Combining Color and Shape Features for Image Retrieval

XiaoFu Lee and Qian Yin

Image Processing and Pattern Recognition Laboratory, Beijing Normal University, Beijing 100875, China
yinqian@bnu.edu.cn, redfox_li@mail.bnu.edu.cn

Abstract. This paper presents a method for extracting and combining features from image that can be used to perform effective image retrieval. First, we use a method of diamond segmentation for all images in database. Then, in HSV color space, we extract color vectors of 256 dimensions for both the segments in and out the diamond. At last, we extract shape feature from all images with the method of moment invariants. Then we compare our method with single color feature and shape feature, as well as two contrast combining methods. Experiments show that our method has better effectiveness and feasibility.

Keywords: Content-Based Image Retrieval (CBIR), HSV color features, moment invariants.

1 Introduction

CBIR has become a vivid research field since 1970's to meet the increasing demand of image retrieval. Generally speaking, CBIR is a system which includes a group of technologies. Extracting features from images is always the first and ultra important step in CBIR system, which has a direct effect on the system.

For its robustness to noise, image degradation, changes in size, resolution, or orientation, color feature is the most widely used feature in CBIR system [2]. Various color spaces are available, definitions of different color spaces can be found in Ref. [1]. Color spaces shown to be closer to human perception and used widely in CBIR system include, RGB, LAB, LUV, HSV, etc.. HSV color space is uniform on visual property, and consonant with human perception [3], and its color difference is proportional to the corresponding color component values of Euclidean distance, hence, we choose it in our implementation.

Color histogram [4], statistically denotes the joint probability of intensities of color channels, and describes the global color distribution in an image. Accumulated color histogram [11] includes single color and multi colors, avoids the disadvantage of zero when subtracted. Color coherence Vector [12] considers the continuousness of color distribution. Dominant color [13] represents image with 20 dominant colors, grasps the main content and despites the less important content. But on the same time, methods based on color histogram drop spatial information of image. A general solution to this issue is to divide the image into several blocks, and to calculate each block color

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2009, LNCS 5616, pp. 569–576, 2009.

histogram separately. Our dividing method is a diamond segmentation which has a consideration of the behavior of human perception.

Different objects of the same category usually have the same shape, and shape has a direct relationship with object. Hence, shape feature could distinguish the difference between objects very well. Shape features have shown to be useful in many domain specific images such as man-made objects. Jain [14] calculates the frequencies of all segments' slopes to represent object's contour. Gudivada [15] uses spline to fit object's edge, and then calculates each control point's information of slope and curvature. Flicker presents shape vector which includes area, eccentricity, roundness and etc in QBIC system. For color images used in most papers, however, it is difficult to apply shape features compared to color and texture due to the inaccuracy of segmentation. Despite the difficulty, shape features are used in some systems and have shown potential benefit for CBIR [6]. We select 7 moment invariants of shape, prompted by Hu [7], as our shape descriptor, for its shift, rotation and scale invariance.

Our motivation for this paper is to find a way by which we can combine different features and overcome the shortcomings of single feature, and improve the performance of CBIR. The rest of this paper is organized as follows: We first discussed the detail implementation of our combining strategy. Then, we provide experimental results to show the property of this combining strategy. Finally, conclusions are drawn.

2 Implementation

This section describes the details in our implementation. The first issue we encounter is that there does not exit a standard image database for image retrieval till now, Rui and Huang [8] believe that even though there are ongoing research efforts to build true image databases, the systems are not at the complete stage yet. Hence, we build an image database including 10 categories, such as flower, horse, bird and etc., from web, each category has 10 images. The images in our database are all in jpg format, within the range of 300×400 to 1024×768, are illustrated by figure 1.

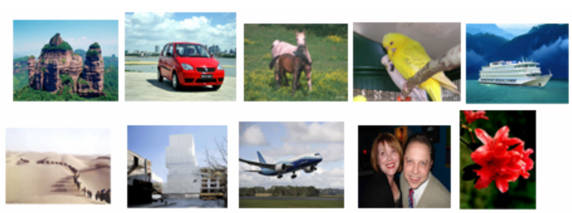

Fig. 1. 10 categories in our image database

2.1 Image Segmentation

People used to put the most interesting object in the center of an image when the image is formed, in the same way; people start to recognize objects from the center of an image. On the contrary, people used to ignore the information at the corners of an image. In order to meet these habits of human perceptions, we prompt a method of segmentation called Diamond Segmentation. It connect the midpoints of 4 edges of a rectangle to form a diamond, the region in the diamond, we believe, covers the most important information in an image, is given the weight of 0.75; while the region out of the diamond contains the information of background, is given the weight of 0.25. The diamond segmentation is illustrated by figure 2.

