Optimal Matching of Images Using Combined Color Feature and Spatial Feature

Xin Huang, Shijia Zhang, Guoping Wang, and Heng Wang

Human-Computer Interaction & Multimedia Lab,
Department of Computer Science and Technology,
Peking University, 100871 Beijing, P.R. China
{hx, zsj, wgp, hengwang}@graphics.pku.edu.cn

Abstract. In this paper[1] we develop a new image retrieval method based on combined color feature and spatial feature. We introduce an ant colony clustering algorithm, which helps us develop a perceptually dominant color descriptor. The similarity between any two images is measured by combining the dominant color feature with its spatial feature. The optimal matching theory is employed to search the optimal matching pair of dominant color sets of any two images, and the similarity between the query image and the target image is computed by summing up all the distances of every matched pair of dominant colors. The algorithm introduced in this paper is well suited for creating small spatial color descriptors and is efficient. It is also suitable for image representation, matching and retrieval.

1 Introduction

Color is a widely used low-level feature in content-based image retrieval systems [1, 4, 5, 6], because of its characteristic of invariance with respect to image scaling and orientation. Smith [1] proposed a method to quantize colors into 166 bins in the *HSV* color space. Zhang [2] gave a new dividing method to quantize the color space into 36 non-uniform bins. It has been observed that the color quantization schemes have a major and common drawback. That is similar colors might be quantized to different bins in the histogram, thus increasing the possibility of retrieving dissimilar images.

Besides color histogram, another commonly used method is to apply clustering based techniques in quantizing the color space. Ma et al. utilized a vector quantization called Generalized Lloyd algorithm (GLA) [3] to quantize the *RGB* color space. Mojsilovic [4] proposed a new quantization scheme in the *Lab* space based on spiral lattice. However, the problem of how to extract semantic information from the image still remains the biggest obstacle in the content-based image retrieval system [13, 14]. Rogowitz performed psychophysical experiments [6] analyzing human perception of image content, showing that visual features have a significant correlation with

[1] The research was supported by No. 2004CB719403 from The National Basic Research Program of China (973 Program), No. 60473100，No. 60573151 from National Natural Science Foundation of China.

V.N. Alexandrov et al. (Eds.): ICCS 2006, Part I, LNCS 3991, pp. 411 – 418, 2006.

semantically relevant information. Mojsilovic indicated that even with the absence of semantic cues, "semantically correct retrievals" [5] can also be achieved by perceptually based features. By exploiting the fact that the human eye cannot perceive a large number of colors at the same time, nor is it to distinguish close colors well, we aim to create a small color descriptor, which is suitable for image representation, matching and retrieval.

In this paper we introduce a color feature extraction method based on ant colony clustering algorithm, which models the behavior of ants' collecting corpses and is self-organizing. The algorithm extracts perceptually dominant colors as the basis for image matching. The spatial information of dominant colors is then taken into account in order to enlarge feature space and increase the retrieval precision. Similarity metric between any two images is established by using an optimal matching algorithm in graph theory.

2 Dominant Color Feature Extraction

Ant colony clustering algorithm [7, 8, 9] has been proposed and applied in various areas since 1990s, while it models the ants' behavior of piling corpses. Researchers found that the ants can assemble the ant corpses into several piles in their studies. Deneubourg proposed a model that explains the ants' behavior of piling corpses, which is commonly called BM (Basic Model) [7] to describe the ants' clustering activity. The general idea is that when an unloaded ant encounters a corpse, it will pick it up with a probability that increases with the degree of isolation of the corpse; when an ant is carrying a corpse, it will drop the corpse with a probability that increases with the number of corpses in the vicinity. The picking and dropping operations are biased by the similarity and density of data items within the ants' local neighborhood.

