CONTENT BASED IMAGE RETRIEVAL - EXTRACTION BY OBJECTS OF USER INTEREST

Mr. D. Kishore Kumar

Dept of Information Technology, GITAM University,Visakhapatnam

E.Usha sree

Dept of Information Technology, GITAM University, Visakhapatnam,

K.Suneera

Dept of Information Technology, GITAM University,Visakhapatnam

P.V. Chaitanya Kumar

Dept of Information Technology, GITAM University, Visakhapatnam

Abstract

Content-based image retrieval (CBIR) systems normally return the retrieval results according to the similarity between features extracted from the query and candidate images. In certain circumstances, however users are most qualified to specify the query "content" or objects(e.g., Eiffel Tower) of their interest , not the computer and only wish to retrieve images containing relevant objects, while ignoring irrelevant image areas (such as the background). Previous work on this normally requires complicated segmentation of the object from the background. In this paper, the user can select "object of user's interest" of different shapes, non homogenous texture containing different colors regardless of many objects present in the same image using varied tools like polygonal, rectangle, circle selector tools. A two-state procedure is used to query the image from the Image database. First, we integrate global color and texture feature vectors to narrow down the search space and in next state we Process using local features. We use color moments and subband statistics of wavelet decomposition as color and texture features respectively. The shape features, generated by mathematical morphology operators, are further employed to produce the final retrieval results.

Keywords: Image retrieval; image database; Selector Tool, Object of users' interest (OUI); Color moments; Wavelet transform.

1. Introduction

In applications like involving the scan of huge volume of images to detect suspect areas, such as to some medical or military purposes, it may be valuable if the object of users' interest can be identified and exhibited when the retrieved images are presented to the user. Even in common CBIR scenarios, considering that the system usually returns a lot of images, explicitly showing the object of interest may be helpful because it could help the user recognize the images he or she really wants more quickly. Unfortunately, at present few CBIR systems can return retrieved images where the object of users' interest has been located.

In general IMAGE databases can be queried in several ways. Of these, Query-By-Example (QBE) is by far the most widely supported method in research prototypes and commercial products. A user formulates a query by selecting an example image from a pool of general image categories. Since this sample set is generally small, the expectation of finding a perfect example (in which the entire content is relevant) is low. Existing CBIR techniques are not effective in processing object of interest queries. For instance, in whole-matching approaches (matching based on the entire image), exclusion of irrelevant regions is not possible because the features of the entire image area are integrated into one or several global feature vectors (e.g., color histograms). Similarity can also be quantified by comparing homogeneous image regions. This is known as the region-based approach (e.g., Query By Image Content(QBIC) , VisualSeek , Netra , Blobworld) Each image is segmented into several

regions but not into objects and image matching proceeds by comparing visual features, such as the dominant color, texture, shape, size, etc., of these regions.
While these simple global descriptors are fast and often do succeed in partially capturing the essence of the users' query, they more often fail due to the lack of higher-level information about what exactly was of interest to the user in the query image and the effectiveness of these methods highly depends on the accuracy of image segmentation techniques, and which are not that reliable. When image regions are incorrectly detected their visual features are inaccurate, degrading the precision of the image retrieval. Moreover, what the user typically thinks of as the "object" is seldom captured by the whole image or its global features. However, in an unconstrained domain, for non-preconditioned imges, the automatic segmentation of image object of interest (OUI) is not always reliable , another problem is that an image may contain several object of which one might be particular interest to the user. Also, today's region-based systems do not support queries that are arbitrarily defined. The user still submits the entire image area as queries.

The present paper addresses this issue by proposing a similarity model for objects of users' interest. This new model enables the user to accurately include only relevant regions when formulating a query. This is illustrated in figure 1. Our system differs from above in one key aspect: there are no pre-segmented regions Instead, the user defines objects of interest (OUI) directly on a query image in order to better communicate to the search engine the intended "content" (which could possibly represent only a subset or partial aspect of the query image selected).Here the region-matching must be done in an online fashion and moreover in "interactive-time" to be tolerated by the user. To support object of interest queries, users are permitted to query arbitrarily-shaped images. In other words, they are able to identify the objects of their interest. We then process that object to extract similar images containing that object. The similarity computation is decoupled from the problem of interference by irrelevant areas, and this increases the effectiveness of retrieving objects of users' interests.

