Transactions Letters

Statistical Modeling in the Wavelet Domain for Compact Feature Extraction and Similarity Measure of Images

Hua Yuan and Xiao-Ping Zhang, *Senior Member, IEEE*

***Abstract*—Image feature extraction and similarity measure in feature space are active research topics. They are basic components in a content-based image retrieval (CBIR) system. In this letter, we present a new statistical model-based image feature extraction method in the wavelet domain and a novel Kullback divergence-based similarity measure. First, a Gaussian mixture model (GMM) and a more systematic generalized Gaussian mixture model (GGMM) are employed to describe the statistical characteristics of the wavelet coefficients and the model parameters are employed to construct a compact image feature space. A nontrivial expectation-maximization (EM) algorithm for the GGMM model is derived. Subsequently, a new Kullback divergence-based similarity measure with low-computation cost is derived and analyzed. The Brodatz texture image database and some other image databases are used to evaluate the retrieval performance based on the presented new methods. Experimental results indicate that the GMM and the GGMM-based image texture features are very effective in representing multiscale image characteristics and that the new methods outperforms other conventional wavelet-based methods in retrieval performance with a comparable level of computational complexity. It is also demonstrated that for image features extracted by the new statistical models, the similarity measure based on Kullback divergence is more effective than conventional similarity measures.**

***Index Terms*—Expectation-maximization algorithm, feature extraction, Gaussian mixture model, generalized Gaussian mixture model, image retrieval, Kullback divergence, similarity measure, wavelet transforms.**

I. Introduction

With the development of the Internet and digital storage techniques, digital image libraries have been widely used for commercial and research purposes. Image feature extraction and similarity measure in feature space are active research topics with many applications such as image classification, clustering and retrieval. For example, they are often basic components in content-based image retrieval (CBIR) systems [1].

Manuscript received October 14, 2008; revised April 26. First version published September 1, 2009; current version published March 5, 2010. This work is supported in part by Natural Sciences and Engineering Research Council of Canada, Grant RGPIN239031. This paper was recommended by Associate Editor Y. Rui.

The authors are with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, Ontario M5B 2K3, Canada (e-mails: hyuan@ee.ryerson.ca; xzhang@ee.ryerson.ca).

Digital Object Identifier 10.1109/TCSVT.2009.2031396

The image texture features, along with the color and shape features, are always important in low-level image feature extraction. Some researchers take advantages of the multiresolution benefits of image transforms in the compressed domain and developed feature extraction methods employing Gabor filters [2], the pyramid-structured wavelet transform (PWT) and the tree-structured wavelet transform (TWT) [3]. In the TWT-based method, energy values of the subbands are calculated and used to extract image features. In the PWT-based method, only the scaling subband is decomposed by the filters at each scale. In the Gabor filter-based method [2], simple statistics, the mean and the variance, of the Gabor transform coefficients are used as image features. In a more sophisticated modeling method in [4], a uni-component generalized Gaussian density (GCD) is used. While a uni-component GCD provides computational ease, wavelet domain distributions of images are known to be better described by mixture densities rather than single component [5].

Based on the extracted image features, a similarity measure computes the distance between the query image and each image in the database so that the top matched images to the query can be retrieved. A commonly used similarity measure is the norm-based distance between two feature vectors, such as the City-block distance and the Euclidean distance. However, it is far from optimal since the extracted image features are different from each other in nature and hardly comparable. For instance, these features often vary in global variances and cannot be compared on a same scale if the simple Euclidean distance is adopted for the similarity measure. In [4], a Kullback divergence similarity measure is proposed to compute the distance of features extracted by a uni-component generalized Gaussian density (GGD) model in the wavelet domain. While the Kullback divergence is more accurate and effective than norm-based distances for the statistical model-based feature space, the GGD is not optimal for the marginal distribution of wavelet coefficients. With more sophisticated and accurate models applied to the wavelet coefficients, the computational complexity of Kullback divergence may become significantly high.

The low-level image features concern little about high-level semantics of human perception. To bridge the gap between low-level image features and high-level semantics, an interactive mechanism known as relevance feedbacks, is

introduced [1]. Recently in [6], methods are proposed to identify low-level features that are more important to relevance feedbacks. Also, the selected low-level image features should reflect as much human perception as possible, e.g., in [7], a semantic subspace projection method is presented to select feature subspace using machine learning on human perceptions.

