A NOVEL HIERARCHICAL APPROACH FOR IMAGE RETRIEVAL USING COLOR AND SPATIAL INFORMATION

Xiuqi Li[1], *Shu-Ching Chen*[2†*], *Mei-Ling Shyu*[3], *Borko Furht*[1]

[1] NSF/FAU Multimedia Laboratory, Florida Atlantic University, Boca Raton, FL 33431

[2]Distributed Multimedia Information System Laboratory
School of Computer Science, Florida International University, Miami, FL 33199

[3]Department of Electrical and Computer Engineering, University of Miami,
Coral Gables, FL 33124

ABSTRACT

A novel hierarchical approach for image retrieval is proposed in this paper. First, a color label histogram is used to effectively filter out the images that are not similar to the query image in color. The color label matrix is built by categorizing the pixel colors into a set of 13 colors and by labeling each pixel based on its color category ID. Next, the class parameters of those images passing the first filter are used to identify the images similar to the query image in spatial layout. These class parameters are obtained automatically from the proposed unsupervised segmentation algorithm. Moreover, the wavelet decomposition coefficients are used to generate the initial partition for the segmentation algorithm. It doubles the segmentation performance. At the last stage, all images passing two filters are ranked based on the total normalized distance in color and spatial layout. The experiments conducted on 500 images show the effectiveness of this approach.

1. INTRODUCTION

Owing to the recent advances in hardware, managing large number of images has become ordinary. This leads to a growing interest in query images based on the content of the images. Traditionally, images are retrieved using a text-based approach. In this approach, each image is manually annotated and then the retrieval process is converted into retrieval of keywords in text descriptions of images. There are several inherent problems in such

† To whom correspondence should be directed (chens@cs.fiu.edu).

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

systems. First, manual annotation is "often subjective, inaccurate, and incomplete" [1]. Secondly, some of the image properties cannot be described using keywords. Because of these reasons, the content-based approach was developed to query images directly based on their visual attributes such as color, texture, shape, layout, object location, etc. without resort to human annotation.

In most of the content-based image retrieval systems, the goal is to find the top N images that are similar to the user query image. [1][3] Because each image has many visual features, similarity comparison based on a single feature is not enough. There have been some hierarchical approaches to content-based image retrieval combining multiple visual features. In [2], a color histogram filter and wavelet-based shape matching were utilized to query and screen objectionable images.

Our approach is different from the previous approaches in three aspects. First, a novel color label histogram is proposed. By categorizing the pixel colors into a set of 13 colors and by labeling each pixel based on the color category ID, a histogram with only 13 bins is obtained. It effectively and efficiently captures the global color information. Secondly, a unique unsupervised segmentation algorithm is applied to images to extract the information about the relationship between the pixel intensities and their spatial layout. Thirdly, wavelet decomposition is used to improve the performance of the segmentation algorithm.

The rest of the paper is organized as follows. In section 2, first, an overview of our query framework is given. Then, the color label histogram filter is presented. Next, the unsupervised segmentation parameter filter and the initial partition generation are presented. Finally, the query ranking is described. Section 3 shows the experimental results. Concluding remarks are given in Section 4.

2. HIERARCHICAL QUERY FRAMEWORK

The hierarchical query framework that we accepted is presented in Figure 1. Before the query, the color label histogram of each image in the image database is extracted offline. The results are stored for later filtering. Each image in the database is also segmented by the SPCPE (Simultaneous Partition and Class Parameter Estimation) algorithm offline. [5][8] The class parameters are generated and stored for later filtering. The query image is processed in the same way as any other image in the database. During the query, the color label histogram and the class parameters of the query image are compared to those of each image in the database. The comparison is performed in two stages. First, a color label histogram filter is used to eliminate all images that are not similar to the query image in color. All images that passed the color filter are further compared to the query image using the class parameter filter. The second filter uses the class parameters obtained from the SPCPE algorithm to filter out those images that are not similar to the query image in spatial layout. At the end, all images that pass the two filters are ranked based on the total normalized color and class parameter distance and the top six (or less) images are displayed in the user interface.