The step of dominant colors extraction based on ant colony clustering is as follows. First an input image is transformed into *CIELAB* color space. We get the training sequence consisting of M source vectors: $T = \{x_1, x_2, \ldots, x_M\}$. The source vector that is three-dimensional consists of L, a, b value in *CIELAB* color space. Then we utilize the ant colony clustering algorithm [7, 8] to extract the dominant colors from the training sequence T. The first step is to randomly project training sequence T onto a plane, and a few virtual ants are generated, randomly placed on the plane. Then the density measure of each ant is computed [8]. Each ant acts according to its current state and corresponding probability. Finally several clustering centers are visually formed through the ants' collective actions. The algorithm is ended with a few clustering dominant colors generated. After using the ant colony clustering algorithm, we extract the dominant color set denoted as $C = \{c_1, c_2, \cdots c_K\}$, $P = \{p_1, p_2, \cdots p_K\}$, where each dominant color $c_i = \{L_i, a_i, b_i\}$ is a three-dimensional Lab color value, and p_i is the corresponding size percentage. In our experiments the number of dominant colors K is assigned the value 16.

3 Combined Color Feature and Spatial Feature

In the procedure of color clustering, we only consider the color feature of each image. Thus it may lose color distribution information and lead to false retrieval. In order to prevent this problem, we introduce the color spatial information to enlarge the feature space of dominant colors. Moment [10, 11] is a simple and effective way for representing the spatial feature in images. It has the prominent property of being invariant to image rotation, shift and scale. We use the centroid of the dominant colors and the second-order central moment [11] to represent spatial features. The centroid represents the location of each dominant color, and the second-order central moment indicates the mass distributing information of dominant colors.

In Statistics, moment represents fundament distributing properties of random variables. The $p+q$ th-order moments of a bounded function $f(x, y)$ with two variables is defined as:

$$M_{pq} = \int\int x^p y^q f(x, y)dxdy \tag{1}$$

where p and q are nonnegative integers [11].

Suppose the size of an image is $m \times n$. After extracting dominant color features by ant colony clustering, the dominant color set is denoted as $C = \{c_1, c_2, \cdots c_K\}$, the $p+q$ th-order moments of dominant color c_i can be defined as follows:

$$M_{pq}^{i} = \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} x^p y^q f(x, y, i) \tag{2}$$

If the pixel at coordinates (x, y) belongs to the dominant color c_i, then $f(x, y, i) = 1$; otherwise $f(x, y, i) = 0$.

Then the centroid coordinates $(\overline{x}_i, \overline{y}_i)$ of each dominant color can be computed using the first-order moment, $\overline{x}_i = \frac{M_{10}^i}{M_{00}^i}, \overline{y}_i = \frac{M_{01}^i}{M_{00}^i}$. The $j+k$ th-order central moments of dominant color c_i can be defined as:

$$\mu_{jk}^{i} = \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} (x-\overline{x}_i)^j (y-\overline{y}_i)^k f(x, y, i) \tag{3}$$

We use the second-order central moment μ_{11} to describe the mass distributing feature of dominant colors.

4 Similarity Measure

In order to define the similarity metric between two images, we first give the formula of computing the distance between two dominant colors c_i and c_j. Both the color feature and spatial feature are considered in defining the distance. According to the dominant color feature and spatial feature defined in section 2 and section 3, we compute four corresponding distances in section 4.1.

4.1 Distance Computation

Distance $dCc(c_i, c_j)$ is the color difference of c_i and c_j in *CIELAB* color space. Distance $dPt(c_i, c_j)$ is the area percentage difference of c_i and c_j . Distance $dCt(c_i, c_j)$ is the centroid coordinates difference of c_i and c_j . Distance $d\mu(c_i, c_j)$ is the second-order central moment difference of c_i and c_j . The formulas of four distances are defined as follows:

$$dCc(c_i, c_j) = \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \tag{4}$$

$$dPt(c_i, c_j) = | P_i - P_j | \tag{5}$$

$$dCt(c_i, c_j) = \sqrt{(\overline{x}_i - \overline{x}_j)^2 + (\overline{y}_i - \overline{y}_j)^2} \tag{6}$$

$$d\mu(c_i, c_j) = | \mu_i - \mu_j | \tag{7}$$

We define the overall distance between c_i and c_j as follows:

$$D(c_i, c_j) = w_1 dCc(c_i, c_j) + w_2 dPt(c_i, c_j) + w_3 dCt(c_i, c_j) + w_4 d\mu(c_i, c_j) \tag{8}$$

w_i is the weight assigned to the corresponding distance. We have assigned different weight to each distance, which is shown in Table 1 in Appendix. According to the weight in each group of Table 1, the performance of image retrieving is evaluated and the result is presented in Fig. 4. From Fig. 4, we can see that the weights assigned the values $w_1 = 0.4$, $w_2 = 0.4$, $w_3 = 0.15$, $w_4 = 0.05$, achieve the best retrieving performance in our experiments. From analysis of Fig. 4, we can see the color feature (the dominant color and its area percentage) is still the most significant part in defining the distance, and the centroid also has more obvious influence on retrieval precision when compared with the second-order central moment.

