Image Retrieval By Color, Texture, And Spatial Information

Xiuqi Li
NSF/FAU Multimedia Laboratory,
Department of Computer Science and Engineering,
Florida Atlantic University, Boca Raton, FL 33431

Shu-Ching Chen*
Distributed Multimedia System Laboratory, School of Computer Science,
Florida International University, Miami, FL 33199, USA

Mei-Ling Shyu
Department of Electrical and Computer Engineering, University of Miami,
Coral Gables, FL 33124-0640, USA

Borko Furht
NSF/FAU Multimedia Laboratory,
Department of Computer Science and Engineering,
Florida Atlantic University, Boca Raton, FL 33431

Abstract

A novel approach to image retrieval using color, texture and spatial information is proposed. The color information of an image is represented by the proposed color homogram, *which takes into account both the occurrence of colors of pixels and the colors of their neighboring pixels. The proposed* Fuzzy Color homogeneity, *encoded by fuzzy sets, is incorporated in the color homogram computation. The texture information is described by the mean, variance and energy of wavelet decomposition coefficients in all subbands. The spatial information is characterized by the class parameters obtained automatically from a unique unsupervised segmentation algorithm in combination with wavelet decomposition.* Multi-stage filtering *is applied to query processing to reduce the search range to speed up the query.* Color homogram filter, wavelet texture filter, *and* spatial filter *are used in sequence to eliminate images that are dissimilar to a query image in color, texture, and spatial information from the search ranges respectively. The proposed texture distance measure used in the wavelet texture filter considers the relationship between the coefficient value ranges and the decomposition levels, thus improving the retrieval performance. The final query ranking is based on the total normalized distance in color, texture, and spatial information of all images passing the three filters. The experimental results show the effectiveness of the proposed approach.*

*This research was supported in part by NSF CDA-9711582.

1 Introduction

Recently, *Content-based Image Retrieval (CBIR)* has been an active research area because of the demand for efficient image retrieval in large image databases. The traditional way of retrieving images is by manually annotated keywords (text-based). There are two main disadvantages. First, it is labor-intensive and therefore time-consuming and expensive. Secondly, the rich semantics of an image is difficult to be precisely described and different people may describe the same image in different ways [1] [2]. To overcome the drawbacks of the text-based approach, the *content-based image retrieval (CBIR)* approach that tries to retrieve images directly and automatically based on their visual contents such as color, texture, and shape was proposed [1] [2]. In a typical content-based image retrieval system, the query pattern is *query by example*, which searches the top N images similar to an example image. Before the retrieval, the visual features are extracted from all images in an image database offline. During the retrieval, the visual features of the example image are compared to those of all images in the image database and the top N images are returned as the query result [1] [2].

There have been many content-based image retrieval systems built in the literature. In the QBIC system [3], color, texture, or shape information is used in image representation. Color information is encoded using two color vectors. One is the 3D average color vector of an image in RGB, YIQ, Lab, and Munsell color spaces. The other one is a 256-bin color histogram in RGB color space. The average color vector acts as a filter to limit the expensive computations required in the color histogram computation. An improved Tamura texture representation is used as texture information. Shape information is described by area, circularity, eccentricity, major axis orientation, and moment invariants [1]. The VisualSEEK system [4] employs color represented by color set, texture described by the energies of wavelet coefficients in all subbands and spatial relationships between image regions. During the retrieval, queries based on individual features are first processed independently. Then the individual results are combined using a weighted sum of distances [1] [4]. The system in [5] screens the objectionable images through an icon filter, a graph-photo detector, a color histogram filter, a texture filter, and a shape filter.

