A STATISTICAL APPROCH FOR IMAGE FEATURE EXTRACTION IN THE WAVELET DOMAIN

Hua Yuan, Xiao-Ping Zhang, Ling Guan
Department of Electrical & Computer Engineering
Ryerson University, Toronto, Ontario M5B 2K3
hyuan, xzhang, lguan@ee.ryerson.ca

ABSTRACT

In this paper, a new image feature extraction method based on the statistical analysis in the wavelet domain is developed for content-based image retrieval (CBIR). A two component Gaussian mixture model is developed to describe the statitistical characteristics of images in the wavelet domain. The model parameters are obtained by an EM (Expectation-Maximization) algorithm and then employed to construct the indexing feature space for CBIR. The new method is applied on the Brodatz image database to demonstrate its performance. The preliminary experimental results indicate that the composed indexing feature space through the statistical approach is very effective in representing image features and provides a high retrieval rate in CBIR. Compared with other CBIR feature extraction methods, the new method achieves comparable retrieval performance with less number of features in the feature space, which means the new method is more computationally efficient.

***Keywords:* Content-based image retrieval, wavelet transforms, feature extraction, EM algorithm**

1. INTRODUCTION

Content-based image retrieval (CBIR) has been very active in recent years with the advancement of the multimedia storage technologies and the development of the World Wide Web. In a CBIR system, low-level image features such as color, texture, and shape are extracted so that users can index and retrieve images in the large image databases. The performance of a CBIR system depends on the effectiveness of the feature extraction method and the composed feature space. Based on low-level feature indexing and matching, a lot of feature extraction tools and image retrieval systems have been developed, e.g. the USC system, the Phonebook system, the Mars system, etc. In this paper, we analyze image features in the wavelet domain and propose a statistical model to extract features and compose the feature space. The extracted features are able to represent image contents (edges, textures) effectively and the composed feature space is compact, which means less computational complexity involved.

This paper is organized as follows. Section 2 reviews the basic concepts and practices of a CBIR system. In section 3, the Gaussian mixture statistical model and related EM algorithm are presented for image feature extraction in the wavelet domain. Section 4 demonstrates some experimental results on the Brotadz image database. Finally, section 5 draws some conclusions.

2. CONTENT-BASED IMAGE RETRIEVAL

Figure 1 shows the architecture of a typical CBIR system. For each image in the image database, its features are extracted and the obtained feature space (or vector) is stored in the feature database. When a query image comes in, its feature space will be compared with those in the feature database one by one and the first K similar images with the smallest feature distance will be retrieved.

Because there is always a gap between the high-level semantics of human perception and the low-level features, the K retrieved images will be passed to the user for discrimination. The user will determine which image is relevant to the query image and which is not among these K images depending on his visual perception. Based on the discrimination result, the feature space of the query image will then be updated by some query models and applied in the next round of feature comparison. Through this iterative process, the gap between human perception and low-level features will be compensated and the retrieval rate will be improved.

CCECE 2003 – CCGEI 2003, Montreal, May/Mai 2003
0-7803-7781-8/03/$17.00 © 2003 IEEE