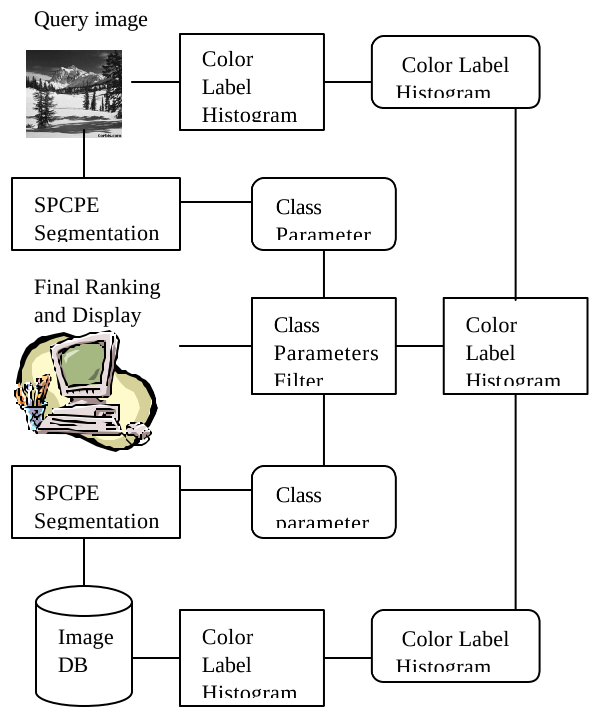

Figure 1. The Hierarchical Framework.

Next, we discuss in more details the color label histogram filter, the unsupervised segmentation parameter filter, and the final query ranking.

2.1. Color Label Histogram Filter

The image color is represented in a 3-channel color space. There are many color spaces, such as RGB, HSV, YCbCr, CIE LAB, and CIE LUV. No color space is dominant in all applications. In our proposed approach, the HSV color space is used because it is a perceptual color space. That is, the three components H (Hue), S (Saturation), and V (Value) correspond to the color attributes closely associated with the way that the human eyes perceive the colors [4][6][7][9]. Hue indicates the type of color, such as red, green, and blue, which corresponds to the dominant wavelength of a given perceived color stimulus. Saturation refers to the strength of a color. A fully saturated color contains a single wavelength. The color becomes less saturated when white light is added. Value (or intensity) is

Color Category ID	Color Category	Hue Range	Saturation Range	Value Range
1	White	Any	< 20	≥ 85
2	Black	Any	Any	< 25
3	Gray	Any	< 20	[25,85]
4	Red	[350°, 25°)		
5	Red-Yellow	[25°, 45°)		
6	Yellow	[45°, 65°)		
7	Yellow-Green	[65°, 85°)		
8	Green	[85°, 160°)		
9	Green-Blue	[160°, 180°)		
10	Blue	[180°, 270°)		
11	Blue-Purple	[270°, 290°)		
12	Purple	[290°, 330°)		
13	Purple-Red	[330°, 350°)		

Table 1. Color Category and HSV Ranges

the amount of light perceived from a given color sensation. White is perceived to be the maximum intensity and black to be the minimum intensity. The color images are often represented in RGB space for easy display. HSV values can be computed from RGB space [6][7].

Color histogram is widely used in image retrieval. It captures the global color information of an image. To compute the color histogram, the quantization of the color space is needed to reduce the computation complexity [1]. Our idea is to perform the quantization through categorization considering all three-color channels together. That is, the pixels are categorized based on their H, S, and V values, and are labeled as the IDs of the color category to which the pixels belong. Then a color label (category) histogram is computed.

In [4], the author used 12 color categories for the representative colors of color regions in an image. All categories are obtained from the experimental results based on the H, S, and V value ranges. Hue is divided into five main color slices and five transition slices. Each transition slice is included into both adjacent main color slices. In our approach, the above categories are modified for color histogram computation. We ignore the difference between bright chromatic pixels and chromatic pixels, regard each transition slice as a separate bin, and add a new category gray. Therefore, totally 13 color categories are produced. Table 1 lists each color category and the corresponding H, S, V value ranges.