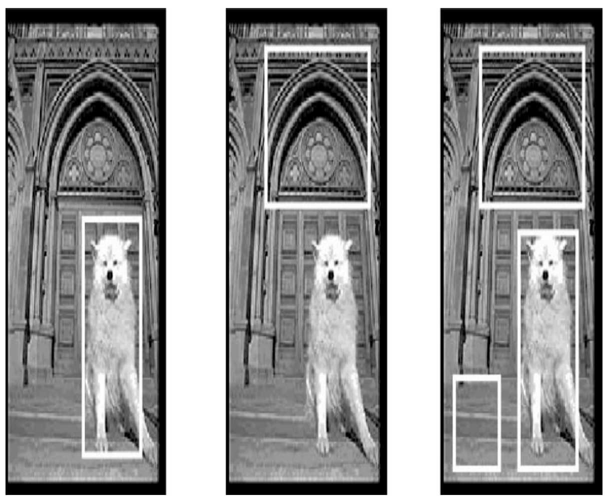

Fig 1. Multiple "contents" in a single image. (Photo courtesy of Philip Greenspun.)

The rest of paper is organized as follows. Section 2 reviews the different types of feature extractions. Section 3, 4 discusses the proposed technique and retrieval procedure. Our experimental study is presented in Section 5 and concluding remarks are set out in Section 6.

2. Feature Extraction

Feature (content) extraction is the basis of content based Image Retrieval. In broad sense, features may include both text-based features (Keywords, annotations, etc,) and visual features (color, texture, shape, faces, etc,). Within the visual feature scope, the features can be further classified as general features and domain specific features. The former include color, texture and shape features while the latter is application dependent and may include, for example, human faces and finger prints.

Because of perception subjectivity, there does not exist a single best presentation for a given feature. For any given feature there exists multiple representations which characterize the feature from different perspectives.

2.1 Color

Color feature is one of the most widely used visual features in image retrieval. It is relatively robust to background complication and independent of image size and orientation. Many Publications focus on color indexing techniques based on global color distributions. However, these global distributions have limited discriminating ability because they are unable to capture local color information. Color correlogram and color coherence vector can combine the spatial correlation of color regions as well as the global distribution of local spatial correlation of colors. These techniques perform better than traditional color histograms when used for content-based image retrieval. However, they require very expensive computation. Color moments have been successfully used in content based image retrieval systems, especially for retrieval of images only containing the objects of user's interest. Because most information can be captured by low-order moments, i.e. the first moment (mean), the second and the third central moments (variance and skewness), color moments can be effectively used as color features. Although simple, these features are inexpensive to calculate. If the value of the i-th color channel at the j-th image pixel is p_{ij}, then the color moments are defined as:

$$\mu_{i=} \sum_{j=1}^{n} p_{ij}; \qquad (1)$$

$$\sigma_i=(1/n \sum_{j=1}^{n}(p_{ij}-\mu_i)2)1/2; \qquad (2)$$

$$S_i=(1/n \sum_{j=1}^{n}(p_{ij}-\mu_i)3)1/3; \qquad (3)$$

where *n* is the number of pixels in an image. These moments may be calculated in different color spaces, e.g. the RGB, HSV and YCbCr space. For each image, a 9-dimensional $_{color}$ feature vector is obtained.

2.2 Texture

Texture refers to the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity. It is an innate property of virtually all surfaces, including clouds, trees, bricks, hair, fabric etc. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. Various techniques like co-occurrence matrix representation for texture feature extraction, tamura texture representation, wavelet transform, Gabor transform have been proposed and proved to be effective in content based image retrieval.

Due to the similarity between wavelet transform and human visual process, wavelet transform are widely used in texture analysis and classification. The subband statistics (e.g. mean and variance) of multi-level wavelet decomposition are essential features. These statistics are commonly used for texture analysis and recognition. In our approach, the statistics obtained from these wavelet transformed images are also used for texture retrieval.

2.3 Shape

In general, the shape representation can be divided into two categories, boundary-based and region-based. The former uses only the outer boundary of the shape while the latter uses the entire region. The most successful representatives for these two categories are Fourier Descriptor and Moment Invariants.

Different objects normally exhibit different shapes. At low level vision, human can identify objects according to their shapes. Therefore, shape is another important feature to describe objects in an image. Many methods have been proposed to represent object shape such as polygonal approximation, finite element models, rectilinear shapes and Fourier-based shape descriptors. However, the prerequisite of the success of these algorithms is that images have been accurately segmented into regions or objects. Since robust and accurate image segmentation is difficult to achieve, the use of shape features for image retrieval has been limited to special applications. In this paper, we focus on the method for efficient image retrieval without complicated image segmentation. Color and texture features are combined with a simple shape representation during retrieval.