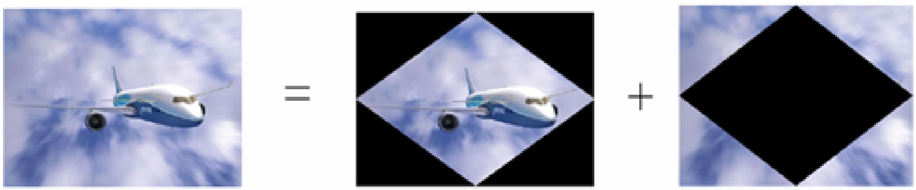

Fig. 2. Diamond segmentation

2.2 Color Feature

HSV color space [16] is consonant with human perception: H for hue, represents colors; S for saturation, represents color purity; V for value, represents the intenseness of color. These three components above are independent from each other. There are many methods converting image from RGB to HSV, here is a set of formulae comparatively easier to carry out.

$$H = \begin{cases} \arccos \dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} & B \le G \\ 2\pi - \arccos \dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} & B > G \end{cases}$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)}$$

$$V = \frac{\max(R,G,B)}{255}$$

We divide the HSV color space into 16×4×4 bins, i.e. 16 hues, 4 saturations, 4 values. Hence, we will get two 256 dimensions vectors of color (both the region in the diamond and out of the diamond) for each image in our database. Then, we normalize both the vectors by the formula 1.

$$F = \frac{F}{Width \times Height} \times 100 \tag{1}$$

In which, F for the color feature we extracted from image originally, Width for the width of the image, Height for the height of the image.

2.3 Shape Feature

Generally, there are two groups of method for extracting shape feature: one is based on edge; the other is based on region. Moment invariant is one of the region based methods. For 2-D continuous function, the p+q moment is defined as formula 2.

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy \qquad p,q = 0,1,2\ldots \tag{2}$$

If f(x,y) represents an image, the formula is reshaped as formula 3.

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y) \tag{3}$$

It is proved that the relationship between m_{pq} and f(x,y) is one to one. The p+q moment of f(x,y) is defined as formula 4.

$$u_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y}) f(x,y) \tag{4}$$

In which, ($\bar{x}$, $\bar{y}$) is the centre of gravity of f(x,y). The normalized central moment of f(x,y) is presented as formula 5.

$$\eta_{pq} = \frac{\mu_{pq}}{\eta^{\gamma}{}_{00}} \qquad \gamma = \frac{p+q}{2} + 1 \qquad p+q = 2,3,\cdots \tag{5}$$

Combining the 2nd moments and 3rd moments, Hu [7] prompted 7 moments with the property of transform, rotation and scale invariant.

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta^2{}_{11} \\
\phi_3 &= (\eta_{30} - \eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}$$

We compute Hu’s invariants for each image in our database as shape features.

2.4 Similarity Measurement

Let Q_a, Q_b be the color feature vectors of the query image, Q_a for the region in the diamond, Q_b for the region out of the diamond. Let B_a, B_b, be the color feature vectors of the image in the database, B_a for the region in the diamond, B_b for the region out of the diamond. The color distance D_c between the query image and the image in the database is given according to the formula 6.

$$D_c = 0.8\sqrt{(Q_a - B_a)^2} + 0.2\sqrt{(Q_b - B_b)^2} \tag{6}$$

Let Q_s be the shape feature vector of the query image, B_s be the shape feature vector of the image in the database, the shape distance D_s between the query image and the image in the database is given according to the formula 7.

$$D_s = \sqrt{(Q_s - B_s)^2} \tag{7}$$

The final distance D between the two images is calculated according to the formula 8.

$$D = 0.6D_c + 0.4D_s \tag{8}$$

3 Experiments and Discussion

We compared our method with color histogram and Hu's invariants using recall and precision [9], which are defined as formula 9 and 10.

$$recall = P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{a}{a+b} \tag{9}$$

$$precision = P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{a}{a+c} \tag{10}$$

where, A stands for all the images which have the same label with the query image in our database, B stands for all the retrieved images, a stands for all the images having the same label with the query image in a query, b stands for all the images not similar to the query image in a query, c stands for all the images similar to the query image but not retrieved.

Figure 3 shows the comparison of precision-recall curves of the method of this paper, color histogram and Hu's invariants. We can see that our method has a higher precision than both the others at the same recall, which means that this method is more efficient, and with less irrelevant images retrieved back.

Figure 4, 5, 6 show the results of the three different methods mentioned in this paper, where the first image is the query image. We can see that the relevant images are closer to the query image with our method and more relevant images are retrieved. In

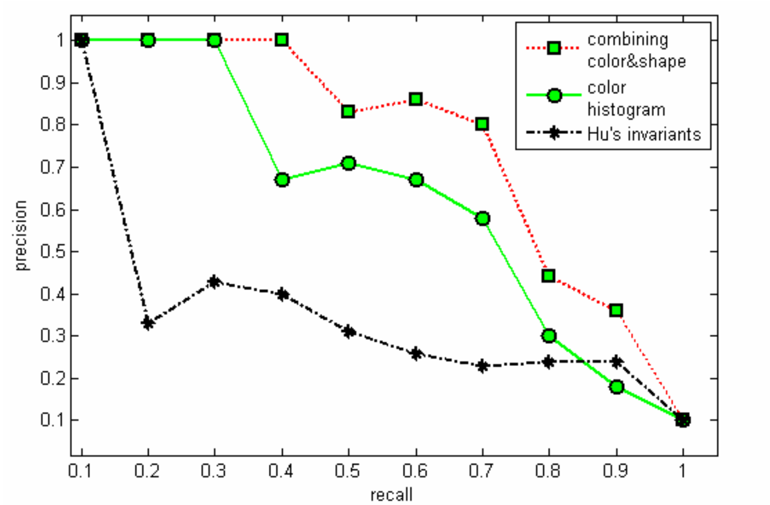

Fig. 3. Comparison of precision-recall curves

Fig. 4. The result of our method

figure 4, there are 3 irrelevant images retrieved back for the reason that giving different weights can only affects retrieved image's order, but can't remove the disadvantage of feature extracting methods.