$D(c_i, c_j)$ is a normalized value so that the value of similarity between c_i and c_j can be defined as:

$$Sim(c_i, c_j) = 1 - D(c_i, c_j) \tag{9}$$

4.2 Optimal Matching

Given two images, a query image A and a target image B , each of them has the dominant color set $C^a = \{c_1^a, c_2^a, \cdots c_K^a\}$ and $C^b = \{c_1^b, c_2^b, \cdots c_K^b\}$ respectively, where K is the number of dominant colors of each image. In order to compute the similarity of the two images, we first have to search the optimal matching dominant colors between the two dominant color sets C^a and C^b.

We use the optimal matching method in graph theory [12] to solve the problem. We construct the bipartite graph as $G = \{C^a, C^b, E\}$, where C^a and C^b are dominant color sets of two images. $E = \{e_{i,j}\}$ is the edge sets, where a weight $w_{i,j}$ is assigned to the edge $e_{i,j}$ in G. $w_{i,j}$ is the value of similarity between two dominant colors c_i^a and c_j^b, computed by formula (9). Given the weighted bipartite graph G (An example is shown in Fig.1), the *Kuhn-Munkres* algorithm [12] can be used to solve the optimal matching problem. This algorithm has been applied in some research such as content-based video retrieval [16] and document similarity search [17]. The computational complexity of *Kuhn-Munkres* algorithm is $O(K^3)$. Based on the optimal matching theory, the similarity measure of the query image and the target image can be computed by the sum of all distances between every matched pair of dominant colors. Then the retrieval result is ranked according to the value of similarity.

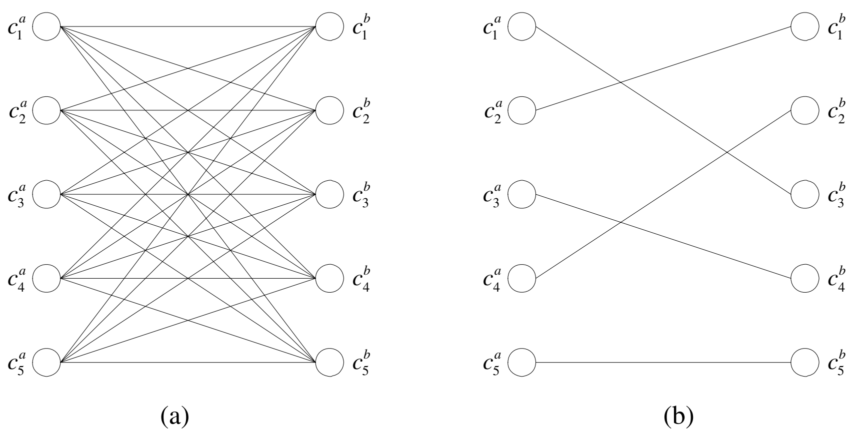

Fig. 1. (a) A bipartite graph of two dominant color sets. (b) The optimal matching result.

5 Experimental Results

We have developed a content-based image retrieval system called PKUQBIC to validate the efficiency of proposed algorithms and techniques in our paper. The image database consists of 4000 images, distributed into 28 different categories. We present the retrieval result of the proposed algorithm in this paper and compare it with other two clustering based algorithms [3] and [4]. The proposed algorithm in this paper is called *CSOP* (Color-Spatial Optimal Matching). From Fig. 2 we can see the method proposed in this paper is well defined, and achieves much better retrieving results than the other two methods.

