Image Retrieval Based On Colour, Texture And Shape Feature Similarity Score Fusion Using Genetic Optimization Algorithm

Abhijit T. Somnathe
Ph. D. Scholar,
Shri JJT University, Jhunjhunu
abhijitsomnathe@gmail.com

Dr. S. Ram Reddy
Professor,
SLRTCE, Mira Road

Abstract: - Images contain information in a very dense and complex form, which a human eye, after years of training, can extract and understand. The main goal is to extract form an image a set of composing objects or real life attributes. Color, texture and shape information have been the primitive image descriptors in content based image retrieval systems. This paper presents a novel framework for combining all the three i.e. color, texture and shape information, and achieve higher retrieval efficiency. This paper proposes an image retrieval method based on multi-feature similarity score fusion using genetic algorithm. Single feature describes image content only from one point of view, which has a certain one-sided. Fusing multifeature similarity score is expected to improve the system's retrieval performance. In this paper, the retrieval results from color feature, shape feature and texture feature are analyzed, and the method of fusing multi-feature similarity score is described. For the purpose of assigning the fusion weights of multi-feature similarity scores reasonably, the genetic algorithm is applied. The experimental results show that the proposed method is superior to other methods.

Keywords*-image retrieval; fusion; genetic algorithm.*

*****

INTRODUCTION

With the rapid development of multimedia and network technology, people can access a large number of multimedia information. For people who want to make full use of multimedia information resources, the primary question is how to query the multimedia information of interest. Text query can be applied to multimedia information retrieval, but it has inherent deficiencies. One hand, text annotation of multimedia information will spend a lot of manpower and resources and it is inefficient. On the other hand, annotated text is usually a person's perception of multimedia information. It is subject to impact of individual difference and state of human and environment, and the described results may be more one-sided. In addition, it is clearly incomplete to describe content-rich multimedia information with a small amount of text. Content Based Image Retrieval (CBIR) techniques appeared in 1990s [1]. It solves the above problems well. It uses low-level features like color, texture and shape to describe image content, and breaks through the limitation of traditional text query technique.

Content-based image retrieval (CBIR) has been an active area of research in computer vision and image processing. CBIR is a system that can search, browse, and navigate similar images from large image databases based on visual content of the images. The content-based image retrieval system retrieves the stored images from the database by comparing the features of the query image against the images in the collection. The system first extracts and stores the features of the query image then it go through all images in the database and extract the features of each image. The results are the images that its features are most similar to the query image.

Color, texture and shape features have been used for describing image content. Different CBIR systems have adopted different techniques. Few of the techniques have used global color and texture features [8, 9, and 10] whereas few others have used local color and texture features [2, 3, 4, 5]. The latter approach segments the image into regions based on color and texture features. The regions are close to human perception and are used as the basic building blocks for feature computation and similarity measurement. These systems are called region based image retrieval (RBIR) systems and have proven to be more efficient in terms of retrieval performance. Few of the region based retrieval systems, e.g., [2], compare images based on individual region-to-region similarity. These systems provide users with rich options to extract regions of interest. But precise image segmentation has still been an open area of research. It is hard to find segmentation algorithms that conform to the human perception. For example, a horse may be segmented into a single region by an algorithm and the same algorithm might segment horse in another image into three regions. These segmentation issues hinder the user from specifying regions of interest especially in images without distinct objects. To ensure robustness against such inaccurate segmentations, the integrated region matching (IRM) algorithm [5] proposes an image-to-image similarity combining all the regions between the images. In this approach, every region is assigned significance worth its size in the image. A region is allowed to participate more than once in the matching process till its significance is met with. The significance of a region plays an important role in the image matching process. In either type of systems, segmentation close to human perception of objects is far from reality because the segmentation is based on color and texture. The problems of over segmentation or under segmentation will hamper the shape analysis process. The object shape has to be handled in an integral way in order to be close to human perception. Shape feature has been extensively used for retrieval systems [14, 15]. Image retrieval based on visually significant points [16, 17] is reported in literature. In [18], local color and texture features are computed on a window of regular geometrical shape surrounding the corner points. General purpose corner detectors [19] are also used for this purpose. In [20], fuzzy features are used to capture the shape information. Shape signatures are computed from blurred images and global invariant moments are computed as shape features. The retrieval performance is shown to be better than few of the RBIR systems such as those in [3, 5, 21].

The discussion above clearly indicates that, in CBIR, local features play a significant role in determining the similarity of images along with the shape information of the objects. Precise segmentation is not only difficult to achieve but is also not so critical in object shape determination. A windowed search over location and scale is shown more effective in object based image retrieval than methods based on inaccurate segmentation [22].