Our approach is unique in several aspects. First, a novel *color homogram* is proposed. It extracts both the global and local color information by considering both the occurrence of colors of pixels and the colors of their neighboring pixels effectively and efficiently. The color homogram computation utilizes the proposed concept *Fuzzy Color Homogeneity*, encoded by fuzzy sets. The color space quantization is done by categorizing the pixel colors into thirteen categories. Secondly, a novel distance measure based on wavelet decomposition technique is proposed. The distance measure considers the relationship between the coefficient value scale and the decomposition level. Thirdly, the spatial information of an image is automatically extracted using a unique unsupervised segmentation algorithm. Fourthly, the wavelet technique is incorporated in generating the initial partition for the segmentation algorithm, improving the segmentation performance twice.

The rest of the paper is organized as follows. The proposed approach is presented in Section 2. In this section, an overview of the entire approach is given first. Then the color homogram filter, the wavelet texture filter, the spatial filter, and the final query presentation are discussed respectively. Section 3 shows the experimental results. The conclusion is presented in Section 4.

2 The Proposed Approach

The architecture of our approach is illustrated in Figure 1. There are three major processing units: offline feature extraction, online query processing and online query presentation. First, the color, texture, and spatial information are extracted by applying color homogram computation, wavelet texture extraction, and unsupervised segmentation algorithm to all images in the *image database* before the retrieval. All extracted visual feature information is stored in another database called *image feature database*. During retrieval, users issue a query by selecting an example image from all images in the image database. The query is processed using color homogram filter, wavelet texture filter, spatial filter, and the final query ranking algorithm sequentially. All images except the query image in the image database are first filtered using the color filter, which excludes images that are much different from the query image in color. Then the wavelet texture filter uses as input the query result of the color filter, compares the texture information of the images and discards the images whose color information is similar to that of the query image but the texture information is much different from that of the query image. Next, the spatial segmentation filter is applied to the query result in this stage. It compares the spatial information of the images in its search range to the query image and eliminates the images that are similar to the query image in color and texture but much disparate from the query image in spatial information. The last step of query processing is the final query ranking based on the total normalized distance in color, texture, and spatial information. The query presentation component displays the top N images among the final query result in the user interface.

2.1 Color Homogram Filter

In this section, we will first discuss the concept of *fuzzy color homogeneity* and the computation of *color homogram*, and then present the filtering scheme using the color homogram.

2.1.1 Color Homogram and Its Computation

Color is commonly considered as the most dominant and distinguishing visual feature in content-based image retrieval [2]. Color is perceived by humans as a combination of three color stimului: Red, Green, Blue, which forms a color space. More color spaces can be obtained by separating the luminance from the chromatic information. There are many color spaces, such as RGB, YIQ, YUV, CIE LAB, CIE LUV, and HSV, none of which is dominant in every application. RGB is the most widely used color space because of the supporting hardware devices. YIQ and YUV are used in color TV broadcasting. CIE LAB and CIE LUV are *uniform* color spaces because the Euclidean distance between two color values in these two spaces matches the human perception. Among all color spaces, only H (Hue) S (Saturation) V (Value) and its variants are *perceptual*, which makes HSV a proven color space particularly amenable to