In this letter, we present new multiscale statistical modeling-based image features that have good reflection to human perception characteristics, and a new similarity measure of two feature vectors based on a novel Kullback divergence approach with low-computational complexity. First, we analyze image texture features in the wavelet domain and present a class of statistical models, including a Gaussian mixture model (GMM) as well as a more generic generalized Gaussian mixture model (GGMM), to help extract the wavelet domain features and compose the feature space. Note that the GMM is a special case of the GGMM. The expectation-maximization (EM) algorithm for the GGMM is derived. The model parameters are then identified by EM algorithms and employed to construct a compact indexing feature space for a CBIR system. The features by themselves reflect certain texture statistics sensitive to human vision. On the other hand, while they are low-level features, the fact that they are obtained through multiscale wavelet decomposition enable them to comply with some high-level semantic traits of the human visual perception as has been proved in many other image applications such as image compression. Therefore, the presented new features conform to human vision characteristics and are able to represent image texture contents effectively. The new method is also computationally efficient because it uses statistical model parameters as features and has a relative small dimension of indexing feature space. Compared to previous work, our work provides more accurate and flexible statistical models and algorithms for image feature extraction. In addition, we present a new similarity measure of two indexing feature vectors based on a novel Kullback divergence approach for image retrieval. Instead of directly using computational expensive Kullback divergence for GGMM, we derived a new separate Kullback divergence-based similarity measure. Our theoretical analysis shows that the new similarity measure has very low-computational cost because it has parameterized close-form, and yet still has comparable performance as the original Kullback divergence in most conditions. The Brodatz image database and some other image databases are used to evaluate the retrieval performance of the CBIR system based on our new methods. Simulation results indicate that the GMM and the GGMM-based image texture features are very effective in representing image characteristics. It is also demonstrated that for image features extracted by the statistical model such as the GMM, the similarity measure based on Kullback divergence is more effective than conventional similarity measures, such as the City-block distance and the Euclidean distance, without sacrificing the computational efficiency. It is shown that the GMM is more accurate than GGD and achieves a better retrieval performance. Overall, it is shown that the new methods outperform most other conventional methods in retrieval performance with a comparable level of computational complexity.

This letter is organized as follows. Section II introduces new statistical modeling-based feature extraction by employing the GMM and the GGMM in the wavelet domain. Section III presents a new Kullback divergence-based similarity measure for image retrieval. Section IV demonstrates some experimental results by applying a CBIR system based on our new methods to some image databases and Section V concludes the letter.

II. New Statistical Modeling-Based Feature Extraction Methods

A. The Gaussian Mixture Model (GMM) in the Wavelet Domain

The image singularities, such as edges and textures, can be represented by the wavelet coefficients [8]. Since the wavelet coefficients have a peaky, heavy-tailed marginal distribution [5] and a near zero mean, their probability density functions (PDF) can be well expressed through a multistate Gaussian mixture. Let $w_i, i = 1, \cdots, K$, represent the wavelet coefficients in a single wavelet subspace. Its PDF, $p(w_i)$, can be well modeled through a multistate Gaussian mixture: $p(w_i) = \sum_{m=1}^{M} P_m \cdot g(w_i, 0, {\sigma_m}^2), \sum_{m=1}^{M} P_m = 1$, where the states of coefficients are represented by subscript "m," and the *a priori* probabilities of the M states are represented by P_m. Note that function $g(w_i, 0, {\sigma_m}^2) = 1/(\sqrt{2\pi}\sigma_m)e^{-{w_i}^2/2{\sigma_m}^2}$, is a Gaussian PDF with distribution $N(0, {\sigma_m}^2)$ for state m.

In this peaky, heavy-tailed marginal distribution, it is observed that only a few coefficients have large values at the positions of image singularities, such as edges and textures, etc., while most others have very small values. Therefore, it is reasonable to simplify the GMM into a two-state representation as shown in (1). One state is used to describe the large coefficient distribution and the other state is used to describe the small coefficient distribution, i.e.,

$$p(w_i) = P_s \cdot g\left(w_i, 0, {\sigma_s}^2\right) + P_l \cdot g\left(w_i, 0, {\sigma_l}^2\right),\ P_s + P_l = 1 \quad (1)$$

where the state of small coefficients is represented by subscript "s" and the state of large coefficients by subscript "l." The zero mean Gaussian component $g(w_i, 0, {\sigma_s}^2)$ corresponding to the small state has a relatively small variance ${\sigma_s}^2$, capturing the peakiness around zero (small coefficients), while the component $g(w_i, 0, {\sigma_l}^2)$ corresponding to the large state has a relatively large variance ${\sigma_l}^2$, capturing the heavy tail (large coefficients).

From a perception point of view, since large coefficients indicate singularity such as image edges or visible texture patterns, the two *a priori* probabilities P_s and P_l represent the density of such singularity within an image, while the two Gaussian variances ${\sigma_s}^2$ and ${\sigma_l}^2$ represent the strength of such singularity. Therefore, they all have significant meaning in image texture content interpretation and are very good candidates of image texture features.

