Image Retrieval Based on Color-Statistic Feature

Jinmei Liu
School of Science and Information
Qingdao Agriculture University
Qingdao, China
Liu_jinmei@qau.edu.cn

Jizhong Li
School of Science and Information
Qingdao Agriculture University
Qingdao, China
apanther@163.com

***Abstract*—Color is the most widely used visual feature in content based image retrieval. The visual coherence color space, HSV, is adopted to represent image. Hue component is used to denote color. Hue difference statistic is proposed to extract color change information as supplement to color feature. The image is divided into sub images equally. Color and change information is extracted in each region. After feature vector clustering and coding, image content can be expressed as vector codes. The text based analysis technology is used for image retrieval. Experiments show that the proposed method can realize efficient retrieval for unconstrained scene images.**

Keywords-color; change statistic; image retrieval

I. INTRODUCTION

With the development of computer and multimedia technology, mass digital image data are widely used in medical, remote sensing, entertainment, education and national defense fields. In order to use these images effectively, the image management system emerged to meet the times' need. The main purpose of image retrieval system is retrieving similar images in database according to inquiry conditions. Content based image retrieval automatically extracts image visual features and establishes a suitable index structure. The query image and the images in the database are matched in the feature space on the basis of appropriate criterions. Some international famous image retrieval systems are QBIC, Virage, MARS, WebSEEK and VisualSEEK system, etc..

Image content includes color, texture and edge. The common used method to present color is color histogram. The frequency information of each color bin is given in color histogram. Digital image source can be regarded as an ergodic random function and an image is a sample of the source. Histogram is the first order joint probability density function [1]. Color change information provides color spatial distribution information. Wavelet transform [2] and gray level co-occurrence matrix [3] can be used to express gray value change. Edge, a basic feature of an image, is defined by its adjacent pixel with step change or roof change. Hiremath used gradient vector flow to represent image edge [4]. Multiple features are used together to improve the efficiency of image retrieval. Color and edge features were combined by Savvas to describe image content [5]. Color histogram and texture feature were integrated to reveal image feature [6]. In this paper, color and spatial distribution are extracted to represent image and text based analysis technology is used in image matching.

II. IMAGE PARTITION AND FEATURE EXTRACTION

A. Image Partition

An image is difficult to be expressed accurately as a whole. It can be divided into sub images. Each sub image represents certain pattern. An image can be expressed as a number of patterns so similar images possess similar patterns.

Segmentation is the process of dividing an image into several parts. The ideal result of segmentation is decomposing an image into different regions. The same region has similar properties and any adjacent regions have different properties. Existing segmentation methods are proposed for a special category images. There is no universal segmentation method for unconstrained scene images. To achieve the unconstrained image retrieval, grid partition is adopted to divide images. Grid partition is different from segmentation. It segments an image into equal sub images by grid without considering the image content. In order to maximize the retention of information in an image, partition scale is chosen as 5×5 [7].

B. Feature Extraction

Color images are stored in RGB color space in computer, but the RGB color space is not the best choice for color representation. Since it does not have vision consistent property, errors occur in subsequent color quantization. The so-called visual consistency is that the distribution of color is smooth and the changes in color and vision are synchronized. HSV space is visual consistent and is selected as image representation space. H (Hue) is the basic attribute of color. It is usually referred to color name, such as red, yellow, blue, etc.. S (Saturation) is the purity of color. V (Value) means brightness. Each component in HSV color space is quantized into 10 bins. As image illumination environment is not restricted, H is the only component for color expression to reduce the dimension of features. Color histogram based on H component is extracted in each sub image and a 10-dimension vector is obtained to represent region color.

Hue difference statistic is proposed to describe color distribution information. Assume h(x,y) is a hue value at row x and column y in an image. The hue difference of the pixel is defined by (1).

$$h(x,y) = \left| h(x,y) - h(x,y+1) \right| \qquad (1)$$

Published by Atlantis Press, Paris, France.