IMAGE FEATURE EXTRACTION

The image content is mainly embodied in color, texture and shape etc. The color feature, texture feature and shape feature describe the image content from different angle. More features will provide more information on the image content. This paper focuses on fusion method of multifeature similarity score. For convenience, this paper only discusses the fusion method of two-feature similarity score. Without loss of generality, the used features are color feature and texture feature. The following part describes the used extraction method of color feature and texture feature.

A. Color feature extraction

HSV color model forms a uniform color space, which uses a linear gauge. The perceived distance between colors is in proportion to Euclidean distance between corresponding pixels in HSV color model, and conforms to eye's feeling about color. So it is very suitable for color based image similarity comparison. In this paper, the color histogram in HSV color space is taken as the color feature describing image content. For calculating color histogram in HSV color space, HSV color space must first be quantified. According to human cognitive about color, three components of HSV space are quantified in non-uniform manner. Hue is quantized into 16 bins and is among [0, 15]. Saturation is quantized into 4 bins and is among [0, 3]. Value is quantized into 4 bins and is among [0, 3]. Among those three components, human cognitive about color is mainly based on hue, and then saturation, finally value. So, quantized results are coded as,

$$C = 16H + 4S + V \qquad (1)$$

where C is an integer between 0 and 255. Thus the color feature can be obtained by calculating histogram of an image in HSV space.

B. Texture feature extraction

In this paper, the statistical properties of image co-occurrence matrix are taken as texture features of an image. Firstly, color image is converted to grayscale image, and the image co-occurrence matrix is gained. Then, the following five statistical properties are calculated to describing image content. They are contrast, energy, entropy, correlation and local stationary. All these statistical properties are calculated in 4 directions, so we can get 20 texture features. At last, we calculated the means and variances of these five kinds of statistical properties, and took the results as the ultimate texture features, denoted as,

$$T = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5; \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5) \quad (2)$$

C. Shape feature extraction

The shape features are used to represent multiple objects in the image sand and hence is the most discriminating feature for effective image retrieval. Before extracting the shape features, images are segmented into meaningful regions using adaptive three-holding. The segmented image is then used to extract centroid distance-based Fourier descriptors [8]. The boundary of all extracted objects is regularizing by using polygon fitting algorithm. The 1-D shape signature of the boundary based on centroid distance is calculated as,

$$r(t) = \sqrt{\left([X(t) - X_c]^2 + [Y(t) - Y_c]^2\right)} \qquad (3)$$

where (X_C, Y_C) are the coordinates of object centroid. Now, Fourier transform of the shape signature can be calculated as,

$$F(k) = 1/N \sum_{t=0}^{N-1} r(t) \exp j(2\Pi kt)/N \qquad (4)$$

where k = -N/2,…N/2-1

For making the complex Fourier coefficient rotation invariant, real part is removed. The magnitude part of first 64 coefficients is taken to make the shape feature vector. All these magnitude values are made scale invariant by dividing each of them by magnitude of first coefficient for constructing the rotation and scale invariant shape feature vector as,

$$f = \left[\frac{Fm_2}{Fm_1}, \frac{Fm_3}{Fm_1}, \ldots\ldots, \frac{Fm_{N-1}}{Fm_1}\right] \qquad (5)$$

In this work, 64 Fourier coefficients are taken so as to represent finer details of shape which improve the accuracy of the system.

MULTI-FEATURE SIMILARITY SCORE FUSION

Since the physical meanings of different features are different, and value ranges are totally different, similarity scores of different features cannot be compared. So, before multi-feature similarity score are fused, they should be normalized. Similarity scores can be normalized through the following ways. Let Q be the query image. By calculating distances between the query image and images in database, similarity score set $\{S_i\}$ can be gotten, where $i = 1,\_, N$, N is the number of images in database. Thus, similarity score normalization can be implemented as,

$$S_{Ni} = \frac{S_i - \min\{S_i\}}{\max\{S_i\} - \min\{S_i\}} \quad (6)$$

The results of multi-feature similarity scores is,

$$S_{Fi} = \frac{S_{NCi} \cdot W_C + S_{NTi} \cdot W_T}{W_C + W_T}, \quad (7)$$

where S_{Fi} is the fused similarity score, S_{NCi} is the normalized color feature similarity score, S_{NTi} is the normalized texture feature similarity score, W_C is the weight of color feature similarity score, and W_T is the weight of texture feature similarity score. By assigning appropriate values to W_C and W_T, a fine similarity score fusion can be gained.

