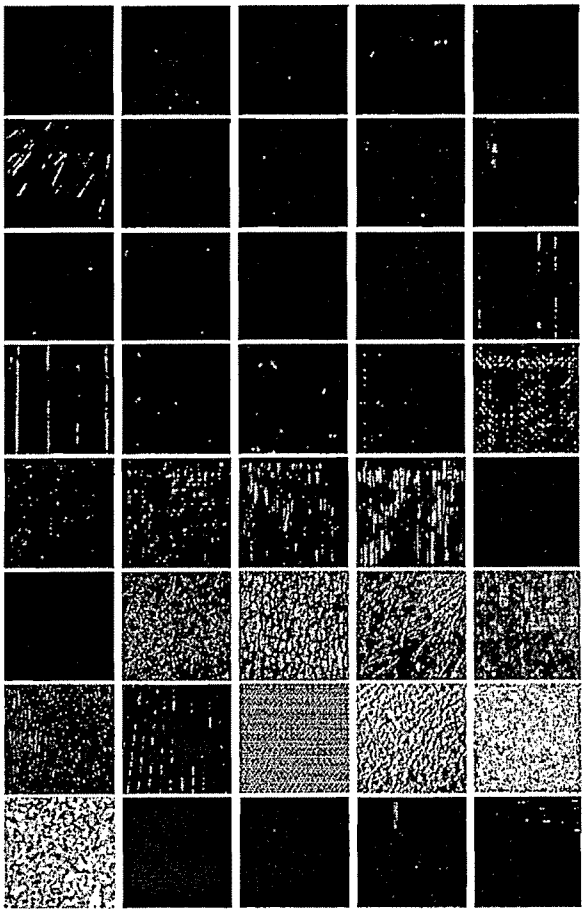


Fig. 3 40 512×512 grayscale Brodatz texture images used in the experiments; from left to right and top to bottom: image 1, image 2, ..., image 40.

$$D_2(Q,C)=MK(Q,C)+\gamma_2 H(Q,C),\tag{11}$$

where γ_1 and γ_2 are the weight factors, which balance the range of two distances to ensure the contribution of dc and ac signatures are equally important. In the following experiments, $\gamma_1=1/2$, $\gamma_2=0.1$; note that γ_2 is quite small because the MK measure is in terms of the Kullback-Leibler distance and is of a smaller scale than the L_1 measure.

4 Performance Evaluation

40 512×512 grayscale Brodatz[27] texture images were selected for evaluation (see Fig. 3). Each image is divided into 16 128×128 nonoverlapping subimages to create a test database of $Z=640$ texture images. In retrieval experiments, a query image is selected randomly from the test database. 16 images are then retrieved based on the smallest distance criterion. Ideally, all 16 images corresponding to the selected query image should be retrieved in each test. However, this does not generally happen in practice. We evaluate the performance of various signatures in terms of average retrieval defined by Eq. (12). Let F denote the number of retrieved images (in this case $F=16$), and m_z denote the number of correctly retrieved images at the z'th test. The average retrieval rate is then calculated as:

Table 1 Average retrieval rate (ARR) for various indexing techniques. TSM means the total times for evaluating the similarity measurements for 640 query images. Δ is the quantization step size.

Indexing techniques	Texture signature	ARR (%)	Length of signature	TSM (seconds)
	FC	23.4	1024	669.92
	HWQCS (L=1)	38.2	4	0.17
Fractal	HNFP	71.2	256	5.72
indexing	dc signature	57.4	64	1.4
techniques		55.3	16	0.45
	ac signature ($L1$)	41.8	4	0.22
	ac signature (KLD)	52.5	4	4.2
	dc+ac signature (D_1)	65.4	20	0.83
	dc+ac signature (D_2)	74.4	20	4.91
Color indexing	Histogram of grayscale	55.8	16	0.46
Wavelet-based	GGD+KLD(D97) (Δ=1)	63.8	18	18.76
indexing	GGD+KLD(D8) (Δ=1)	78.8	18	18.69
technique	GGD+KLD(D8) (Δ=2)	72.7	18	18.53

$$\text{average retrieval rate} = \frac{\sum_{z=1}^{Z} m_z}{F \times Z}. \quad (12)$$

All programs are implemented using VC++6.0 and are integrated using the VC++ graphical user interface. In our experiment, all texture images are encoded using the fast fractal image compression technique in Ref. 24 with $2^2 \times 2^2$ range block and the step size $\delta=B=2^2$, which means that we have two ac components. The range block mean $\bar{r}$ and contrast scaling s are quantized to 6 and 2 bits, respectively, using the optimal quantization scheme described in Ref. 24.