(a) Retrieval results of car image

(b) Retrieval results of flower image

Fig. 2. Retrieval results of three methods, with (1) Proposed method in [3], (2) Proposed method in [4], (3) *CSOP* method

We also use average retrieval rate (ARR) and average normalized modified retrieval rank (ANMRR) [15] to evaluate the performance of our proposed technique in the 4000-image database, which is shown in Fig.3. ARR and ANMRR are the evaluation criterions used in all of the MPEG-7 color core experiments [15]. ANMRR measure coincides linearly with the results of subjective evaluation about retrieval accuracy. To get better performance, ARR should be larger and ANMRR should be smaller. We also give the ARR and ANMRR evaluation of the two methods [3] and [4] in order to compare them with *CSOP* algorithm. From Fig.3 we can see that *CSOP* gets a significant improvement in retrieval performance compared with the other two

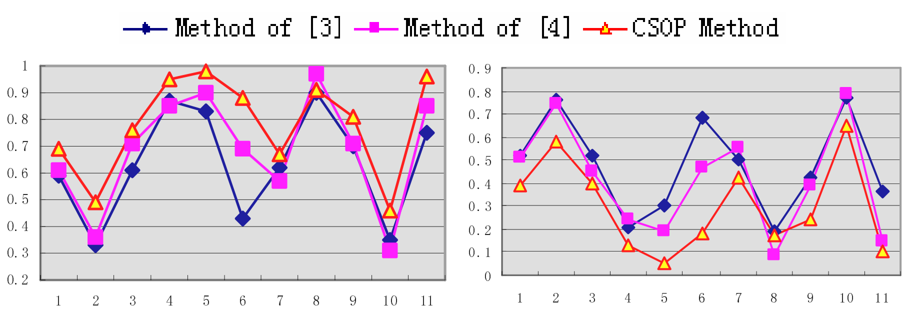

Fig. 3. ARR performance (left) and ANMRR performance (right) of the three methods

methods. The horizontal axes in Fig.3 (including Fig.4) denote corresponding image category, listed as: 1-fruit, 2-cup, 3-building, 4-sky, 5-face, 6-car, 7-hill, 8-fire, 9-bird, 10-dog, 11-sea, and the vertical axes denote the ARR and ANMRR performance.

6 Conclusion

Along with the fact that visual features have a significant correlation with semantic information of image, this paper proposes an ant colony clustering scheme to extract the dominant color features that well match human perception of images. Spatial feature combined with the dominant color feature is taken into account to measure the similarity. Besides we develop a perceptually based image similarity metric based on optimal dominant color matching algorithm, which is used to search the optimal matching pair of dominant color sets of any two images. The future work is to extend the proposed algorithm *CSOP* to include other spatial information such as the texture feature or shape feature to measure similarity, and a larger image database should be employed to evaluate the performance of the proposed scheme.

Acknowledgements

The authors would like to thank Mr. Yuxin Peng from Institute of Computer Science & technology of Peking University and Ms. Jianying Hu from IBM T.J. Watson Research Center, for their advice and help during the research.

References

1. J. R. Smith, S. F. Chang, A fully automated content-based image query system, Proceedings of ACM Multimedia, pp. 87－98, 1996
2. L. Zhang, F. Lin, and B. Zhang, A CBIR method based on color-spatial feature, IEEE Regin10 Annual International Conference, pp. 166-169, 1999
3. W.Y. Ma, Y. Deng, and B.S. Manjunath, Tools for texture/color base search of images, Proc. SPIE, vol. 3016, pp. 496–505, 1997
4. A. Mojsilovic, J. Hu, and E. Soljanin, Extraction of perceptually important colors and similarity measurement for image matching, retrieval, and analysis, IEEE Trans. Image Processing, vol. 11, pp. 1238-1248, Nov. 2002
5. A.Mojsilovic, J.Kovacevic, J.Hu, etc., Matching and retrieval based on the vocabulary and grammar of color patterns, IEEE Trans. Image Processing, vol. 9, pp. 38–54, Jan. 2000
6. B. Rogowitz, T. Frese, J. Smith, C. A. Bouman, and E. Kalin, Perceptual image similarity experiments, Proc. SPIE, 1997
7. J.L. Deneubourg, S. Goss, N. Frank, The dynamics of collective sorting: robot-like ants and ant-like robots, Proc. Of the 1st Int. Conference on Simulation of Adaptive Behavior: From Animals to Animats, MIT Press/ Bradford Books, Cambridge, MA, pp.356-363, 1991
8. E. Lumer and B. Faieta, Diversity and adaption in populations of clustering ants, In Proc. of the 3rd International conference on Simulation of Adaptive Behaviour: From Animals to Animats 3, MIT Press, Cambridge, MA, pp. 501−508, 1994
9. R.S. Parpinelli, H.S. Lopes, and A.A. Freitas, Data mining with an ant colony opti-mization algorithm, IEEE Trans. on Evolutionary Computing, vol. 6, pp. 321–332, Aug. 2002
10. Yap, P.- T., Paramesran, R., Seng-Huat Ong, Image analysis by Krawtchouk moments, IEEE Transactions on Image Processing, Vo. 12, pp. 1367 – 1377, 2003