B. An EM Algorithm for the GMM

The GMM parameters $[P_s, P_l, {\sigma_s}^2, {\sigma_l}^2]$ can be obtained based on an EM algorithm. The E-step calculates individual

state probabilities for each wavelet coefficient, $P_{s,i}$ and $P_{l,i}$, and the M-step involves simple closed-form updates for the variances ${\sigma_s}^2$ and ${\sigma_l}^2$ and the overall state probabilities P_s and P_l:

EM Algorithm for the GMM

Initialization: Select an initial model estimate: $\Theta(0) = [P_s(0), P_l(0), {\sigma_s}^2(0), {\sigma_l}^2(0)]$, where Θ represents the GMM parameter set $[P_s, P_l, {\sigma_s}^2, {\sigma_l}^2]$. Set iteration counter $n = 0$.

E step: Calculate the state probabilities for each wavelet coefficient w_i

$$P_{m,i} = \frac{P_m(n) \cdot g(w_i, 0, {\sigma_m}^2(n))}{P_s(n) \cdot g(w_i, 0, {\sigma_x}^2(n)) + P_l(n) \cdot g(w_i, 0, {\sigma_l}^2(n))}, \quad i = 1, \cdots, K, m = s, l \tag{2}$$

where K is the total number of wavelet coefficients.

M step: Update the model parameters: $\Theta(n) \rightarrow \Theta(n+1)$

$$P_m(n+1) = \frac{1}{K}\sum_{i=1}^{K} P_{m,i}, \ {\sigma_m}^2(n+1) = \frac{\sum_{i=1}^{K} {w_i}^2 \cdot P_{m,i}}{K \cdot P_m(n+1)}, \ m = s, l. \tag{3}$$

Set $n = n + 1$. Stop if converge; Otherwise, return to **E step**.

C. The Generalized Gaussian Mixture Model (GGMM)

The GMM is a special case of the GGMM. Inspired by the uni-component GGD model [4] and the success of the GMM in the wavelet domain, we define a two-state representation of a mixture GGMM as follows:

$$p(w_i) = P_s \cdot h_s(w_i, \alpha_s, \beta) + P_l \cdot h_l(w_i, \alpha_l, \beta) \tag{4}$$

where $h(w_i, \alpha, \beta) = \beta/2\alpha\Gamma(1/\beta)e^{-(|w_i|/\alpha)^\beta}$, and $P_s + P_l = 1$. The state of small coefficients is represented by subscript "s" and the state of large coefficients by subscript "l." The generalized Gaussian component $h_s(w_i, \alpha_s, \beta)$ corresponding to the small state has a relatively small variance α_s, while the component $h_l(w_i, \alpha_l, \beta)$ corresponding to the large state has a relatively large variance α_l. Here $\Gamma(t) = \int_0^\infty u^{t-1}e^{-u}du$ is the Gamma function and β is a preset fixed exponent value. When $\beta = 1$, the GGMM becomes a Laplacian mixture model (LMM) and when $\beta = 2$, the GGMM is exactly the GMM. It is possible that for special applications, certain GGMMs with exponent values other than 2 may have a better description of wavelet coefficients than the GMM. Therefore, we will compare their retrieval performance with that of the GMM through experiments.

The EM algorithm for GGMM is nontrivial since the exponents are variables. In the following, we develop an EM algorithm to estimate the GGMM parameters $[P_s, P_l, \alpha_s, \alpha_l]$.

In each iteration (say g), the objective of the EM algorithm is to obtain an updated parameter set Θ from the current parameter set Θ^g by maximizing the following log-likelihood function Q:

$$\begin{aligned} Q(\Theta, \Theta^g) &= \sum_{m=s,l}\sum_{i=1}^{K} \log(P_m h_m(w_i|\alpha_m)) p(m|w_i, \Theta^g) \\ &= \sum_{m=s,l}\sum_{i=1}^{K} \log(P_m) p(m|w_i, \Theta^g) \\ &\quad + \sum_{m=s,l}\sum_{i=1}^{K} \log(h_m(w_i|\alpha_m)) p(m|w_i, \Theta^g) \end{aligned} \tag{5}$$

where Θ^g represents the current model parameters $[P_s^g, P_l^g, \alpha_s^g, \alpha_l^g]$ and K represents the total number of coefficients in the wavelet subspace. The two state probabilities for each wavelet coefficient w_i are represented by $p(m|w_i, \Theta^g)$, which can be easily calculated using Bayes rule

$$p(m|w_i, \Theta^g) = \frac{P_m^g h_m(w_i|\alpha_m^g)}{p(w_i|\Theta^g)} = \frac{P_m^g h_m(w_i|\alpha_m^g)}{\sum_{k=s,l} P_k^g h_k(w_i|\alpha_k^g)}, \quad m = s, l. \tag{6}$$

To maximize the log-likelihood expression in (5), we can maximize the term containing P_m and the term containing α_m independently since they are not related.