Hue values are in the range of 10 levels. Hue differences are calculated in sub images. The probability of difference value i is represented by p(i). Four statistics can be computed by (2-5).

$$CON = \sum_i i^2 p(i) \quad (2)$$

$$ASM = \sum_i [p(i)]^2 \quad (3)$$

$$ENT = -\sum_i p(i) \lg p(i) \quad (4)$$

$$MEAN = \frac{1}{m} \sum_i ip(i) \quad (5)$$

Four statistics are extracted in each sub image and 4 parameter values are used to describe color changes in a region. Considering the prior 10-dimension color feature, a region is represented by a 14 dimensional vector.

III. FEATURE CLUSTERING AND CODING

In order to obtain the entire patterns of sub-images, some images in image database are selected as training images. The training images are segmented by a fixed scale grid and feature extraction is performed as follows. Clustering is implemented for all feature vectors. Finally all the sub image patterns are acquired.

There are a large number of Clustering algorithms, such as the nearest neighbor clustering algorithm, K-means algorithm, ISODATA algorithm, hierarchical clustering algorithm, genetic algorithm and artificial neural network clustering algorithm. The K-means clustering algorithm and Kohonen neural network are generally used in large data clustering, because control parameters are relatively easy to set in those algorithms and their execution speed is satisfactory [8]. As K-mean clustering algorithm is easy to fall into local optimal clustering results and clustering results depend on the initial clustering centers, Kohonen network is used in the experiments.

Kohonen network is a kind of unsupervised learning neural network. The network can learn input vectors' rules and relations by adjusting the weights of network. It identifies the input vector distribution and topological structure. The input layer of the neural network includes 14 neurons, corresponding to the 10-dimension color and 4-dimension color change statistical information. Euclidean distance is used as distance function and random-layer function is adopted initialize the network topology. Network learning rate is 0.9 and the learning step is 5000. More than 37 thousands feature vectors were clustered into 360 cluster centers. Those cluster centers can be used to represent all the regional characteristics.

These cluster centers can represent the training vectors, which are called as entity. An image in database corresponds to a set of certain entities in ordered distribution. In order to simplify the representation of images, the entities are coded. An image is encoded as a two-dimensional code table. Image coding simplifies the subsequent processing and saves storage space. The mature text retrieval technique can be applied to the image retrieval.

Vector model is used to compute image similarity. Assume d_j is image j in the database and q is a query image. There are t code words in the code book. The similarity measure is defined by (6). The distance function is show in (7).

$$S(q, d_j) = \frac{1}{1 + dis(q, d_j)} \quad (6)$$

$$dis(q, d_j) = \sum_{i=1}^{t} \frac{|f_{ij} - f_{iq}|}{1 + f_{ij} + f_{iq}} \quad (7)$$

Here, f_{ij} is the probability of code i in image d_j and f_{iq} is the probability of code i in image q.

IV. EXPERIMENTAL RESULTS

The images used in experiments were downloaded from the image retrieval database (http://www.cs.washington.edu/research/imagedatabase/groundtruth). There are 11 image sets in the database. 256 images were selected as the experimental database and 36 images were selected as training samples in experimental database. The size of image is 200×300 or 300×200.

Partition was performed for all of the images in the training set. Color and change information was integrated as a feature vector from each image blocks. Since a large number of feature vectors are difficult to operate, feature vectors were divided into several subsets. Clustering was carried out in each subset. Then all the cluster centers were combined together and clustered again. Eventually 360 cluster centers were obtained, which can be used to represent all the features. The cluster centers were coded, so images were mapped to two-dimensional code tables. By comparing the code tables, the similarity was calculated.

Published by Atlantis Press, Paris, France.