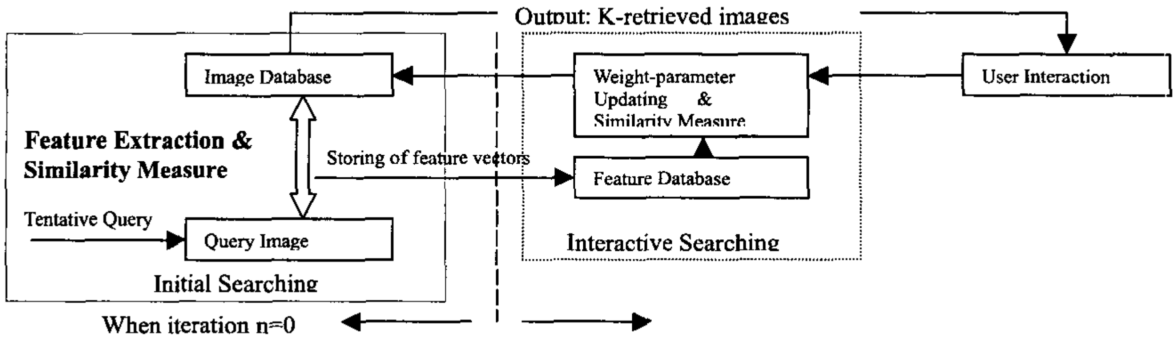


Fig. 1. Architecture for interactive content-based image retrieval

3. FEATURE EXTRACTION BY THE MIXED GAUSSIAN MODEL IN THE WAVELET DOMAIN

The proposed feature extraction method is based on a multi-scale statistical modeling of images in the wavelet domain. The obtained model parameters of the wavelet subspaces at multiple scales are employed to construct the indexing feature space. The feature extraction process contains three steps.

In the first step, a Gaussian mixture statistical model is developed to describe the distribution characteristics of wavelet coefficients in the wavelet subspaces. The wavelet coefficients record image texture and edge information at different wavelet scales and have a peaky, heavy-tailed marginal distribution [1]. Only a few coefficients have large values at the positions where edges occur, while most others have small values. This statistical characteristic can be expressed by using a two component Gaussian mixture:

$$p(w_i) = p_s \cdot g(w_i, 0, \sigma_s^2) + p_l \cdot g(w_i, 0, \sigma_l^2) \qquad (1)$$

$$p_s + p_l = 1, \qquad (2)$$

where the class of small coefficients is represented by subscript "*s*" and the class of large coefficients by subscript "*l*". The *a priori* probabilities of the two classes are represented by p_s and p_l, respectively. The Gaussian component $g(w_i, 0, \sigma_s^2)$ corresponding to the small coefficients has a relatively small variance σ_s^2, capturing the peakiness around zero, while the component $g(w_i, 0, \sigma_l^2)$ corresponding to the large state has a relatively large variance σ_l^2, capturing the heavy tails. Note w_i, $i=1,\ldots,K$, represents the wavelet coefficient.

In the second step, we apply an EM algorithm [2] on the Gaussian mixture model to obtain the model parameters $[p_s, p_l, \sigma_s^2, \sigma_l^2]$ pertaining to a certain wavelet subspace. The EM algorithm is iterative and consists of two steps for each iterative cycle. For the *n*-th iterative cycle, the E step calculates the individual hidden state probabilities for each wavelet coefficient:

$$p_{si}(n) = \frac{p_s(n) g(w_i, 0, \sigma_s^2(n))}{p_s(n) g(w_i, 0, \sigma_s^2(n)) + p_l(n) g(w_i, 0, \sigma_l^2(n))} \qquad (3)$$

$$p_{li}(n) = \frac{p_l(n) g(w_i, 0, \sigma_l^2(n))}{p_s(n) g(w_i, 0, \sigma_s^2(n)) + p_l(n) g(w_i, 0, \sigma_l^2(n))} . \qquad (4)$$

The M step involves simple closed-form updates for the variances $[\sigma_s^2, \sigma_l^2]$ and the overall *a priori* probabilities $[p_s, p_l]$:

$$p_s(n+1) = \frac{1}{K}\sum_{i=1}^{K} p_{si}(n),\ p_l(n+1) = \frac{1}{K}\sum_{i=1}^{K} p_{li}(n) \qquad (5)$$

$$\sigma_s^2(n+1) = \frac{\sum_{i=1}^{K} w_i^2 p_{si}(n)}{K p_s(n+1)},\ \sigma_l^2(n+1) = \frac{\sum_{i=1}^{K} w_i^2 p_{li}(n)}{K p_l(n+1)} . \qquad (6)$$