To find the expression for P_m, the Lagrange multiplier λ is introduced with the constraint that $\sum_m P_m = 1$, and the following equation should be satisfied:

$$\frac{\partial}{\partial P_m}\left[\sum_{m=s,l}\sum_{i=1}^{K} \log(P_m) p(m|w_i, \Theta^g) + \lambda\left(\sum_m P_m - 1\right)\right] = 0. \tag{7}$$

As a result, we obtain

$$P_m = \frac{1}{K}\sum_{i=1}^{K} p(m|w_i, \Theta^g). \tag{8}$$

To find an analytical expression for α_m, we input the generalized Gaussian function $h(w_i, \alpha, \beta)$ as given in (4) into the right part of (5) and obtain

$$\begin{aligned} &\sum_{m=s,l}\sum_{i=1}^{K} \log(h_m(w_i|\alpha_m)) p(m|w_i, \Theta^g) \\ &= \sum_{m=s,l}\sum_{i=1}^{K} \left(\log\beta - \log 2\alpha_m - \log\Gamma\left(\frac{1}{\beta}\right) - \left(\frac{|w_i|}{\alpha_m}\right)^\beta\right) \\ &\quad p(m|w_i, \Theta^g). \end{aligned} \tag{9}$$

Taking the derivative of (9) with respect to α_m and setting it to zero, we can solve α_m

$$\alpha_m = \left(\frac{\sum_{i=1}^{K} \beta|w_i|^\beta p(m|w_i, \Theta^g)}{\sum_{i=1}^{K} p(m|w_i, \Theta^g)}\right)^{\frac{1}{\beta}}. \tag{10}$$

Using (6), (8), and (10), the update of model parameters in the EM algorithm is an iterative procedure until a final converged set of parameters are found. In each iteration, the EM algorithm has two steps: the E-step and the M-step. The E-step calculates the individual state probabilities for each wavelet coefficient $p(m|w_i, \Theta^g)$ and the M-step involves the updates for the model parameters $[P_s, P_l, \alpha_s, \alpha_l]$. The complete EM algorithm for the GGMM with a preset β is described as follows.

EM Algorithm for the GGMM

Initialization: Select an initial model estimate: $\Theta(0) = [P_s(0), P_l(0), \alpha_s(0), \alpha_l(0)]$, Set iteration counter $n = 0$.

E step: Calculate the state probabilities for each wavelet coefficient $p(m|w_i, \Theta(n))$

$$p(m|w_i, \Theta(n)) = \frac{P_m(n) \cdot h_m(w_i, \alpha_m(n), \beta)}{P_s(n) \cdot h_s(w_i, \alpha_s(n), \beta) + P_l(n) \cdot h_l(w_i, \alpha_l(n), \beta)}, \quad i = 1, \cdots, K, \ m = s, l. \tag{11}$$

M step: Update the model parameters: $\Theta(n) \rightarrow \Theta(n+1)$,

$$P_m(n+1) = \frac{1}{K} \sum_{i=1}^{K} p(m|w_i, \Theta(n)), \ \alpha_m(n+1) = \left(\frac{\beta \sum_{i=1}^{K} |w_i|^\beta \cdot p(m|w_i, \Theta(n))}{K \cdot P_m(n+1)} \right)^{1/\beta}, m = s, l. \tag{12}$$

Set $n = n + 1$. Stop if converge; Otherwise, return to **E step**.

D. The Indexing Feature Space for Image Retrieval

The GMM and the related EM algorithm need to be applied to each wavelet subspace to extract the texture features. The indexing feature space integrates the GMM parameters extracted from all decomposed subspaces of j scales, with the following form of representation:

$$F = [W_{1H}, W_{1V}, W_{1D}, \cdots, W_{jH}, W_{jV}, W_{jD}] \tag{13}$$

where W represents the GMM parameter set $[P_s, P_l, {\sigma_s}^2, {\sigma_l}^2]$ of a single wavelet subspace. Subscripts H, V, and D represent the three different directions (horizontal, vertical, and diagonal) of the wavelet transform at each scale and subscript j represents the number of scales the image is decomposed to. For the GGMM, W representing the GGMM parameter set $[P_s, P_l, \alpha_s, \alpha_l]$.

III. A New Similarity Measure Based on the Kullback Divergence

Euclidean distance is a commonly used similarity measure but it is often ineffective because features are so different. The normalized Euclidean distance [9] is one way to alleviate the problem. However, it is still arbitrary since the features in the feature vector represent measurements of different characteristics. With a new wavelet domain parametric statistical model, we use Kullback divergence to measure the similarity of distributions represented by our feature vectors.