SIMILARITY SCORE FUSION USING GENETIC ALGORITHM

During the course of similarity score fusion, a key problem is how to assign the weights of similarity score. It affects directly the retrieval performance of the system. It can be considered as an optimization problem to assign reasonably the weights of color feature similarity score and texture feature similarity score. That is to find the optimum in weight value space. So, this problem can be resolved by genetic algorithm. This paper proposed a similarity score fusion method using genetic algorithm. With genetic algorithm the weights of color feature similarity score and texture feature similarity score are assigned optimally.

A. Determination of solution space

The aim of fusing similarity scores is to assign the weights of color feature similarity score and texture feature similarity score to gain a better image retrieval performance. With the consideration of (4), the weight of color feature similarity score W_C can be an integer between 0 and I, where I is a positive integer. Without loss of generality, the weight of texture feature similarity score can be assigned to $I - W_C$. The positive integer I determined the accuracy of solution. The bigger the value of I is, the higher the accuracy of solution is. But this may take a long time to resolve, and vice versa. To resolve using genetic algorithm, the weights should be encoded. The solution should be expressed as a binary number. So generally the value I is taken as $2L$, where L is a positive integer, the encoding length of the solution.

B. Population Initialization

In genetic algorithm, the number of individuals in population and the initial values of the individuals will influence the solution greatly. In this paper, the number of individuals in population N is taken as $\sqrt{I}$. N is set a bigger value, the aim of which is to gain the optimal solution quickly. The individuals are initialized as follows. The solution space is divided into N equal portions, the centers of which are taken as the initial values of the individuals.

C. Determination of fitness function

The fitness of individuals can be evaluated as follows. According to the weights W_C and W_T of N individuals, we can get N groups of image retrieval results. For every group, the top M images are considered. Total number of images is MN. By calculating occurrence frequency of images of every group in all images, the fitness of every individual is evaluated. Specific operations are as follows.

Let N_{ikj} denote if kth image A_{ik} of ith group G_i is in jth group G_j or not. That can be formulated as,

$$N_{ikj} = \begin{cases} 1, & A_{ik} \in G_j \\ 0, & A_{ik} \notin G_j \end{cases}. \quad (8)$$

Then the occurrence frequency of kth image A_{ik} of ith group G_i in all MN images is

$$N_{ik} = \sum_{j=0}^{N} N_{ikj}. \quad (9)$$

The occurrence frequency of all images of ith group G_i in all MN images is

$$N_i = \sum_{k=1}^{M} N_{ik}. \quad (10)$$

The normalized version of it is

$$P_i = \frac{N_i}{\sum_{l=0}^{N} N_l}. \quad (11)$$

The bigger P_i indicates that the images in ith group G_i possess a high proportion in all MN images, and the solution is considered a good one.

In this paper, it is taken as fitness function.

D. Solving for optimal solution

The genetic algorithm is implemented in classic mode.
The condition for ending the iteration is that the number of iteration is equal to 3. When the iteration is ended, the maximum $P* = \max Pi$ is taken as the optimal solution. According to the optimal solution the weights W_C^* and W_T^* are assigned, then the image retrieval results with these two weights are taken as the ultimate retrieval results.

EXPERIMENTS AND ANALYSIS

The proposed system is implemented in MATLAB. To establish the validity of the proposed method, it is tested on

Wang database [10] of 1,000 images. All images of this dataset are divided into 10 categories, for example, African people, beaches, buildings, buses, dinosaur, horses, flowers, elephants, mountains, and food. Each category contains 100 images each. The efficacy of the proposed system is evaluated on the basis of standard parameter, i.e., precision and recall.

Experiments are conducted by taking each image in each category as the query Image and setting the number of output images to 20. Figure 1 shows the comparison of the proposed method with the traditional histogram method and the Wang method [6]. It is obvious that the proposed method outperforms all other method for all categories in terms of average precision.

Figure 2 gives the performance comparison in terms of average recall. It can be easily concluded from the graph that the proposed method have achieved high average recall in all categories listed above.