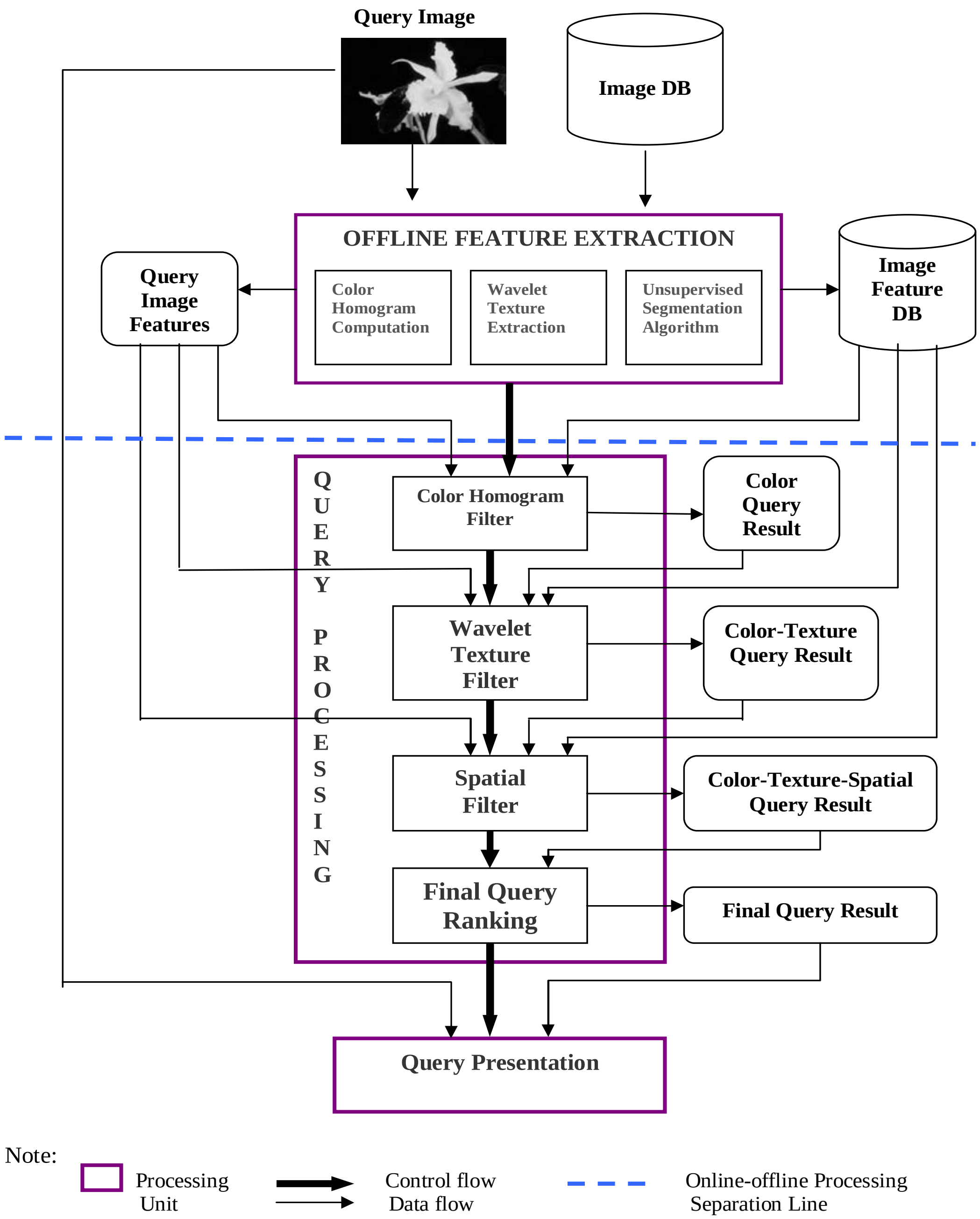

Figure 1. The architecture of the multi-filter retrieval framework.

color image analysis. Therefore, we chose the HSV color space [6] [12] [13] [16]. We also select CIE LAB space to compute the color homogram.

Many schemes have been developed to extract the color information from an image. Color histogram is the most commonly used one because it is computationally fast [2] [12] [13]. The disadvantage of color histogram is that it does not consider the local color information. To overcome this shortcoming, two typical schemes, color coherence vector and color correlogram among others, were proposed. Color coherence vector divides each histogram bin into two bins: one with coherent pixels and the other with non-coherence pixels. A pixel is considered as coherent if there is at least one neighboring pixel of the same color. Color correlogram uses co-occurrence matrices, where each element at position (i, j) represents the number of pixels of color C_j found at distance d from pixels of color C_i [2] [17].

A new approach is proposed in this paper. We include the local color information by considering the color homogeneity of each pixel with respect to its neighboring pixels. Our approach is different from the previous approaches in several aspects. First, we consider the neighboring pixels of similar colors, not just of the same colors. Second, the fuzzy set technique is utilized to encode the degree of color similarity (homogeneity). Third, the color categorization is incorporated in color space quantization.