A. A Novel Kullback Divergence-Based Similarity Measure

Suppose that there are two PDFs, $p(x)$ and $q(x)$, for wavelet coefficients in a wavelet subspace. Their Kullback divergence is calculated as $d(p(x), q(x)) = \int p(x) \ln p(x)/q(x) dx$. To measure the similarity between an image and the query, we need to compute the Kullback divergence between their Gaussian mixture distributions for each decomposed wavelet subspace and then sum them up to get an overall distance. Using (1), we obtain the Kullback divergence for a single wavelet subspace for the GMM

$$d(p_1(x), p_2(x)) = \int \left(P_{s_1} g(x, 0, {\sigma_{s_1}}^2) + P_{l_1} g(x, 0, {\sigma_{l_1}}^2) \right) \cdot \ln \left(\frac{P_{s_1} g(x, 0, {\sigma_{s_1}}^2) + P_{l_1} g(x, 0, {\sigma_{l_1}}^2)}{P_{s_2} g(x, 0, {\sigma_{s_2}}^2) + P_{l_2} g(x, 0, {\sigma_{l_2}}^2)} \right) dx \tag{14}$$

where $p_1(x)$ is the Gaussian mixture distribution of the image (used to compare to the query), and $p_2(x)$ is the Gaussian mixture distribution of the query.

However, there is no closed form for the Kullback divergence given in (14) and it can only be numerically calculated [10]. Its computational complexity is so high that the approach is not practical. It is observed that by dividing the Gaussian mixture distribution into two separate Gaussian distributions, a simple closed form of the Kullback divergence for each separate Gaussian distribution can be easily computed. Therefore, we present a new Kullback divergence-based similarity measure for the GMM as follows:

$$d_k(p_1(x), p_2(x)) \triangleq F_s + F_l \tag{15}$$

$$F_m \triangleq \int P_{m_1} g(x, 0, {\sigma_{m_1}}^2) \ln \frac{P_{m_1} g(x, 0, {\sigma_{m_1}}^2)}{P_{m_2} g(x, 0, {\sigma_{m_2}}^2)} dx, m = s, l. \tag{16}$$

The two separate Kullback divergences F_s and F_l have simple closed forms: $F_m = P_{m_1} \ln (P_{m_1}\sigma_{m_2}/P_{m_2}\sigma_{m_1}) + P_{m_1}/2(({\sigma_{m_1}}^2/{\sigma_{m_2}}^2) - 1), m = s, l$. Similarly, for the GGMM, the similarity measure becomes

$$F_m \triangleq \int P_{m_1} h(x, \alpha_{m_1}, \beta) \ln \frac{P_{m_1} h(x, \alpha_{m_1}, \beta)}{P_{m_2} h(x, \alpha_{m_2}, \beta)} dx = P_{m_1} \ln \left(\frac{P_{m_1}\alpha_{m_2}}{P_{c_2}\alpha_{m_1}} \right) + \frac{P_{m_1}}{\beta} \left(\frac{\alpha_{m_1}^\beta}{\alpha_{m_2}^\beta} - 1 \right), m = s, l. \tag{17}$$

Therefore, the similarity measure based on the proposed separate Kullback divergence approach can be calculated very efficiently using GMM or GGMM parameters. In fact, its computational complexity is retained at the same level as other conventional similarity measures using Minkowski distances.

B. Analysis and Discussion of the Separate Kullback Approach

The separate Kullback divergence can approximate the complete Kullback divergence in distance computation. The closeness of the two divergences can be justified through the following mathematical analysis.

Assume there are two generalized Gaussian mixture distributions

$$p_1(x) = P_{s_1} h_{s_1}(x) + P_{l_1} h_{l_1}(x) = A + B, \quad p_2(x) = P_{s_2} h_{s_2}(x) + P_{l_2} h_{l_2}(x) = C + D. \tag{18}$$

For simplicity, $A, B, C,$ and D are used to denote the four Gaussian components. The complete Kullback divergence (14) can be rewritten as

$$d(p_1(x), p_2(x)) = \int A \ln \frac{A + B}{C + D} dx + \int B \ln \frac{A + B}{C + D} dx \tag{19}$$

while the separate Kullback divergence (15) becomes

$$d_k(p_1(x),\, p_2(x)) = \int A \ln \frac{A}{C} dx + \int B \ln \frac{B}{D} dx. \qquad (20)$$

If we assume that the two Gaussian mixture distributions are similar to each other, i.e., $A/C \approx B/D$, it can be further derived that: $(A + B)/(C + D) \approx A/C$, $(A + B)/(C + D) \approx B/D$. If we substitute the above into (19) and compare it with (20), we will get the conclusion: $d_k(p_1(x), p_2(x)) \approx d(p_1(x), p_2(x))$. That means the separate Kullback divergence is able to approximate the complete Kullback divergence when the two Gaussian mixture distributions are similar to each other. The detailed analysis is shown in Appendix. Since any two images from the same class will have relatively similar Gaussian mixture distributions, their distance can be computed accurately using the separate Kullback divergence instead of the complete Kullback divergence.

IV. Experimental Results

A CBIR system is used to test our methods. The features of each image in the image database are extracted by the GMM or the GGMM and the composed indexing feature vectors are stored in the feature database. The similarity measure between the query and the images in the database is based on a proposed Kullback divergence approach.