All experiments were conducted on a Pentium 4 (2 Ghz) PC, and each experiment consists of 640 query images from the test database of 640 texture images. The experimental results are summarized in Table 1. The actual fractal compression time is around 1 msec per image, but we store the texture signature for each image in the database, and hence the computational complexity for an image retrieval system is just measured by the total times for evaluating the similarity measurements (referred to as TSM in Table 1).

4.1 *Results of the Existing Fractal-Based Signatures*

The length of the signature in the FC technique is proportional to the image size, which means the FC technique is very time consuming (see Table 1). But still, the FC technique yielded the lowest retrieval rate (23.4%). This is probably attributable to the fact that the fractal codes are in general very sensitive to even minor changes to the image. Thus, images that look similar and have the same texture would nevertheless have drastically different fractal codes. The underwhelming performance of the FC technique strongly suggests that one should avoid using image features directly as signature, but instead call for some sort of statistical characterization of the texture images.

The HWQCS technique exploits the statistical property of the scaling parameter. As a result, compared to the FC technique, the retrieval rate is enhanced, and computational complexity is significantly reduced. However, its retrieval rate is still a mere 38.2%, far lower than what one would demand for a practical system.

The HNFP technique exploits the joint statistical property of range block mean and scaling parameter, which significantly enhances the retrieval rate (71.2%), but at the expense of a sizeable increase in complexity and storage requirement: the retrieval time (as measured by the TSM given in Table 1) increased more than 30 fold.

4.2 *Results of Our Proposed Signature*

As mentioned in Sec. 2, the decoded image is the sum of the dc component and ac components. The dc component is a coarse scale representation of the image, and hence the dc signature (histogram of $\bar{r}$) is expected to yield a retrieval rate close to that by the grayscale histogram of the original image. Indeed, from Table 1, using a bin size of 16, the dc signature yields a retrieval rate of 55.3% compared to 55.8% from that of the grayscale histogram (so-called color indexing technique). In this sense, our dc+ac signature can be thought of as an extension of the grayscale histogram technique, with the ac components providing global edge information to supplement the dc component.

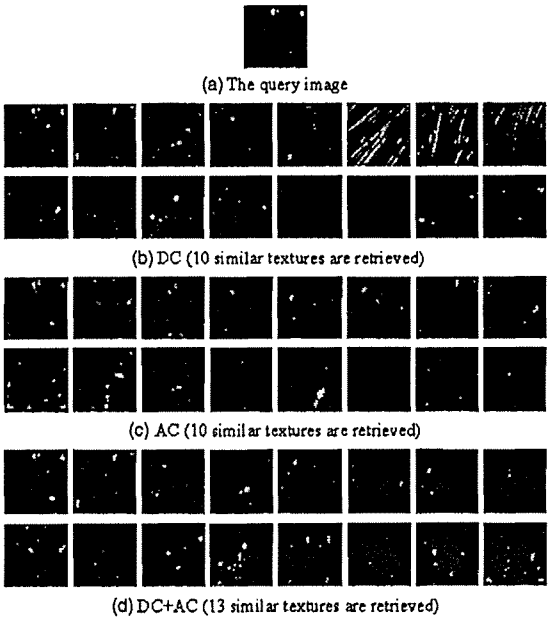


Fig. 4 The top 16 retrieved images using dc, ac, and dc+ac signature.