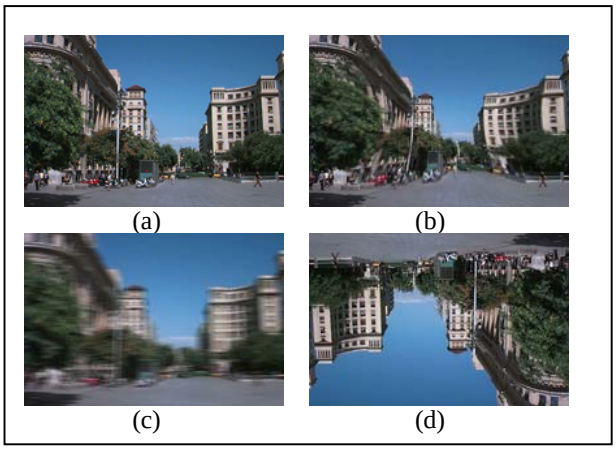

Figure 1. Query images.

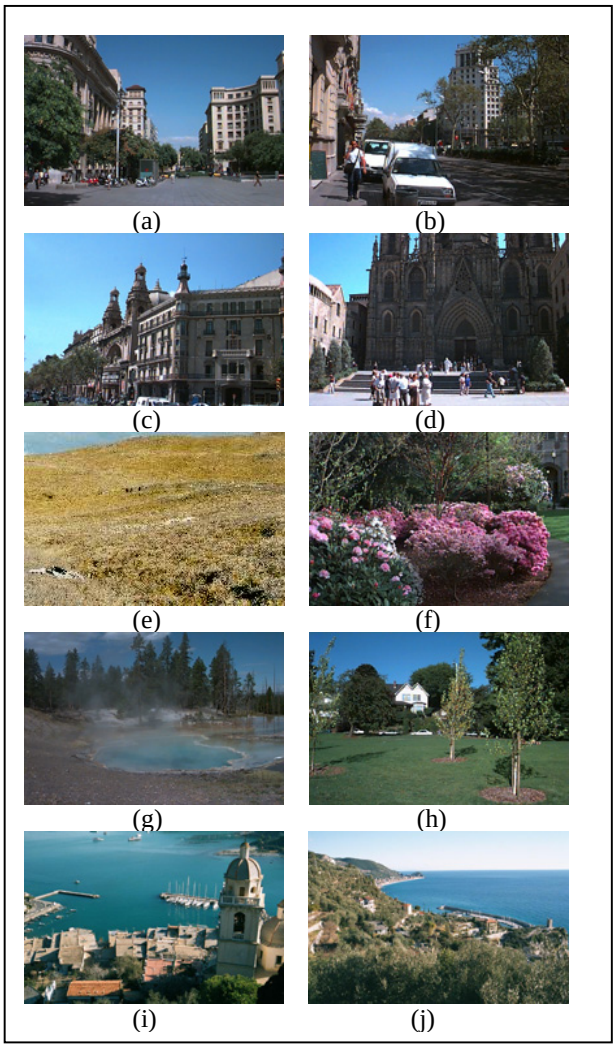

Figure 2. Test images.

An image in database was selected as model image for query. Images may be polluted for some reasons in the processing of collection and transmission so the robustness of the algorithm is very important. In order to verify the robustness of the proposed algorithm, the model image was processed by distortion, motion blur and rotation operation, as shown in Fig. 1. Fig. 1(a) is the original image. Fig. 1(b) is Fig. 1(a) squeezed by 50%. Fig. 1(c) is Fig. 1(a) blurred in horizon at a step of 7 pixels. Fig. 1(d) is Fig. 1(a) rotated in 180 degree.

Ten images were selected as test images, which were showed in Fig. 2. Fig. 2(a) and Fig. 1(a) are the same image. There are buildings, trees, streets, pedestrians and blue sky in the image. Seen from the image content, Fig. 2(b) and Fig. 2(d) are more similar to Fig. 2(a). There the same components in them although each component has different proportions. Fig. 2(c) and Fig. 2(i) are related to Fig. 2(a). There are buildings and green trees in the two images. Moreover, there are pedestrians and blue sky in Fig. 2(c). The original image and the processed image were inputted into the system as query images one by one. The similarity values of images were computed in Tab. 1.