The Brodatz image database, along with some other image databases, is used to demonstrate the effectiveness of the proposed image retrieval method. The Brodatz database consists of 1856 images in 116 different classes, with each class containing 16 similar images. Given a query from any class, the ideal condition is that all 16 images in the same class as the query are retrieved in the top 16 matches. The retrieval performance of the presented new method is evaluated by the overall retrieval rate which is defined as the average percentage of images belonging to the same class as the query in the top 16 matched images [2]. Since the first round *initial retrieval rate* reflects objectively the effectiveness of the applied feature extraction technique and the similarity measure, the *initial retrieval rate* is used to benchmark the retrieval performance and no relevance feedback is applied in this experiment.

One observation from the experiment is that the retrieval performance is affected by the number of wavelet scales used to extract the image features. With more levels of wavelet decomposition, the retrieval performance tends to get better because more image texture information can be represented by the indexing feature space. However, the feature number will increase accordingly, lowering the computational efficiency of indexing and retrieval. The total number of features in the indexing feature space is $6 \times 4 = 24$ for two scales, $9 \times 4 = 36$, and $12 \times 4 = 48$ for four scales.

A retrieval example is given in Fig. 1 to illustrate the effectiveness of the new method. The correctly retrieved images are marked with check boxes. Further results also indicate that the extracted features are appropriate and effective. Besides, the retrieval rate can be further enhanced if a proper relevance feedback mechanism is introduced in the subsequent rounds of retrieval.

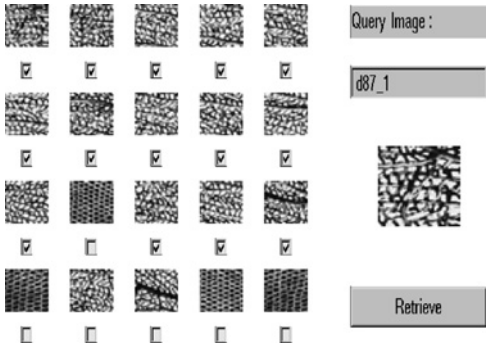

Fig. 1. Initial Retrieval of a Query from Class D87. The correctly retrieved images are marked with check boxes.

TABLE I
Initial Retrieval Rates of the GMM-Based CBIR System Using Different Similarity Measures and Wavelet Decomposition Scales

Scales Used	Type of the Similarity Measure		
	NED	CKD	SKD
2	69.12%	71.63%	71.87%
3	73.72%	75.68%	75.50%
4	51.58%	74.42%	54.97%

TABLE II
GMM Compared With the GGMM (Separate Kullback Similarity Measure is Used)

GMM	GGMM				
$\beta = 2.0$	$\beta = 1.0$	$\beta = 1.5$	$\beta = 2.5$	$\beta = 3.0$	$\beta = 3.5$
75.50%	69.15%	73.75%	75.80%	73.63%	68.44%

To demonstrate the effectiveness of the new similarity measure, the initial retrieval results using three different kinds of similarity measures [normalized Euclidean distance (NED), complete Kullback divergence (CKD), and separate Kullback divergence (SKD)] are compared in Table I. In the experiments, two, three and four scales of wavelet decomposition are used for feature extraction, respectively.

It can be seen that with three scales of decomposition, the retrieval rate is enhanced by 4%, and the number of features also increases by nearly 50%, compared to that of two scales.

Also from Table I, the similarity measure based on the Kullback divergence achieves a better retrieval rate than the normalized Euclidean distance approach, because the GMM is a statistical model that uses PDF to describe image texture features. The closeness of two images, represented by the similarity of their PDFs, can be more accurately measured using the Kullback divergence than using other approaches. It can be concluded from Table I that the presented new similarity measure based on the separate Kullback calculation has the same effectiveness as that based on the complete Kullback computation. Meanwhile, the new similarity measure has much lower computational complexity and is practical in a CBIR system.

Note that in scale 4, the transformed image size becomes only 8×8 from the original size of 128×128. It appears

TABLE III
GMM (WITH THE SEPARATE KULLBACK SIMILARITY MEASURE) COMPARED WITH OTHER TRADITIONAL METHODS

GMM & Kullback		GGD & Kullback		PWT	Gabor
Two Scales	Three Scales	Two Scales	Three Scales		
71.87%	75.50%	49.82%	56.26%	68.70%	74.37%
24	36	12	18	24	48

Last row represents the number of features.

that an 8×8 image size is too small to have reliable mixture model estimation. We observed that the variance changes significantly. That leads to the performance degradation since the variance features and mean features are less comparable due to unstable estimations in scale 4. Also the separate Kullback method cannot well approximate the continuous Kullback method because the distributions at scale 4 lose the similarity we assumed due to instability of the parameter estimation. On the positive side, we do see that the Kullback-based method shows stability in terms of similarity since we are comparing distributions rather than different types of features. We therefore suggest that, when selecting the number of scales, the size of wavelet subspaces should be larger than 8×8 such that the statistical parameter estimation is reliable. The following analysis will focus on three-scale feature space.