3. Calculating Color, Texture and Region with Object of Users' Interest

3.1 Color feature vector

Color feature vectors are extracted by calculating color moments and defined as:

$CV=[\mu_{ci} ;\sigma_{ci};Sci];$

where i=1,2,3 is the three channels of a color space (for example RGB,HSV and YCbCr)

3.2 Texture feature

Texture feature vectors are constructed using subband statistics of wavelet multi-scale decomposition. In order to clarify the use of symbols, we briefly describe the general process of wavelet transforms and the notations of subbands. The wavelet transform of a 2D image involves recursive filtering and sub-sampling. At each level, there are three detail images. Following , we denote the detail images (subbands) as LH (containing horizontal information in the high frequency), HL(containing vertical information in the high frequency), and HH (containing diagonal information in the high frequency).The decomposition /transform also produce one approximation image, denoted as LL, which contains the low frequency information.

The wavelet transform can recursively decompose the LL band. Since 2 level wavelet decomposition yields 6 detail images, we use LH1, HL1, HH1, LH2 , HL2, HH2, and an additional approximation image LL2 to denote all the subband images. The mean and standard deviation of the coefficients in each subband uniquely characterize a texture. Mean and standard deviation are defined as

$$\mu_{ti}=1/(M*N)\sum^{M}_{j=1}\sum^{N}_{k=1}x(j,k) \qquad (4)$$

$$\sigma_{ti}=1/(M*N)\sum^{M}_{j=1}\sum^{N}_{k=1}([x(j,k)-\mu_{ti}]2)^{1/2} \qquad (5)$$

Where: i=1, 2, 3 …10 is the 10 different subband images and M,N are their sizes; x(j, k) is the subband's coefficient of wavelet decomposition; j and k represent the row and column values of the subband images respectively. Texture feature vector is represented as:

$CT=[\mu_{ti}; \sigma_{ti}];$

i=1, 2, 3…10;

3.3. Region With object Of user's interest

In this paper, object shape and size are not extracted using complicated image segmentation methods. Instead, users selects different tool for selecting the object depending upon the object .the user is provided with different tools like circle selector tool for circle shaped objects, rectangle selector tool for rectangle shaped objects, polygonal selector tool.These are illustrated in figure2.We form a template with the outline of the object so that it can be used for template matching . The dilation and erosion operations are applied further to fill any holes in the outline.

Fig 2

4. THE RETRIEVAL PROCEDURE

A two-state procedure is used to query the image from the image database. First, we integrate global color and texture feature vectors to narrow down the search space. The weighted L2 distance is used to measure the similarity. It is defined as follows:

$Dis\_1= w_1 *Dis\_CV/mean1+ w_2 *Dis\_CT/mean2,$
$w_1+ w_2 =1,$

where mean1 and mean2 are the mean values of all distances between color and texture features of the query and candidate images respectively;
Dis_CV is the L_2 distance between color feature vectors of the query image and a candidate image;
Dis_CT is the L_2 distance between texture feature vector;
W_1 and W_2 are the weights of the color feature and texture feature, respectively.

Secondly, we combine shape, color and texture features of the object of user’s interest (OUI) for image retrieval. We define templates produced by erosion and dilation operations. These templates are then used to find the images with those objects from the database that have been narrowed down in the first step, with the goal of finding the matched shapes. We may also apply templates with multiple sizes, to find matches at various scales. This scheme aims at finding the most similar objects to user’s interest.

Dilation of morphology can find images containing similar but larger objects. Erosion of morphology is used to find images containing similar but smaller objects. In this step,we also extract the color moments and texture features from the approximate shape region containing OUI to capture local color and texture information, in order to overcome the influence of irrelevant image areas (such as background).

5. Performance Analysis

As introduced in section 4, we first integrated global color and texture feature vectors to narrow down the search space. Then, we combine shape, color and texture features of the object of user’s interest (OUI) for image retrieval. In addition to the proposed algorithm, global color moments are also used for the comparison purpose. We also assessed the performance of the method in different color spaces, including RGB, HSV and YCbCr . A variety of queries using different methods are performed in our experiments. we evaluated the performance in terms of the average rate of retrieving relevant images as a function of the number of top correctly retrieved images . The average precision P(n) is defined as the number of correct images in the first n returned images. In the experiments, the precision P(n) for the first n (n = 10 ,20,and 30) retrieval results of three different methods are calculated. The new approaches are compared with the traditional methods in different ways ,as listed in Table 1

Average precision of different methods (g: global color moment ; g&t : global color and texture feature ; g&t&s : integrating shape, size, color and texture feature.)