In [15], the authors introduced fuzzy homogeneity for gray values (instead of color values) using the fuzzy set theory and applied this concept to histogram thresholding for gray-scale image segmentation. A gray-scale homogram of 256 bins was used. In [16], the gray-scale fuzzy homogeneity and 256-bin gray-scale homogram are applied to each channel of a color space to segment color images. We adapt the gray-scale fuzzy homogeneity and gray-scale homogram to color images and propose fuzzy color homogeneity and color homogram in our approach. First, we define the degree of color homogeneity of two pixels of two colors based on the Euclidean distance between the two color values in CIE LAB color space. Second, we quantize the color space using color categorization based on H S V value ranges to reduce the computation complexity, storage space, and dimension of the feature vector. The following three definitions are derived from [15] [16].

Definition 1. Let $C(i,j)$ and $C(k,l)$ be the color values in CIE LAB space of two pixels at location (i,j) and (k,l) in an image of size M-by-N. The *degree of homogeneity in terms of color (color homogeneity)* between these two pixels is denoted by $\delta(\Delta C)$, where ΔC is the Euclidean distance between $C(i,j)$ and $C(k,l)$. Let L be the maximum distance between two colors in LAB space. $\delta(\Delta C)$ is defined as follows:

$$\delta(\Delta C) = Z(\Delta C, a, b, c)$$

$$= \begin{cases} 1 & 0 \le \Delta C \le a; \\ 1 - 2\left(\frac{\Delta C - a}{c-a}\right)^2 & a \le \Delta C \le b; \\ 2\left(\frac{\Delta C - c}{c-a}\right)^2 & b \le \Delta C \le c; \\ 0 & c \le \Delta C \le L. \end{cases}$$

$Z(\Delta C, a, b, c)$ is the standard fuzzy Z function [15]. We assign 0 to a, L/2 to b, and L to c and hence $Z(\Delta C, 0, L/2, L)$ is used in our approach.

The color homogeneity between pixels and their neighbors can be computed by considering the neighbors at different angels θ separately and summing up all individual homogeneity values.

Definition 2. A fuzzy color homogeneity vector at angle θ, denoted by $h(C,\theta)$, is defined as:

$$h(C,\theta) = \sum_{P_c} \delta(\Delta C(P_c, N_\theta))$$

, where P_c refers to a pixel of color c, N_θ indicates the specific neighbor at angle θ, $\theta \in \{0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°\}$.

As we can see, the fuzzy color homogeneity vector $h(C,\theta)$ represents the sum of the color homogeneity values of all pixels of color C and their neighbors at angle θ. Considering the symmetricity, we can reduce the total number of fuzzy color homogeneity vectors to four, which are $h(C,0°)$, $h(C,45°)$, $h(C,90°)$ and $h(C,135°)$. By normalizing, summing up and averaging these four color homogeneity vectors for each color C, the color homogram can be obtained.

Definition 3. A color homogram H of an image of size M-by-N is defined as a function of color values C in LAB color space, the function value at specific color C is denoted by $H(C)$.

$$H(C) = \frac{1}{4}\left(\begin{array}{c} \frac{h(C,0°)}{M(N-1)} + \frac{h(C,45°)}{(M-1)(N-1)} + \\ \frac{h(C,90°)}{(M-1)(N)} + \frac{h(C,135°)}{(M-1)(N-1)} \end{array} \right)$$

There are so many possible color values in LAB color space. To compute a practical color homogram, we must quantize the color space. We employ the idea of color categorization for color space quantization. In [6], the author experimentally categorized the pixels based on the H, S, and V value ranges and used these color categories for the representative colors of the color segments generated from his color segmentation algorithm. Twelve categories are obtained. They are black, white, red, bright red, yellow, bright yellow, green, bright green, blue, bright blue, purple and bright purple. The Hue is divided into five main color slices and five transition color slices. Each transition color slice is considered in both adjacent main color slices. We adopted this categorization method in the color homogram computation. We disregard the difference between the bright chromatic pixels and the chromatic pixels to reduce the total number of colors. Each transition color slice is treated as

Figure 2. Query result of image 301.

a separate category instead of being combined into both adjacent main color slices. A new category "gray" is added to count all possible value ranges. Therefore, totally thirteen color categories are generated in our approach.