We also compare our method with two contrastive methods. Contrastive method 1 combines accumulated color histogram and edge orientation autocorrelogram [10];contrastive method 2 uses wavelet color and shape feature.

Table 1 show the experimental result of average precision when the number of returned images is 10. Comparing with contrast method 1 and 2, our method has a higher precision when the query image is in the categories of horse, flower and human. These categories have a common characteristic that color feature dominates the images, while shape feature is supplementary. But when the query image is in the categories of ship, desert and plane, the precision of our method is low, also the other two. These categories have a common characteristic that there is too much noise when shape feature is extracting, and color features vary in color space. Totally, our method is more effective and feasible than the other two.

Fig. 5. The result of color histogram

Fig. 6. The result of Hu's invariants

Table 1. Comparing our method with contrastive methods

	Our Method	**Contrast method 1**	**Contrast method 2**
Category	$\overline{P}(N=10)\%$	$\overline{P}(N=10)\%$	$\overline{P}(N=10)\%$
Horse	92	84	62
Bird	71	65	43
Human	79	57	67
Plane	55	43	49
Building	78	82	62
Desert	54	63	47
Car	74	73	63
Mountain	81	82	60
Ship	79	85	43
Flower	84	64	66
Average	74.7	69.8	56.2

4 Conclusions

This paper studies the combination of color and shape features. Images are divided into central section and corner section by a diamond before color features are extracted, and then we extract Hu's invariants for the entire images in our database. By giving different weights, we combine color feature and shape feature. Experimental results show that the effect of our method is effective and feasible.

Acknowledgements

The research work described in this paper was fully supported by grants from the National Natural Science Foundation of China (Project No. 60675011 and 90820010).

References

1. Plataniotis, K.N., Venetsanopoulos, A.N.: Color Image Processing and Applications. Springer, Berlin (2000)
2. Gevers, T.: Color Based Image Retrieval. Springer, Heidelberg (2001)
3. Liu, Y., Zhang, D.S., Lu, G., Ma, W.Y.: Region-based Image Retrieval with Perceptual Colors. In: Proceedings of the Pacific-Rim Multimedia Conference (2004)
4. Swain, M.J., Ballard, D.H.: Color Indexing. International Journal of Computer Vision (1991)
5. Mehrotra, R., Gary, J.E.: Similar-shape Retrieval in Shape Data Management. IEEE Computer 28(9), 57–62 (1995)
6. Liu, Y., Zhang, D., Lu, G., Ma, W.Y.: A Survey of Content-based Image Retrieval with High-level Semantics. Pattern Recognition 40, 262–282 (2007)
7. Hu, M.K.: Visual Pattern Recognition by Moment Invariants. In: Computer Methods in Image Analysis. IEEE computer Society, Los Angeles (1977)
8. Rui, Y., Huang, T.S.: Image Retrieval: Current Techniques, Promising Directions, and Open Issues. Journal of Visual Communication and Image Representation 10, 39–62 (1999)
9. Smith, J.R., Li, C.S.: Image Retrieval Evaluation. In: Workshop Content-based Access of Image and Video Libraries (1998)
10. Mahmoudi, F.: Image Retrieval Based on Shape Similarity by Edge Orientation Autocorrelogram. Pattern Recognition 36, 1725–1736 (2003)
11. Stricker, M.A., Orengo, M.: Similarity of Color Images. In: Proc. of SPIE: Storage and Retrieval for Image and Video Databases III, San Jose, CA, pp. 381–392 (1995)
12. Pass, G., Zabih, R., Miller, J.: Comparing Image Using Color Coherence Vectors. In: Proc. of ACM Intern Conf. Multimedia, Boston, MA (1996)
13. Zhang, H.J., Low, C.Y., Smoliar, S.W., Wu, J.H.: Video Parsing, Retrieval and Browsing: An Integrated and Content-based Solution. In: ACM Multimedia 1995, pp. 15–24 (1995)
14. Jain, A.K., Vailaya: Shape-based Retrieval: A Case Study with Trademark Image Database. Pattern Recognition 31, 1369–1390 (1998)
15. Gudivada, V.N., Raghavan, V.V.: Content-based Image Retrieval System. IEEE Computer 28(9), 18–22 (1995)
16. Ulrich, H.G.: Pairwise classification and support vector machines. In: Bernhard, S. (ed.) Advances in Kernel Methods: Support Vector Learning, pp. 255–268. The MIT press, Massachusetts (1999)