Average Precision(%)	P(10)	P(20)	P(30)
RGB: g	18.31	15.75	11.48
RGB: g&t	41.26	32.17	21.34
RGB: g&t&s	53.74	40.12	31.54
HSV: g	19.33	17.57	13.61
HSV: g&t	43.25	34.43	24.92
HSV: g&t&s	55.25	41.43	32.36
YCbCr:g	22.50	17.52	12.67
YCbCr: g&t	44.11	42.87	22.54
YCbCr: g&t&s	54.84	42.87	32.23

Table1

From Table 1, it can be seen that the result based on integrating global color and texture features outperforms traditional algorithms based on global color moments. For texture is another important property of images. Furthermore, the integration of shape, color and texture features for non homogenous, non uniform objects can greatly improve the retrieval performance. As listed in the Table 1, the result produced by using global features is poor due to the influence of irrelevant regions (such as the background). In addition, the order of the retrieved similar images appears to be low. When we combine shape, color and texture features of the OUI, in our experiments, obviously, the retrieval result is much more effective in finding similar images and the order of the returned similar images becomes higher. Besides, many incorrectly retrieved images with high rankings have similar texture but different semantic meaning with the query image. These experimental results show that the large amount of irrelevant content in the candidate images does not affect their rankings, and the proposed method that combines shape, color and texture features of the OUI is effective in different color spaces and non homogenous regions. Table 1. shows the average retrieval results which can illustrate the effectiveness of the proposed method.

If suppose there are about 100 images in the database , Among them 20 similar images are successfully retrieved when RGB, HSV, YCbCr color spaces, and texture calculations are used, whereas only 5 similar images which have the user selected objects are retrieved using this system. The number of images retrived after second step depends mainly on the object selected by the user. This method retrieves only those images which have that object after getting those images obtained from first step.

The proposed method that combines shape, color and texture features of the OUI is effective in different color spaces and also non homogenous regions. In addition, the performances of the proposed method in different color spaces were evaluated, including RGB, HSV and YCbCr.

6. CONCLUSION AND DISCUSSION

CBIR has been widely investigated in the past years. Although many CBIR systems have been developed, few of them can return relevant images where the oui has been located. We have studied the problem caused by irrelevant regions in querying using an example environment. One way to tackle the problem is to determine similarity based solely on objects identified by the user. Such queries are called object of users' interest queries. These have been inspired by allowing user to select the object using different tools with respect to different shapes of the object. We have presented a model which combines color, texture and region with objects of users' interest. Using this method we can find effectively the objects with non uniform color and non homogenous reasons. A two-step procedure is introduced to gradually narrow down the retrieval range. It should be noted that the proposed method may be further improved by combining more complex similarity metrics.

References

[1] Flickner, M.: Query by image content: the QBIC system. In: IEEE Computer Magazine. Volume 28. (1995)
[2] Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Trans. PAMI 22 (8), 888–905.

[3] Ma, W.Y., Manjunath, B.S., 2000. EdgeFlow: A technique for boundary detection and image segmentation. IEEE Trans.Image Process. 9 (8), 1375–1388.

[4] Lau, H.F., Levine, M.D., 2002. Finding a small number of regions in an image using low-level features. Pattern Recognition 35 (11), 2323–2339.

[5] Carson, C., Belongie, S., Greenspan, H., 2002. Blobworld: Image segmentation using expectation-maximization and its application to image querying. IEEE Trans. PAMI 24 (8), 1026– 1038.

[6] Huang, J., Kumar, S., Mitra, M., Zhu, W., Zabih, R.: Image indexing using color correlograms. In: Proc. of IEEE Conf. on Com. Vis. and Pat. Rec. (1997) 762-768

[7] Long F., Zhang H.J., and Feng D.D. (2003) "Fundamentals of content-based image retrieval" in Multimedia Information Retrieval and Management- Technological Fundamentals and Applications, D. Feng, W.C. Siu, and H.J. Zhang (Eds.), Springer, 2003

[8] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: The QBIC system." IEEE Computer, Vol.28, No.9, pp. 23-32, Sept. 1995.

[9] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, Computer graphics: principles and practice, 2nd ed., Reading, Mass, Addison-Wesley, 1990.

[10] G. Pass, and R. Zabith, "Histogram refinement for contentbased image retrieval," IEEE Workshop on Applications of Computer Vision, pp. 96-102, 1996.

Authors Profile

1. Mr.D.Kishore Kumar , M.Tech , Assitant Professor . Over 5 years of Teaching experience with GITAM University and 4 years of Industry experience. Handled Courses for B.Tech, M.Tech. Research areas includes Computer Networks and Data Communications. Published 4 papers in various National and International Conferences and Journals.

2. E.Usha Sree Pursuing B.Tech(IT) from GITAM university.

3. K.Suneera Pursuing B.Tech(IT) from GITAM university.

4. P.V. Chaitanya Kumar pursuing B.Tech(IT) from GITAM university.