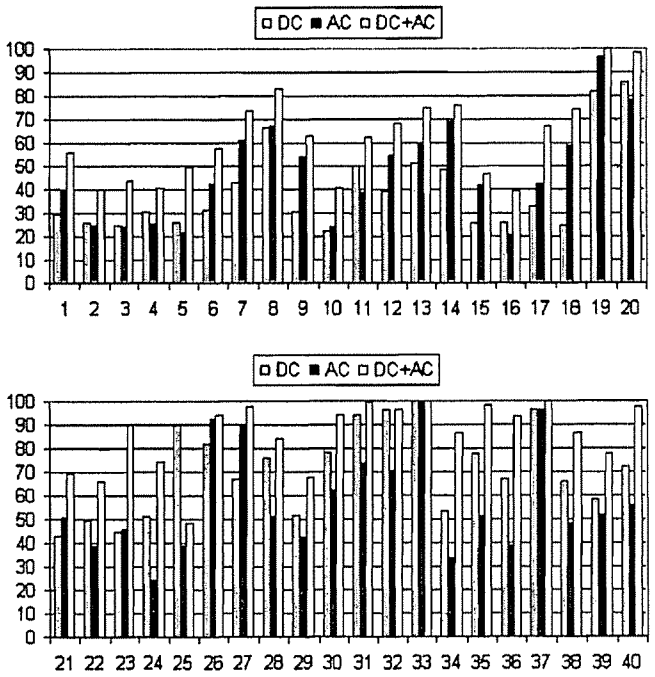


Fig. 5 Average retrieval rates of dc, ac, and dc+ac signatures for individual texture classes. The horizontal axis represents the texture class labels, while the vertical axis represents the retrieval rate in percentages.

As explained in Sec. 3.3, the ac signature can be used in two ways, depending on which similarity measures we adopt. Our results (in Table 1) showed that the KLD is a much more appropriate measure than L_1-norm. For the ac signature using the KLD measure, the retrieval rate is 52.5%, whereas, the retrieval rate using the L_1-norm is only 41.8%, representing a difference of more than ten percentage points. This finding is consistent with the experience in wavelet-based retrieval.[4]

For the dc+ac signature (under the D_2 distance measure), the retrieval rate using KLD for the ac signature yields a retrieval rate of 74.4%, whereas the corresponding retrieval rate using the L_1-norm (i.e., the dc+ac signature with D_1 distance measure) is 65.4%, again around ten percentage points difference. However, the KLD measure requires costly computation (compared to L_1-norm), because it involves operations such as logarithm, exponential, and gamma functions, etc. [see Eq. (9)]. Nevertheless, overall the dc+ac signature (D_2) still outperforms the HNFP technique, with a higher retrieval rate (74.4 versus 71.2%), achieved at a faster retrieval time (4.91 sec for the batch of experiments versus 5.72 sec). Moreover, since our new dc +ac signature is only 20 components long as compared to the length of 256 components in the HNFP signature, our method requires a much smaller storage than the HNFP method. We are also currently working on a new similarity measure for the dc+ac signature that may be able to exploit the shorter signature for much faster retrieval time while maintaining comparable retrieval rates.

4.3 *Quality of the Retrieval*

The dc signature and the ac signature provide complementary statistical models for the texture images. Figure 4 shows a query image and the top 16 images retrieved using the dc signature, ac signature, and dc+ac signature, respectively. Using either the dc or ac signature alone, only ten similar texture images were obtained in the top 16 retrieval. In particular, some visually quite different texture images were retrieved using the dc signature alone. By contrast, the composite dc+ac signature retrieves 13 textures similar to that of the query image. Moreover, the three different textures retrieved were ranked in the last four of the 16 retrievals. Figure 5 further reports average retrieval rates of dc, ac, and dc+ac signatures for the 40 individual texture classes, and demonstrates that the dc+ac signature consistently and significantly improves on either individual signature alone.

4.4 *Comparison with Wavelet-Based Methods*

Our dc+ac signature provides performance comparable with the state of the art wavelet-based indexing technique (GGD+KLD). The retrieval rate GGD+KLD provides depends on the choice of wavelet filters. For the biorthogonal wavelet (Daubechies 9-7 filter),[28] the retrieval rate is 63.8%, somewhat lower than our method. For the orthogonal wavelet (Daubechies D8),[28] the retrieval rate is 78.8%, which is higher than the 74.4% achieved by the dc+ac signature. But the GGD parameters are extracted from almost nonquantized wavelet coefficients ($\Delta=1$).[29] If the GGD parameters are extracted from the quantized wavelet coefficients ($\Delta=2$), the retrieval rate becomes 72.7%, marginally lower than the retrieval rate of the dc+ac signature. It should be pointed out that for the texture images used in this study, the compression ratio for $\Delta=2$ is close to 6.4 times, which is the compression ratio achieved by fractal block coding with a 4×4 range block. Thus, the $\Delta=2$ comparison may be more meaningful. Furthermore, the retrieval speed of the wavelet-based GGD+KLD method is very slow, and its similarity measurement time (TSM) is