11. Kenneth R. Castleman, Digital Image Processing, Prentice-Hall International, Inc. 1996
12. L. Lovász and M. D. Plummer, Matching Theory. Amsterdam, The Netherlands: North Holland, 1986
13. Wei Jiang, Guihua Er, and Qionghai Dai, Multilayer semantic representation learning for image retrieval, International Conf. on Image Processing, vol. 4, pp. 2215-2218, 2004
14. Feng Jing, Mingjing Li, Hong-Jiang Zhang, and Bo Zhang, A unified framework for image retrieval using keyword and visual features, IEEE Trans. on Image Processing, vol. 14, pp. 979-989, 2005
15. B.S Manjunath, J-R Ohm, V.V. Vasudevan, and A. Yamada, Color and texture descritors, IEEE Trans. Circuits Syst. for Video Technol., Vol. 11, pp. 703-715, June 2001
16. Peng Y.X, Ngo C.W, Dong Q.J, Guo ZM, Xiao JG, An approach for video retrieval by videoclip, Journal of Software, 14(8):1409~1417, 2003
17. X. J. Wan, Y. X. Peng, A new retrieval model based on texttiling for document similarity search[J], Comput. Sci. & Technol, VOL.20, NO.4, 2005

Appendix

Different weights are used to combine the color feature and spatial feature to compute the distance between any two images. We construct Table 1, in which five typical weight groups (T1, T2, T3, T4, T5) are assigned to coordinate the color feature and spatial feature. Retrieving performance of *CSOP* assigned with the five weight groups is evaluated using the 4000-image database of PKUQBIC, shown in Fig. 4. We can see weight group T3 achieves the best retrieving performance. The appendix shows that using combined features performs better than using either mainly spatial feature or only color feature.

Table 1. Five typical weight groups

	w_1	w_2	w_3	w_4
T1	0.1	0.1	0.4	0.4
T2	0.3	0.3	0.2	0.2
T3	0.4	0.4	0.15	0.05
T4	0.4	0.4	0.1	0.1
T5	0.5	0.5	0	0

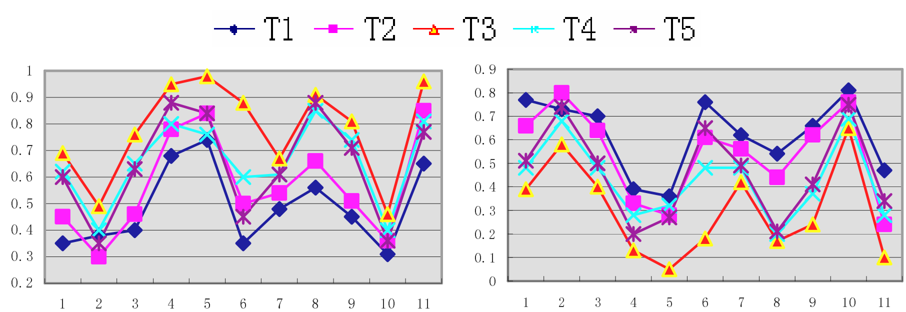

Fig .4. ARR performance (left) and ANMRR performance (right) according to different weights assigned in Table 1