In the third step, the model parameters of different wavelet subspaces are integrated to construct the indexing feature space. As known, the 2-D wavelet transform decomposes an image into three wavelet subspaces (Horizontal, Vertical and Diagonal) at each wavelet scale. The three wavelet subspaces contain different image information. Therefore, the statistical model needs to be applied to each of them to extract the features. Besides the wavelet subspace, the 2-D wavelet transform also generate a scaling subspace at the coarsest scale. Since the scaling subspace is a low-frequency approximation to the original image, we can take the mean value of its coefficients as a stand-alone feature. Thus, the integrated feature space can be expressed as follows:

$$F = \left[W_{1H} W_{1V} W_{1D} \ldots W_{jH} W_{jV} W_{jD} S_j \right] \qquad (7)$$

where W is the model parameters $\left[p_s, p_l, \sigma_s^2, \sigma_l^2\right]$ of the wavelet subspace and S is the mean value of the coefficients in the scaling subspace. Subscripts H,V,D represent the three wavelet subspaces (Horizontal, Vertical and Diagonal) at each scale and subscript j represents the number of scales the image is decomposed.

4. EXPERIMENTAL RESULTS

The Brodatz image database is used to demonstrate the effectiveness of the proposed feature extraction method. The database consists of 1,856 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 our experiment, each image in the database is decomposed to two wavelet scales for feature extraction. Therefore it has a total of 6 wavelet subspaces and 1 scaling subspace. Since each wavelet subspace has 4 parameters (or features), the total number of features in the feature space is $6\times4+1=25$. Fig.2-Fig.5 shows the retrieval results of 4 image queries, each from a different class. Table 1 lists the retrieval results:

Table 1. Retrieval results of 4 different queries

Query Class	Correctly Retrieved Images Among Top 16	
	Initial Retrieval	Second Round Retrieval
1	14	16
37	10	15
66	9	16
40	8	15

Figure 2 and 3 shows the retrieval results of two simple queries which are from class 1 and 37 respectively. The images in these two classes are not complicated so the retrieval results are very satisfactory. Figure 4 shows some interesting staffs. The retrieved images are mainly from two classes and the images in the two classes are very similar to each other. The only big difference is at the texture colors. One is white while the other is black. To distinguish this kind of difference, a good method is to add the mean of the pixel values as a new feature in the feature space. Figure 5 gives a complicated query image which contains rich image information. The retrieval result is satisfactory as far as the complicacy of the image is concerned. In all above 4 cases, if we pick out the relevant images and perform a second round retrieval, we can get a perfect retrieval result with nearly all relevant images being retrieved. This fact explains that the features we selected are appropriate and effective.

The performance of the proposed feature extraction 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. Table 2 lists the overall retrieval rate with the method applied on the Brodatz image database:

Table 2. Overall retrieval rate

Wavelet Scales	Retrieval Rate (%)			
	R = 1	R = 2	R = 3	R = 4
2	73.12	84.8	87.6	88.7

R is the round of retrieval. In this test, each image is decomposed to only 2 wavelet scales. With more scales decomposed, the retrieval rate will be further improved.

We also compared the experimental result with other feature extraction methods, such as the PWT (Pyramid-structured Wavelet Transform) based, TWT (Tree-structured Wavelet Transform) based, and Gabor filter based methods. Table 3 shows the comparison result. The new method outperforms PWT and TWT based methods. When compared with the Gabor filter based method, the new method achieves a comparable performance with only half number of features in the feature space, which means it is more computationally efficient.

Table 3. Retrieval rate and feature space dimensions

	Gaussian	PWT	TWT	Gabor
Retr. Rate	73.12%	68.70%	69.41%	74.37%
Feature Num	25	24	104	48

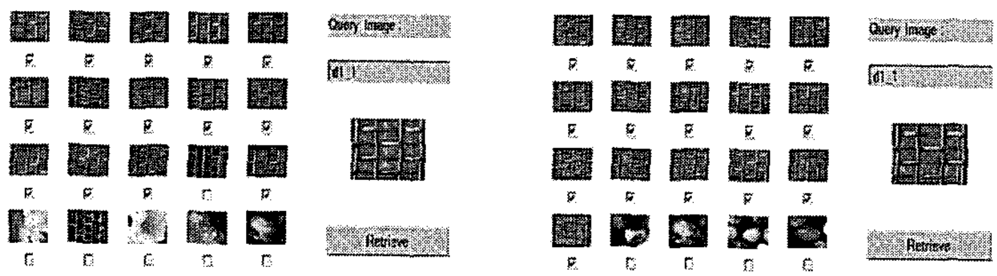
5. CONCLUSIONS

In this paper, we proposed a statistical method to extract image features in the wavelet domain. The obtained features and feature space are quite effective in representing image contents. The new method also achieves comparable or better retrieval performance with a more compact feature space, if compared with other feature extraction methods.