2.1.2 Color Homogram Fileter

The color homograms of all images in the image database are generated offline before the retrieval. During the retrieval, the color filter compares the color homogram of the query image to those of all images in the database using the L_1-*Distance*, also called *city block distance* [2]. The L_1-*Distance* between the color homograms of a query image q and an image d in the image database, denoted by $D_{chomog}^{(q,d)}$, is defined as $D_{chomog}^{(q,d)} = \sum_{k=1}^{B} \left| h_k^{(q)} - h_k^{(d)} \right|$, where h_k is the homogeneity value of the kth color category and B is the total number of color categories.

The color homogram filter eliminates the images whose color homogram distance is larger than a threshold. It effectively removes around eighty-five percent of the images in the database from its search range. This dramatically saves the computation time taken unnecessarily for feature comparison in later stages.

2.2 Wavelet Texture Filter

Wavelet transform converts a function $f(x)$ to a combination of two families of basis functions called wavelets, represented by $\psi_{mn}(x)$ and $\varphi_{mn}(x)$. These wavelets are generated by the following transation and dilation of a *mother wavelet* $\psi_{mn}(x)$ and a *father wavelet* $\varphi_{mn}(x)$ respectively [8] [9] [10]:

$$\psi_{mn}(x) = 2^{-m/2}\psi(2^{-m}x - n),$$
$$\varphi_{mn}(x) = 2^{-m/2}\varphi(2^{-m}x - n)$$

where m and n are integers, m indicates the scale, and n indicates the position. A signal $f(x)$ can therefore be represented by

$$f(x) = \sum_m \sum_n (a_{mn}\varphi_{mn}(x) + c_{mn}\psi_{mn}(x))$$

where a_{mn} and c_{mn} are the coefficients after wavelet transformation. In a one-level wavelet transformation, a low-pass filter (LPF) and a high-pass filter (HPF) are applied to a signal and the outputs of the filers are downsampled by two. The LPF generates approximation coefficients a_{mn}, while the HPF produces fluctuation coefficients c_{mn} [8] [9] [10].

With respect to an image, the first-level wavelet transform is done by first applying the one-level wavelet transform to each row of the original image and then applying one-level transform to each new column of the intermediary result matrix. This generates four sub-images (subbands) of the original image. They are the approximation sub-image A_1, the horizontal detail sub-image H_1, the vertical detail sub-image V_1, and the diagonal detail sub-image D_1. Applying the same procedure to the sub-image A_1 generates the second level wavelet transform consisting of four sub-images of A_1: A_2, H_2, V_2, and D_2. A multi-level pyramid wavelet transformation is then obtained by applying the same procedure to the approximation sub-image at lower levels [8] [9] [10].

Among all kinds of wavelets, Daubechies wavelets are proven to be good for image analysis and synthesis because of their compact support, more continuous derivatives, and zero integral of mother wavelets [10]. Therefore, Daubechies wavelets are chosen in our approach. The maximum decomposition level is three. The wavelet texture computation algorithm transforms the gray-scale versions of a query image and all images in the image database and stores three feature vectors for each image. The three vectors are mean vector, variance vector, and energy vector representing the mean, variance, and energy of coefficients in each subband respectively.