To examine the impact of different exponent values in the GGMM on the retrieval performance, different GGMM cases with the exponent value β ranging from 1.0 to 3.5 are compared. The model becomes a LMM if $\beta = 1$ and a GMM if $\beta = 2$. Some researchers studied wavelet coefficients using a uni-component GGD model and suggested that the appropriate exponent is around 0.5 for low or middle frequency subspaces [11]. In fact, the appropriate exponent value selected for retrieval is dependent on the model applied, the subspaces decomposed and the image database tested. In our case, as far as the GGMM is adopted and some high-frequency subspaces are used for feature extraction, the appropriate exponent value for retrieval is around 2.0. In our experiments, image features from three wavelet scales are extracted and the similarity measure is based on the separate Kullback divergence. It can be concluded from Table II that the retrieval performance achieves the highest when β takes a value between 2.0 and 2.5, which means that the selection of the GMM is appropriate and near optimal among all GGMM cases when the experiment is conducted on the Brodatz image database.

The new CBIR method with GMM and the separate Kullback-based similarity measure is further compared to other traditional methods, such as the pyramid wavelet transform (PWT)-based and the Gabor filter-based methods. Like the GMM-based method, these methods also extract image texture features from the compressed domain. We also compare our method with the GGD model-based Kullback divergence approach presented in [4]. The GGD is a statistical model in the wavelet domain, but is a uni-model with only one generalized Gaussian component. In our experiment, we test and compare all methods using the entire Brodatz image database (1856 images in 116 classes). Table III shows comparison results. The new method achieves a higher retrieval rate than the PWT and Gabor methods, with equal or fewer features in the feature vector. When compared with the GGD model-based Kullback approach, the new method has more features but also a much higher retrieval rate. That is largely due to the fact that the GMM describes the marginal distribution of wavelet coefficients more accurately than GGD.

TABLE IV
AVERAGE PRECISION RATE IN THE INITIAL RETRIEVAL FOR THE COREL DATABASE USING THE GMM

$N = 10$	$N = 20$	$N = 30$	$N = 40$	$N = 50$
82.47%	74.83%	69.59%	65.51%	62.00%

The presented statistical modeling in the wavelet domain captures the distribution of image singularities. The distributions of singularities, including edges, shapes, smoothness as well as texture, etc., are also important features for general images. Actually, as shown in [8], singularity features alone may represent all image information. Therefore, the presented CBIR method is effective not only for texture images, but also for other nontexture image databases. More experiments are performed on two additional databases. One database consists of 171 natural scene images such as landscapes, architectures, animals, etc. The other one is the Corel database that consists of 600 images in six classes, with 100 similar images in each class. For the Corel database, an average precision rate is used to evaluate the retrieval performance. The average precision rate measured by the percentage of correctly retrieved images in the top N matched positions, is used to evaluate the retrieval performance, as shown in Table IV. Results indicate that the proposed method is effective in retrieving relevant Corel images.

V. CONCLUSION

In this letter, a new statistical modeling-based image feature extraction and a novel Kullback divergence similarity measure have been developed. The GMM and GGMM are presented to help extract new image features associated with image singularities, such as textures, etc., in the wavelet domain and construct the feature space. Specifically, the related EM algorithm for the GGMM is derived to identify model parameters. In addition, a new efficient separate Kullback divergence-based similarity measure is developed for the feature space based on the GMM or the GGMM. Compared with conventional norm-based distances (City-block or Euclidean), the Kullback divergence is more appropriate and efficient in the similarity

measure of features extracted by statistical models. Simulation results indicate that, compared with Euclidean distance-based similarity measures, the new Kullback divergence-based similarity measures achieve a higher retrieval rate with the same level of computational complexity in a CBIR system. It is shown that the CBIR system using the new statistical modeling-based features and the new Kullback divergence-based similarity measure outperforms many other existing methods for texture image retrieval. Additional experiments show that the new methods are also effective in retrieving general natural images and have potential to a broad range of applications. Also, it is worth to point out that our two-component mixture model, though effective in creating a compact feature space, have a relative simple structure. More mixture components could generate better discriminative features for image classification, clustering, and retrieval depending on types of images. Our parameter estimation methodology and similarity measure can be generalized to multiple components. The performance of such multicomponent mixture model is worth further investigation in the future.