References

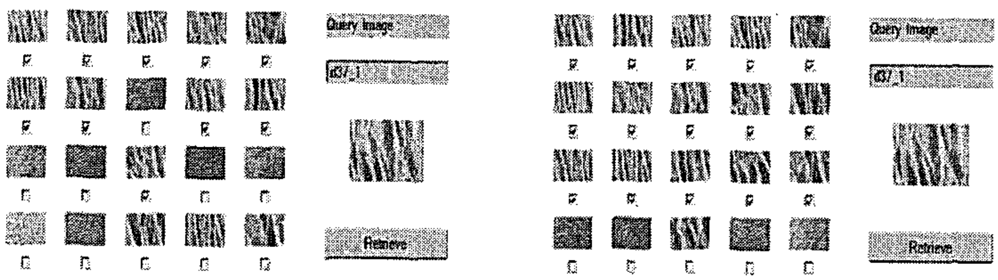
[1] Justin K. Romberg, Hyeokho Choi, Richard G.. Baraniuk, "Bayesian Tree-Structured Image Modeling Using Wavelet-Domain Hidden Markov Models", IEEE Transactions on Image Processing, Vol.10, No.7, July 2001

[2] Matthew S. Crouse, Robert D. Nowak, Richard G. Baraniuk, "Wavelet-Based Statistical Signal Processing Using Hidden Markov Models", IEEE Transactions on Signal Processing, Vol.46, No.4, April 1998



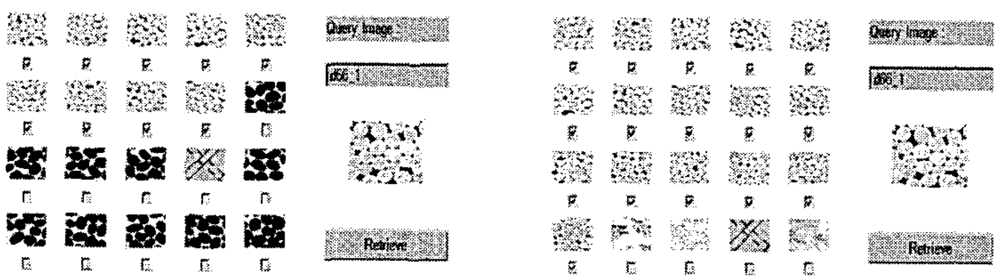
(a) Initial retrieval (b) First iteration

Fig. 2. Image query from class 1



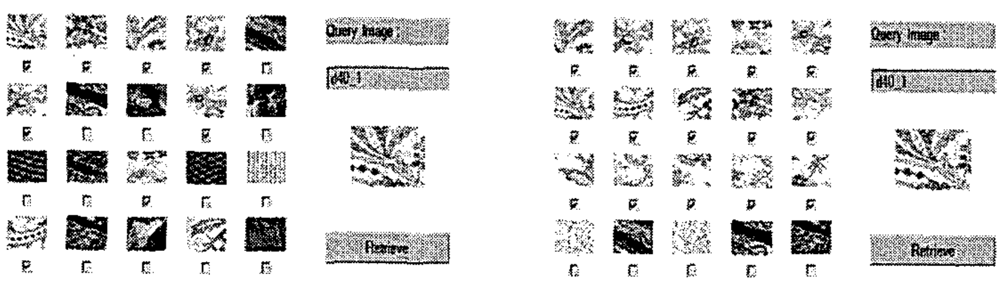
(a) Initial retrieval (b) First iteration

Fig. 3. Image query from class 37



(a) Initial retrieval (b) First iteration

Fig. 4. Image query from class 66



(a) Initial retrieval (b) First iteration

Fig. 5. Image query from class 40