The wavelet texture filter compares the mean, variance, and energy vectors of a query image to those of an image in the query result obtained from the color filter using the following scheme. First, a Euclidean distance L_2 between the corresponding components of each feature vector at each decomposition level is computed. Three distance values per feature vector and totally nine distance values are produced. Considering that the coefficients at higher decomposition level are usually much larger than those at the lower level, the nine distance values are then normalized by dividing them by their maximum distance values. A total distance is computed for each feature (mean, variance, or energy) using the mean of the three normalized distance values at three levels. The total texture distance between a query image q and the dth image in the search range is computed as the mean of the three texture feature distances:

$$D_{texture}^{(q,d)} = \frac{1}{3}\sum_{k=1}^{3} D(q,d)^{(k)}$$

where $D(q,d)^{(k)}$ represents the distance between the kth texture feature.

The wavelet texture filter removes images whose texture distances exceed a threshold from its search range. Around half of the images in the query result obtained from the color filter are eliminated. The query result of the wavelet texture filter is sent to the spatial segmentation filter for further comparison.

2.3 Spatial Filter

The spatial information of a gray-scale image is represented by the class parameters generated by our Simultaneous Partition and Class Parameter Estimation (SPCPE) algorithm because a class is characterized by a polynomial function between the intensities of pixels in this class and their spatial locations with the class parameters being the coefficients of the function. The SPCPE algorithm partitions a gray-scale image into b *segments* belonging to s *classes*. With an initial partition, the estimation of the partition and class parameters is performed iteratively and simultaneously through Bayesian learning. [7] [11] [12] [13].

The initial partition is very important in generating the final segmentation result. The previous version of the SPCPE algorithm generates the initial partition randomly, which is unstable and not very effective. To solve this problem, the wavelet decomposition technique is utilized. The basic idea is to generate a candidate partition for each decomposition subband in a one-level wavelet decomposition based on the prominent coefficients in that subband. The one with the least cost $J(C_1, C_2, \theta_1, \theta_2)$ is chosen as the final result.

$$J(C_1, C_2, \theta_1, \theta_2) = \sum_{y_{ij} \in C_1} -ln p_1(y_{ij}|\theta_1) + \sum_{y_{ij} \in C_2} -ln p_2(y_{ij}|\theta_2)$$

, where y_{ij} represents the intensity of the pixel at (i,j), $C_1, C_2, \theta_1, \theta_2$ refer to the partitions and class parameters for class 1 and 2, and $p_1(y_{ij}|\theta_1)$ and $p_2(y_{ij}|\theta_2)$ are the probabilities of the pixel at (i,j) belonging to class 1 and class 2 given class parameters θ_1,θ_2 respectively [12] [13].

The SPCPE algorithm partitions all images in the image database and stores two vectors (one for each class) for each image in the feature database. Each vector is composed of all parameters for a class. The spatial filter computes the spatial information similarity between a query image q and an image d in its search range according to the Euclidian distance between their corresponding class parameter vectors:

$$D_{spatial}(q,d) = \sum_{m=1}^{2} \sqrt{\sum_{j=0}^{3} \left(a_{mj}^{(q)} - a_{mj}^{(d)}\right)^2},$$

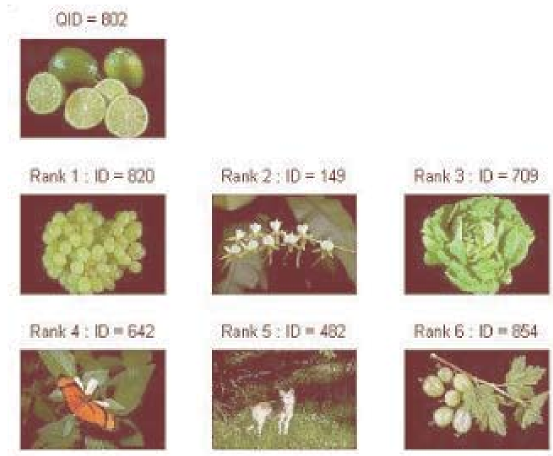

Figure 3. Query result of image 802.

where m refers to the class m and a_{mj} refers to the jth class parameter for class m. The spatial filter uses a threshold to eliminate the images that are not similar to the query image in the spatial information. It effectively eliminates approximately half of the images from its search range.