APPENDIX

When the two Gaussian mixture distributions are different from each other, it is worth of study that how the separate Kullback divergence and the complete Kullback divergence will respond to this difference respectively. To simplify the problem, we assume that $p_1(x)$ and $p_2(x)$ have the same Gaussian components but slightly different state probabilities $p_1(x) = P_s h_s(x) + P_l h_l(x)$, and $p_2(x) = (P_s + \delta)h_s(x) + (P_l - \delta)h_l(x)$, where δ is a very small quantum compared with P_s and P_l. The difference between the complete Kullback divergence and separate Kullback divergence is given by

$$\begin{aligned}\Delta &= d(p_1(x), p_2(x)) - d_k(p_1(x), p_2(x)) \\ &= \int P_s h_s(x) \left(\ln \frac{P_s h_s(x) + P_l h_l(x)}{(P_s + \delta)h_s(x) + (P_l - \delta)h_l(x)} \right. \\ &\qquad \left. - \ln \frac{P_s h_s(x)}{(P_s + \delta)h_s(x)} \right) dx \\ &\qquad + \int P_l h_l(x) \left(\ln \frac{P_s h_s(x) + P_l h_l(x)}{(P_s + \delta)h_s(x) + (P_l - \delta)h_l(x)} \right. \\ &\qquad \left. - \ln \frac{P_l h_l(x)}{(P_l - \delta)h_l(x)} \right) dx. \end{aligned} \tag{21}$$

Applying Taylor's series, since

$$\begin{aligned}&\ln \frac{P_s h_s(x) + P_l h_l(x)}{(P_s + \delta)h_s(x) + (P_l - \delta)h_l(x)} - \ln \frac{P_s h_s(x)}{(P_s + \delta)h_s(x)} \\ &= \ln \left(1 + \frac{\delta}{P_s}\right) - \ln \left(1 + \frac{\delta(h_s(x) - h_l(x))}{P_s h_s(x) + P_l h_l(x)}\right) \\ &= \frac{\delta}{P_s} - \frac{\delta(h_s(x) - h_l(x))}{P_s h_s(x) + P_l h_l(x)} + o(\delta) \\ &= \frac{\delta(P_s + P_l)h_l(x)}{P_s(P_s h_s(x) + P_l h_l(x))} + o(\delta) \end{aligned} \tag{22}$$

and similarly

$$\begin{aligned}&\ln \frac{P_s h_s(x) + P_l h_l(x)}{(P_s + \delta)h_s(x) + (P_l - \delta)h_l(x)} - \ln \frac{P_l h_l(x)}{(P_l - \delta)h_l(x)} \\ &= \frac{-\delta(P_s + P_l)h_s(x)}{P_l(P_s h_s(x) + P_l h_l(x))} + o(\delta) \end{aligned} \tag{23}$$

we substitute (22) and (23) into (21) and obtain

$$\begin{aligned}\Delta &= \int P_s h_s(x) \frac{\delta(P_s + P_l)h_l(x)}{P_s(P_s h_s(x) + P_l h_l(x))} dx \\ &\quad + \int P_l h_l(x) \frac{-\delta(P_s + P_l)h_s(x)}{P_l(P_s h_s(x) + P_l h_l(x))} dx + o(\delta) \\ &= \int \frac{\delta(P_s + P_l)h_l(x)h_s(x)}{P_s h_s(x) + P_l h_l(x)} dx \\ &\quad - \int \frac{\delta(P_s + P_l)h_s(x)h_l(x)}{P_s h_s(x) + P_l h_l(x)} dx + o(\delta) \\ &= o(\delta) \approx 0. \end{aligned} \tag{24}$$

The result reveals that the separate Kullback divergence is exactly the same as the complete Kullback divergence if the two generalized Gaussian mixture distributions under comparison are slightly different in parameters. Therefore, the separate Kullback divergence can be regarded as a very good approximation to its complete counterpart. Note that since the GMM is a special case of the GGMM, the above analysis and results also hold for the GMM.

REFERENCES

[1] Y. Rui, T. S. Huang, and S. F. Chang, "Image retrieval: Past, present, and future," *J. Visual Commun. Image Representation*, vol. 10, pp. 1–23, 1999.

[2] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.

[3] 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.

[4] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Process.*, vol. 11, no. 2, pp. 146–158, Feb. 2002.

[5] J. Romberg, H. Choi, and R. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1056–1068, Jul. 2001.

[6] D. Tao, X. Tang, and X. Li, "Which components are important for interactive image searching?" *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 3–11, Jan. 2008.

[7] J. Yu and Q. Tian, "Semantic subspace projection and its applications in image retrieval," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 4, pp. 544–548, Apr. 2008.

[8] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 617–643, Mar. 1992.

[9] H. Yuan, X.-P. Zhang, and L. Guan, "Content-based image retrieval using a Gaussian mixture model in the wavelet domain," *Soc. Photographic Instrum. Eng. Visual Commun. Image Process.*, vol. 5150. no. 3, pp. 422–429, Jul. 2003.

[10] N. Vasconcelos, "On the efficient evaluation of probabilistic similarity functions for image retrieval," *IEEE Trans. Inform. Theory*, vol. 50, no. 7, pp. 1482–96, Jul. 2004.

[11] Q. Cheng and T. S. Huang, "Robust optimum detection of transform domain multiplicative watermarks," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 906–924, Apr. 2003.