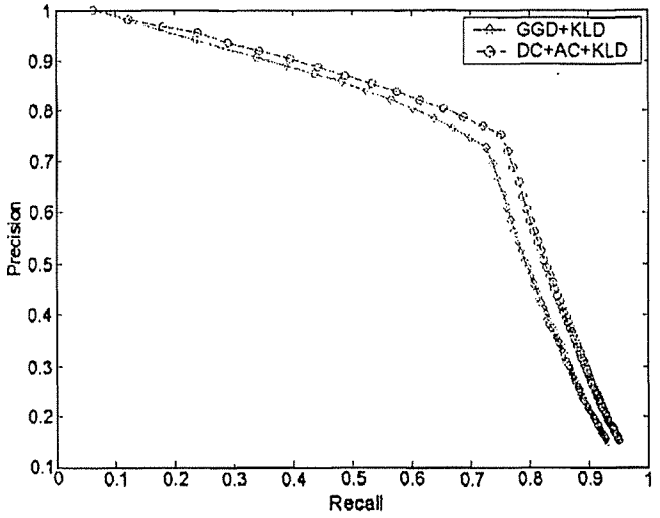


Fig. 6 Precision-recall curves for our dc+ac+KLD signature (i.e., using D_2) versus the wavelet-based GGD+KLD (D8, $\Delta=2$) signature.

3.7 times more than that of the dc+ac signature. Finally, we compare the precision-recall curves for our method with the wavelet-based method ($\Delta=2$ case) in Fig. 6. Our curve is always above the wavelet curve, showing that our method has a consistently better retrieval performance than the wavelet-based signature.

5 Conclusions

Using the orthogonal fractal decoding algorithm, the decoded image is decomposed into the sum of dc and ac components. We propose different but complementary statistical models for the dc and ac components: we characterize the dc component by its histogram, and the ac components by GGD parameters. On that basis, we present the dc+ac signature as a statistical index of the proposed indexing technique, which can be viewed as an extension of the color indexing technique based on a grayscale histogram.

The experimental results show that the dc+ac signature provides better performance in both retrieval rate and retrieval time compared to the existing fractal indexing techniques. Moreover, the dc+ac signature provides performance comparable with the current state of the art indexing technique based on wavelet transform. At similar compression levels, our method slightly outperforms the wavelet approach in retrieval rate. Our method also has the advantage that it is faster than the wavelet-based method, and also does not require fine tuning of the wavelet filters.

Although our discussion is focused on single-level fractal block coding, our method can be extended to multilevel quadtree fractal coding. We are also investigating new similarity measures for the dc+ac signature that may be able to exploit the shorter signature for much faster retrieval time while maintaining comparable retrieval rates.

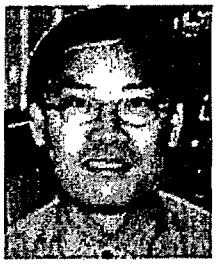
Acknowledgment

The authors thank Dr. M. L. Tang for helpful discussions, comments, and suggestions. This research was partially supported by the HKBU grant FRG/02-03/II-04 and the CERG grant HKBU 2021/04P.