Histogram comparison between the query image *q* and the *jth* image in the database is based on the L1 distance [1], which is defined as follows:

$$D_{colorlabelhist}(q,j) = \sum_{i=1}^{N} \left| X_i^{(q)} - X_i^{(j)} \right|$$

where X_i is the *ith* bin and *N* is the total number of bins.

The color label histogram filter lets images with small histogram distance (a threshold) pass. It eliminates one-twentieth of the images in the database on average.

2.2. Unsupervised Segmentation Class Parameter Filter

Given a gray-scale image, the SPCPE algorithm [5][8] partitions it into s *regions* (*classes*) that are mutually exclusively and totally inclusive. Each class consists of one or more *segment*s that are similar to each other in some sense and may not be spatially contiguous. Therfore, each image is partitioned into s classes and b segments. In the SPCPE algorithm, both the *class parameters* q and the *partitions* $\boldsymbol{C}$ are considered as random variables. The algorithm estimates $\boldsymbol{C}$ and q to be that which *maximize*s the *a-posterior probability (MAP)* of the partition variable and class parameter variable given the image data Y. Specifically, the algorithm begins with an initial partition, estimates $\boldsymbol{C}$ and q iteratively and simultaneously**,** and stops when the partition cannot be further optimized (the cost f

2.2.1. Initial Partition Generation

The SPCPE algorithm starts with an initial partition and optimizes it using the least square technique and re-labeling rule. It is found, in our experiment, that the initial partition is very important and different initial partitions lead to different segmentation results. To produce a better result, the wavelet decomposition coefficients are used in the initial partition generation.

Assume that there are 2 classes. The algorithm estimates $\boldsymbol{C}$, q to be that which has the least cost *J* [5][8].

$$\left(\hat{c},\hat{q}\right) = Arg \min_{(C,q)} J(C_1, C_2, q_1, q_2) =$$

$$Arg \min_{(C,q)} \sum_{k=1}^{2} \sum_{y_{ij} \in C_k} -2 \ln p_k(y_{ij} \mid q_k)$$

Under the assumption $p_k(y_{ij} \mid q_k) \sim$ Gauss(m_k, r_k),

$$-2 \ln p_k(y_{ij} \mid q_k) = ((y_{ij} - m_k)^2 / r_k) + \ln 2pr_k$$

Our idea is to label pixels as different classes based on the wavelet coefficient values. Images are first decomposed using wavelet at level one. Next, salient points in horizontal, vertical and diagonal subbands are extracted by thresholding. For each of the three subbands, all pixels in the original image that correspond to the salient points in that subband are labeled as one class, and the rest of the pixels are labeled as the other class. This generates three candidate initial partitions. The final initial partition is the one with the least cost J among the three candidates. Compared to the random initial partition generation, the segmentation precision is doubled with the help of wavelet technique.

2.2.2. Class Parameter Filter

The unsupervised segmentation filter applies the SPCPE algorithm to the query image and all the images in the database to generate the class parameters. Then the filter compares the class parameters of the query image to those of the images passing the color label histogram filter. It filters out on average one-fourth images whose class parameters are much different from those of the query image.

Class parameter comparison is based on the sum of the Euclidian Distances of each class parameters between the query image q and the *jth* image in the database.

$$D_{classpar}(q, j) = \sum_{k=1}^{2} \sqrt{\sum_{i=0}^{3} \left(a_{ki}^{(q)} - a_{ki}^{(j)}\right)^2}$$

where k indicates there are two classes and a_{ki} is the *ith* class parameter for class k.

2.3. Final Query Ranking

After passing the above two filters, images are sorted in descending order based on the sum of the color label histogram distance and the class parameter distance. The top six (or less) images are returned and displayed in the user interface. It is found that the class parameter distance is much larger than the color label histogram distance. Therefore, the two distances need to be normalized before the sum is computed. Normalization is implemented by dividing each color/parameter distance by the maximum color/parameter distance among all color/parameter distances between the query image and all the images in the database.

3. EXPERIMENTAL RESULTS

The experiments were conducted on 500 natural scene images, which were downloaded from *yahoo (www.yahoo.com)* and *corbis (www.corbis.com)*. They vary in color and spatial layout. Their sizes are 256x192.