2.4 Final Query Ranking and Presentation

At the last stage, a total normalized distance including the color, texture, and spatial information is used to rank the images in the query result obtained from the spatial filter. Normalization is required because the distance values of different visual features may be at different scales. The ratio of each visual feature distance to its corresponding maximum distance is used in our approach. The total similarity distance between a query image q and the image d in the final query result is given as:

$$\begin{aligned} D_{sim}(q,d) &= D_{color}(q,d)/MAX_{d\in[1,K]}(D_{color}(q,d)) \\ &+ D_{texture}(q,d)/MAX_{d\in[1,K]}(D_{texture}(q,d)) \\ &+ D_{spatial}(q,d)/MAX_{d\in[1,K]}(D_{spatial}(q,d)), \end{aligned}$$

where K is the number of images that passed all three filters: the color homogram filter, the wavelet texture filter and the spatial filter. The images in the query result of the spatial filter are sorted based on their total normalized distances. The top six or less images that are most similar to the query image are displayed in the query user interface as the final query result.

3 Experiments

There are 2,000 images selected from Corel Gallery published in [14] in the image database in our current system. There are twenty semantic categories, each of which consists of 100 images. All images are of size 384x256.

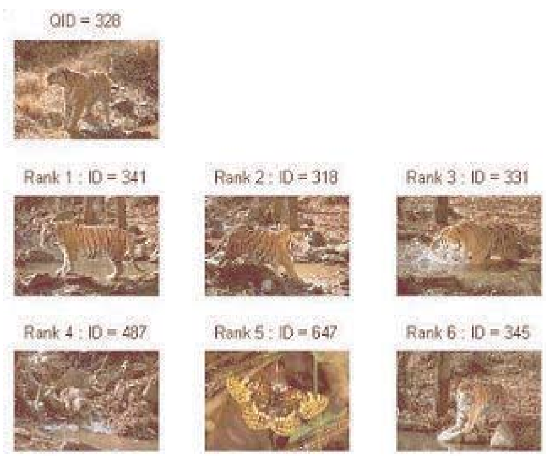

Figure 4. Query result of image 328.

Figure 2 is the query result of Image 301 that consists of the top six similar images. The image in the first row is the query image. The images in the next two rows are displayed in descending ranks. Clearly, all images have a green background and a brown object as the foreground. The foreground objects are located in similar positions in the images. Although the Rank 5 image contains some amount of white color, it is still similar to the query image in that it contains large amount of green color and large amount of brown color like the query image. The Rank 6 image is also similar to the query image though containing some amount of gray color since it contains a large amount of green color and the location and color of the foreground object are similar to those of the query image. Moreover, all the foreground objects are in the same semantic categories. They are all animals.

Figure 3 gives the query result of Image 802. Six images are returned and displayed similarly to that in Figure 2. It can be clearly seen that all images contain a black background and some amount of green color. All of them contain foreground objects located approximately at the center of each image although they have different colors. The Rank 4 image contains some amount of red color that is not present in the query image. However, it is still similar to the query image because of the black background and the green color shared with the query image.

The query result of Image 328 is shown in Figure 4. The top 6 images are returned and displayed in the same format as that in the previous figures. It is easily observed that the scenes in all images are complex. They do not have a clear background while all images except the Rank 4 image contain a foreground brown object located in the similar position in each image. All images contain some amount of black color, some amount of brown color and some amount of gray colors. The Rank 4 image is returned because it is much similar to the query image in texture and similar to the query image in color and spatial information although the color of the object in it is not brown. In addition, though the Rank 5 image contains some amount of yellow color, it is chosen because it is much similar to the query image in texture and spatial information.