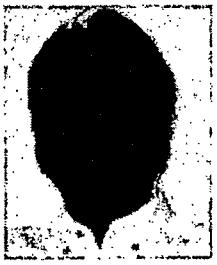
References

1. M. Swain and D. Ballard, "Color indexing," *Int. J. Comput. Vis.* **7**(1), 11–32 (1991).
2. R. Mehrotra and J. E. Gary, "Similar-shape retrieval in shape data management," *IEEE Computer* **28**(9), 57–62 (1995).
3. A. Laine and J. Fan, "Texture classification by wavelet packet signature," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**, 1186–1193 (1993).
4. M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Process.* **11**(2), 146–158 (Feb. 2002).
5. T. Chang and C. C. J. Kuo, "Texture analysis and classification using tree-structured wavelet transform," *IEEE Trans. Image Process.* **2**(4), 429–441 (Oct. 1993).
6. M. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Process.* **46**(4), 886–902 (Apr. 1998).
7. B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(8), 837–842 (Aug. 1996).
8. A. Pentland, "Fractal-based description of natural scenes," *IEEE Trans. Pattern Anal. Mach. Intell.* **6**(6), 661–674 (1984).
9. B. Schouten and P. Zeeuw, "Image databases, scale and fractal transforms," *Proc. ICIP* **2**, 534–537 (2000).
10. M. Pi, M. Mandal, and A. Basu, "Image retrieval based on histogram of new fractal parameters," *Proc. ICASSP* (2003).
11. P. Aigrain, H. Zhang, and D. Petkovic, "Content-based representation and retrieval of visual media: a state of the art review," *Multimedia Tools Appl.* **3**, 179–202 (1996).
12. F. Idris and S. Panchanathan, "Review of image and video indexing techniques," *J. Visual Commun. Image Represent* **7**(1), 28–43 (Mar. 1996).
13. M. K. Mandal, F. Idris, and S. Panchanathan, "A critical evaluation of image and video indexing techniques in the compressed domain," *Image Vis. Comput.* **17**(7), 513–529 (May 1999).
14. J. Eakins and M. Graham, "Content-based image retrieval," Report to JISC Technology Applications Programme (Jan. 1999), see http://www.unn.ac.uk/iidr/report.html.
15. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(12), 1349–1380 (2000).
16. M. L. Kherfi, D. Ziou, and A. Bernardi, "Image retrieval from the World Wide Web: issues, techniques, and systems," *ACM Comput. Surv.* **36**(1), 35–67 (Mar. 2004).
17. M. Shneier and M. A. Mottaleb, "Exploiting the JPEG compression scheme for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(8), 849–853 (Aug. 1996).
18. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(7), 674–693 (Jul. 1989).
19. K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," *IEEE Trans. Circuits Syst. Video Technol.* **5**, 52–56 (1995).
20. A. E. Jacquin, "Fractal image coding: A review," *Proc. IEEE* **81**(10), 1451–1465 (1993).
21. E. W. Jacobs, Y. Fisher, and R. D. Boss, "Image compression: a study of iterated transform method," *Signal Process.* **29**(3), 251–263 (Dec. 1992).
22. A. Zhang, B. Cheng, and R. Acharya, "An approach to query-by-texture in image database system," *Proc. SPIE* **2606**, 338–349 (1995).
23. C. S. Tong and M. Pi, "Fast fractal image compression using adaptive search," *IEEE Trans. Image Process.* **10**(9), 1269–1277 (Sep. 2001).
24. G. E. Øien and S. Lepsøy, "Fractal-based image coding with fast decoder convergence," *Signal Process.* **40**, 105–117 (1994).
25. Y. H. Moon, H. S. Kim, and J. H. Kim, "A fast fractal decoding algorithm based on the selection of an initial image," *IEEE Trans. Image Process.* **9**(5), 941–945 (May 2000).
26. M. Pi, A. Basu, and M. Mandal, "A new decoding algorithm based on range block mean and contrast scaling," *IEEE Intl. Conf. Image Process. (ICIP)* **2**, 271–274 (2003).
27. See http://www.cipr.rpi.edu/resource/stills/brodatz.html.
28. I. Daubechies, *Ten Lectures on Wavelets*, SIAM Publishers, Philadelphia, PA (1992).
29. "JPEG2000 part 2 final committee draft," *ISO/IEC JTC1/SC29/WC1 N2000* (Dec. 2000).

Ming Hong Pi received the BA, MS, and PhD degrees from Hubei Normal University, XiangTan University, and Huazhong University of Science and Technology (HUST), China. From 1988 to 1992, he was with the Institute of Parallel Computation Research and developed petroleum exploration software. From 1992 to 1998, he was a lecturer and then associate professor with the Department of Mathematics, HUST. He is currently a research associate fellow. His current research interests include statistical image processing, adaptive image segmentation, content-based image retrieval, watermarking, image compression, and parallel and distributed systems. He has published more than 30 papers in refereed conference proceedings and journals.



C. S. Tong received a BA degree in mathematics and a PhD degree (on mathematical modeling of intermolecular forces) from Cambridge University. After graduation, he joined the signal and image processing division of General Electric Company-Marconi's Hirst Research Centre as a research scientist, working on image restoration and fractal image compression. He joined the Department of Mathematics at Hong Kong Baptist University in 1992 and is currently an associate professor. He is a member of the IEEE, a Fellow of the Institute of Mathematics and Its Application, and a chartered mathematician. His current research interests include image processing, fractal image compression, pattern recognition, and neural networks.



Anup Basu completed his BS in mathematics/statistics and MS in computer science from the Indian Statistical Institute. He then worked at Tata Consultancy Services in New Delhi and the biostatistics division of Strong Memorial Hospital, Rochester, New York. Subsequently he completed a PhD in computer science at University of Maryland, College Park. He is currently a professor at the University of Alberta, Edmonton, Canada. He has published more than 100 papers, patents, book and book chapters, in leading conferences and journals. He has also helped startup several technology companies. His current research interests include computer vision, graphics, multimedia, and multimedia communications over heterogeneous networks. He pioneered the use of foveation in image and video compression and stereo visualization in 1992 to 1993. He also introduced several new panoramic image sensors in the mid-1990s. These research directions have been subsequently pursued by many researchers in leading institutes around the world.