The query result of Image 162 is shown in Figure 2. The image in the first row is the query image. The top three similar images and their ranks and image IDs are displayed in the next two rows. There are only three images returned. As can be seen from this figure, the result is quite good. The query image and the top three images contain two major colors: red and black. As for the spatial layout, the query image is very similar to the images with Rank 1 and Rank 2. All of them consist of a top area and a bottom area. The image with Rank 3 image is a little bit different. There are several small dark areas on the top half of the image. However, the major areas are still the top and bottom ones.

Figure 3 is the query result of Image 102. Similarly, the image in the first row is the query image. The top five images with their ranks and image IDs are displayed in the interface. There are five images returned. Clearly, all five images contain blue, black, and white colors. They are similar in color. Regarding the spatial layout, the query image is composed of three areas, the top, middle and bottom ones. Among the top 5 images returned, we can easily identify the three areas in the Rank 2, 4, 5 images. The rank 1 and 3 images also contain three approximately horizontal intensity areas located vertically next to the other because the segmentation algorithm is based on the gray-scale image, not the color image.

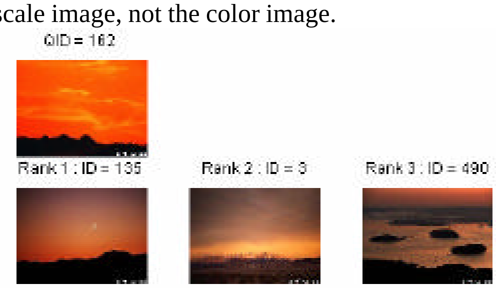

Figure 2. Query Result of Image 162.

4. CONCLUDING REMARKS

In this paper, a hierarchical framework for content-based image retrieval is proposed. A novel color label histogram, a unique unsupervised segmentation algorithm and the wavelet technique are integrated in our framework. Before the query process, the color label histogram and the class parameters are extracted from all database images offline.

During the query process, the color label histogram filter and the class parameter filter are used to filter out images that are not similar to the query image in color and spatial layout, respectively. All images passing the two filters are ranked based on the total normalized distance at the final stage. The top six (or less) images are returned in the interface. The experimental result demonstrates the effectiveness of our framework.

We plan to extend our framework to a large-scale image database. We also intend to incorporate the segmentation algorithm with more classes into our framework to improve the performance of queries on more complicated natural scene images.

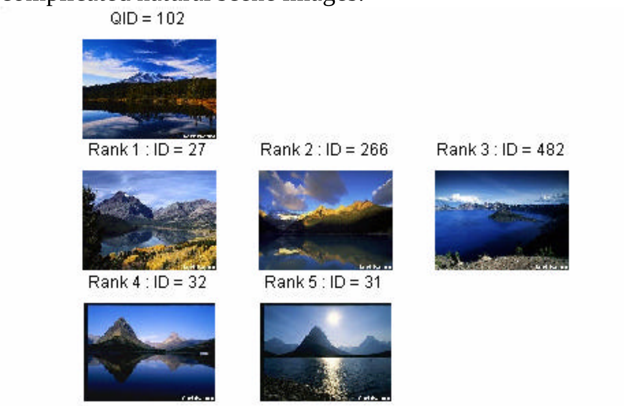

Figure 3. Query Result of Image 102.

5. REFERENCES

[1] W. Ma and H.J Zhang, "Content-Based Image Indexing and Retrieval," *Handbook of Multimedia Computing*, CRC Press. 1999

[2] 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.*

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

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

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

[6] H.D. Cheng, X. H. Jiang, Jingli Wang. "Color Image Segmentation: Advances and Prospects," *Pattern Recognition,* 2001.

[7] H.D. Cheng and Y. Sun. "A Hierarchical Approach to Color Image Segmentation Using Homogeneity," *IEEE Transactions on Image Processing*, Dec. 2001.

[8] S.-C. Chen, S. Sista, M.-L. Shyu, and R.L. Kashyap, "An Indexing and Searching Structure for Multimedia Database Systems," IS&T/SPIE Conference on Storage and Retrieval for Media Databases 2000, pp. 262-270, 2000.

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