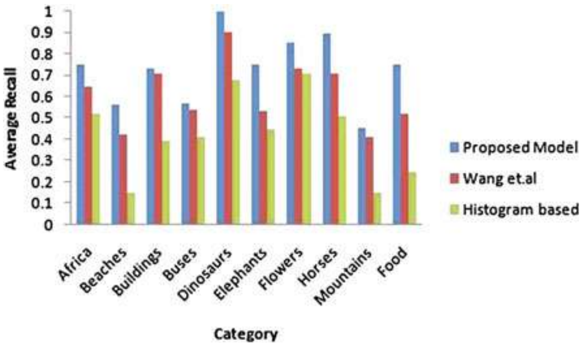

Fig. 1: Comparison based on average precision

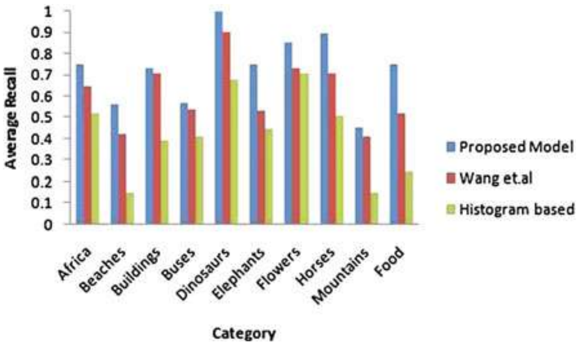

Fig.2: Comparison based on average recall

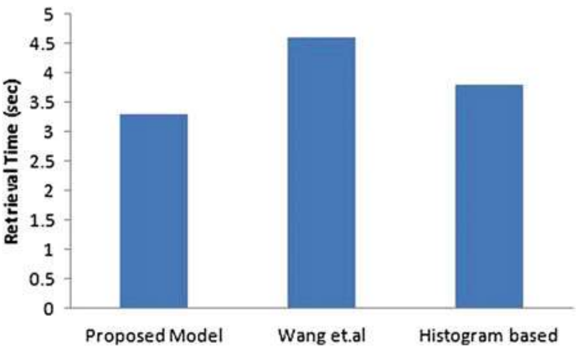

Fig.3: Comparison based on average retrieval time

Figure 3 shows the comparison on the basis of average retrieval time. It is obvious from the figure that the proposed method has comparable retrieval time which is on an average less than the other methods in comparison.

CONCLUSION

The paper presents a novel and effective combination of features for efficient CBIR system. The experimental results have shown that the proposed methods are quite effective and are better than some of the traditional methods in retrieving user intended images. In future, the validity of the proposed technique will be tested on larger image databases and work will be done to further improve the method by utilizing more effective feature set in the process of retrieval.

REFERENCES

[1] Gudivada V. N., Raghavan V. V., "Content based image retrieval systems," IEEE Computer, Vol 28, pp. 18-22, 1995.

[2] Ritendra Datta, Dhiraj Joshi, Jia Li, James Z. Wang, “Image retrieval: Ideas, influences, and trends of the new age,” ACM Computing Surveys, vol. 40, pp. 1-60, 2008.

[3]. Y. Chen and J. Z. Wang, “A Region-Based Fuzzy Feature Matching Approach to Content-Based Image Retrieval,” in *IEEE Trans. on PAMI*, vol. 24, No.9, pp. 1252-1267, 2002.

[4]. A. Natsev, R. Rastogi, and K. Shim, “WALRUS: A Similarity Retrieval Algorithm for Image Databases,” in *Proc. ACM SIGMOD Int. Conf. Management of Data*, pp. 395–406, 1999.

[5] MATLAB Image Processing Toolbox, User’s Guide

[6]. Wang, X.-Y., Yu, Y.-J., Yang, H.-Y.: An effective image retrieval scheme using color, texture and shape features. Comput. Stand. Interfaces 33, 59–68 (2011) http://wang.ist.psu.edu/

[7] Zoran Pe¡cenovi´c, Minh Do, Serge Ayer, Martin Vetterli, *New Methods for Image Retrieval,* Laboratory for Audio-Visual Communications, Swiss Federal Institute of Technology (EPFL) CH-1015 Lausanne, Switzerland.

[8]. W. Niblack *et al.*, “The QBIC Project: Querying Images by Content Using Color, Texture, and Shape,” in *Proc. SPIE*, vol. 1908, San Jose, CA, pp. 173–187, Feb. 1993.

[9]. A. Pentland, R. Picard, and S. Sclaroff, “Photobook: Content-based Manipulation of Image Databases,” in *Proc. SPIE Storage and Retrieval for Image and Video Databases II*, San Jose, CA, pp.34–47, Feb. 1994.

[10]. M. Stricker, and M. Orengo, “Similarity of Color Images,” in *Proc. SPIE Storage and Retrieval for Image and Video Databases*, pp. 381-392, Feb. 1995.

[11] http://wang.ist.psu.edu/