The highest similarity is written in bold in Tab. 1. It is not difficult to find that the original image has high similarity with the processed or transformed image. Although the first four images are related in the content, Fig. 2(c) is more similar with Fig. 2(a) in content and the similarity is higher. The other two similar images are Fig. 2(g) and Fig. 2(j). From the subjective judgment, the two images are not similar to Fig. 2(a), but their hue is very similar to Fig.2 (a), so the similarity is high. The matching results of blurred image are worst in experiment. The reason for the results may be that the main feature in the proposed method is color and blurring has serious damage to the color information.

TABLE I. MATCHING RESULTS FOR THE QUERY IMAGES

Test Image	**Similarity**			
	Original image	***Distortion***	***Blurring***	***rotation***
Fig. 2(a)	**1.00000**	**0.69568**	0.53213	**0.66814**
Fig. 2(b)	0.52884	0.52108	0.49852	0.53378
Fig. 2(c)	0.56564	0.54118	0.47986	0.56644
Fig. 2(d)	0.52336	0.50917	0.43523	0.52126
Fig. 2(e)	0.39002	0.39010	0.40325	0.38861
Fig. 2(f)	0.48256	0.47775	0.42594	0.47992
Fig. 2(g)	0.51971	0.53733	0.53488	0.52448
Fig. 2(h)	0.48598	0.50252	0.48243	0.48193
Fig. 2(i)	0.49665	0.50777	0.50413	0.49674
Fig. 2(j)	0.52686	0.53417	**0.54109**	0.51650

V. CONCLUSION

A new method for image retrieval is presented in this paper. Ten color components in visual consistent space are used to represent color and four statistical components of hue changes are adopted to represent spatial information. The feature vectors are clustered and coded. Vector model is used to calculate image similarity. The proposed retrieval method shows robustness in case of distortion, blurring and rotation in content based image retrieval.

Published by Atlantis Press, Paris, France.

ACKNOWLEDGMENT

This work was financially supported by Shandong Provincial Natural Science Foundation, China under Grant BS ZR2009GL008 and ZR2012FM009. The authors would like to thank the anonymous reviewers for their comments.

Corresponding author: Jizhong li.

REFERENCES

[1] Z. John, S. S. Iyengar, B. Jacob, "Content based image retrieval and information theory: a general approach," Journal of the American Society for Information Science and Technology, vol. 52, pp. 840–852 , 2011.

[2] Q. Yu-Long, Z. Chun-Hui, S. Chun-Yan, "Complex wavelet based texture classification," Neurocomputing, vol. 72, pp. 3957-3963, 2009.

[3] M. Partio, B. Cramariuc, M. Gabbouj, A. Visa, "rock texture retrieval using gray level co-occurrence matrix," unpublished.

[4] P. S. Hiremath, J. Pujari, "Content based image retrieval based on color, texture and shape features using image and its complement," International Journal of Computer Science and Security, vol. 1, pp. 25-36, 2007.

[5] A. Savvas., S. Yiannis Boutalis, "CEDD: color and edge directivity descriptor, a compact descriptor for image indexing and retrieval," ICVS Proceedings, Santorini, Greece, vol. 5008, pp. 312–322, May, 2008,.

[6] H.B.Kekre, K.Tanuja Sarode,D. Sudeep, "Image retrieval using color-texture features from DCT on VQ codevectors obtained by Kekre's fast codebook generation," ICGST-GVIP Journal, Vol.9, pp. 1-8, 2009.

[7] L. Jinmei., W. Guoyu. "A novel partition-based color image retrieval," Journal of Information & Computational Science, vol. 3, pp. 811-816, 2006.

[8] A.K. Jain, M.N. Murty, P.J. Flynn, Data clustering: a review, ACM , 1999.