4 Conclusion

In this paper, we proposed a novel image retrieval approach that consists of three components: offline feature extraction, online query processing and on-line query presentation. The proposed approach includes multi-stage filtering and therefore greatly reduces the query search ranges at different stages. This way saves a large amount of time for unnecessary feature comparison. Our approach is unique in several aspects. First, a novel fuzzy color homogeneity was proposed. Fuzzy set theory is used to map the color distance in LAB spaces to the degree of the color similarity (homogeneity) between pixels and their neighboring pixels. Secondly, a novel color homogram was proposed, which considers both the global and local color information. Thirdly, color categorization is used in color space quantization to effectively reduce computation complexity and the storage space in color homogram computation. Fourthly, a novel texture distance measure considering the relationship between the wavelet coefficient value ranges and the decomposition levels is proposed. Fifth, a unique unsupervised segmentation algorithm together with the wavelet technique is employed to automatically extract the spatial information of the images. The wavelet technique used in the initial partition generation doubles the performance of the segmentation algorithm. In the future, supervised learning algorithm will be included into the existing framework to further improve the retrieval performance.

References

[1] Yong Rui and Thomas S. Huang, “Image Retrieval: Current Techniques, Promising Directions and Open Issues,” *Journal of Visual Communication and Image Representation*, 1999.

[2] Wei-Ying Ma and HongJiang Zhang, “Content-Based Image Indexing and Retrieval,” *Handbook of Multimedia Computing*, CRC Press, 1999.

[3] M. Flicker, H. Sawhney, etc., “Query by Image and Video Content: The QBIC System,” *IEEE Computers*, Sept. 1995.

[4] John R. Smith and Shih-Fu Chang, "VisualSEEK: A Fully Automated Content-Based Image Query System," *Proc. of ACM Multimedia 96*, 1996.

[5] J. Wang, G. Wiederhold, and O. Firschein, "System for Screening Objectionable Images Using Daubechies' Wavelets and Color Histograms," *Proc. of Interactive Distributed Multimedia Systems and Telecommunication Services*, 1997.

[6] D. Androutsos, "Efficient Indexing and Retrieval of Color Image Data Using a Vector-Based Approach," *Ph.D. Dissertation*, University of Toronto, 1999.

[7] S. Sista and R. Kashyap, "Bayesian Estimation For Multiscale Image Segmentation," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, March 1999.

[8] "Wavelet Toolbox Users Guide," *The Mathworks Inc.*, 2000.

[9] Ingrid Daubechies, *Ten Lectures on Wavelets*, Capital City Press, 1992.

[10] James-Ze Wang, Gio Wiederhold, Oscar Firschein, Sha-Xin Wei, "Content-based Image Indexing and Searching Using Daubechies Wavelets," *International Journal of Digital Libraries*, P311-328, 1997.

[11] Shu-Ching Chen, Mei-Ling Shyu, Chengcui Zhang, and R.L. Kashyap, "Video Scene Change Detection Method Using Unsupervised Segmentation and Object Tracking," *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 57-60, 2001.

[12] Xiu-Qi Li, Shu-Ching Chen, Mei-Ling Shyu, and Borko Furht, "An Effective Content-Based Visual Image Retrieval System," *Proc. of 26th IEEE Computer Society International Computer Software and Applications Conference (COMPSAC)*, Oxford, 2002.

[13] *http://wang1.ist.psu.edu/docs/related*

[14] Xiuqi Li, Shu-Ching Chen, Mei-Ling Shyu, and Borko Furht, "An Efficient Multi-Filter Retrieval Framework For Large Image Databases", *Proc. of 2002 International Symposium on Information Systems and Engineering*, San Diego, 2002

[15] H.D. Cheng, C. H. Chen, H. H. Chin, and Huijuan Xu, "Fuzzy Homogeneity Approach to Multilevel Thresholding," *IEEE Transactions on Image Processing*, Vol.7, No.7, July 1998.

[16] H.D. Cheng, X. H. Jiang, Jingli Wang, "Color Image Segmentation Based on Homogram Thresholding and Region Merging," *Pattern Recognition*, 2001.

[17] J. Huang, "Color-Spatial Image Indexing and Applications," *Ph.D. Dissertation*, Cornell University, 1998.