A New Image Retrieval Method Based on Optimal Color Matching

Xin Huang, Shijia Zhang, Guoping Wang, Heng Wang
Human-Computer Interaction & Multimedia Lab,
Computer Science Department, Peking University,
Beijing, P. R. China 100871

Abstract - *Color is one of the most important low-level features used in image analysis and retrieval. In this paper we introduce an algorithm for extracting compact color feature in the CIELAB color space. Our method is based on an ant colony clustering algorithm, which helps us develop a perceptually based image similarity metric. We propose a dominant color and area percentage model, which provides a very compact color feature representation of images. Besides image similarity metric is developed based on dominant colors matching. The algorithm introduced in this paper is well suited for creating small color descriptors and is efficient. It is suitable for image representation, matching and retrieval. It also has some important advantages for content-based image retrieval system, such as great reduction of the feature file sizes needed for storage and indexing.*

Keywords: Ant-colony clustering, dominant color set, optimal matching.

1 Introduction

Colors [1], textures [9], and shape of objects [10] are widely used low-level features in content-based image retrieval systems. Color histogram, due to its good performance in characterizing the global color information, is widely used in many image retrieval systems. Histogram Intersections, proposed by Swain [1] in 1991, was a stable and efficient technique for matching model and image histograms. Smith [2] proposed a method to quantize colors into 166 bins in the HSV color space, with 18 Hues, 3 Saturations and 3 Values, plus 4 grays. Zhang [3] gave a new dividing method to quantize the color space into 36 non-uniform bins, which is more reasonable to the human vision model. Although the color histogram has shown to be an effective method for retrieving based on the entire image property, the large volume of feature vectors makes the storage and indexing problem more complicated. It has been observed that the color quantization schemes commonly used in computing the color histogram, such as the above two methods, have a major and common drawback. That is, neglecting the similarity of the colors near the boundary of the bins, so that similar colors might be quantized to different bins in the histogram, thus increasing the possibility of retrieving dissimilar image.

Besides color histogram, another commonly used method is to apply clustering based technique to quantize the color space. Ma et al. utilized a vector quantization called Generalized Lloyd algorithm (GLA) [4] to quantize the RGB color space, in order to minimize the mean square error and generate a compact color feature representation. But it has two drawbacks: the number and position of the initial clustering centers must be set before the algorithm starts; another problem is that it usually produces the local optimal clustering result instead of the global optimal result. Mojsilovic [5] proposed a new quantization scheme in the *Lab* space, that is firstly sampling the luminance axis into N_L levels, then for each discrete luminance value, choose N_P colors in the corresponding (a,b) as points z_n of a hexagonal spiral lattice in the complex plane. Wan [7] proposed a new approach to image retrieval with hierarchical color clustering scheme based on a pruned octree color representation.

The problem of how to extract semantic information from the image still remains the biggest obstacle in the content-based image retrieval system [18], [19]. In [8] Rogowitz performed psychophysical experiments analyzing human perception of image content, showing that visual features have a significant correlation with semantically relevant information. Mojsilovic [5] indicated that even with the absence of semantic cues, “semantically correct retrievals” 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 small color descriptor, which is suitable for image representation, matching and retrieval.

The contribution of this paper is that we propose a new ant-based clustering technique, which models the behavior of ants’ collecting corpses and is self-organizing, to extract perceptually dominant colors as the foundation for image matching. Corresponding similarity metric is established by using an optimal matching algorithm in graph theory. In addition to this, we also address the issue of GLA quantization scheme used in [4]. That is avoiding clustering getting into local optimality and the sensibility to initial clustering centers. It also has better average retrieval rate (ARR) and average normalized modified retrieval rank (ANMRR) compared with other techniques shown in section 5, according to our experiment in the image retrieval system PKUQBIC.

2 Basic model of ant colony algorithm

Ant colony algorithm has been proposed and applied in various areas since 1990s [11, 12, 13, 14, 15, 16]. It is a distributed, heuristic searching algorithm, which models the behavior of ants' prowling. Many achievements have been attained when ant colony algorithm applied in solving complex combinatorial optimization questions, such as Traveling Salesman Problem, Job Shopping Scheduling Problem, and Graph Coloring Problem. Another type of ant colony algorithm is ant-based clustering, which models the ants' behavior of piling corpses [11]. Researchers found that the ants can assemble the ant corpses into several piles in their studies. If there is an ant corpse pile, it will attract the other ants to put more and more corpses in the pile. So a few bigger corpse piles are constructed in the end. Deneubourg [11] proposed a basic model that explains the ants' behavior of piling corpses, which is commonly called BM (Basic Model) 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 it 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: unloaded ants are likely to pick up data items that are either isolated or surrounded by dissimilar ones; loaded ants tend to drop them in the vicinity of similar ones. Two functions [11] define the probabilities of picking up and dropping an object in the following.

The probability that an ant which is not carrying an object will pick up an object:

$$P_{pickup} = \left(\frac{K^+}{K^+ + f}\right)^2 \qquad (1)$$

The probability that an ant which is carrying an object will drop an object:

$$P_{drop} = \left(\frac{f}{K^- + f}\right)^2 \qquad (2)$$

f is defined as the fraction of neighboring sites occupied by objects of the same type, and K^+, K^- are threshold constants.

Lumer and Faieta [12] proposed LF algorithms, which improve Deneubourg's basic model by introducing a measure of dissimilarity between data objects. The dissimilarity d between objects in the attribute space of the object is defined as: if two objects o_i and o_j are similar objects, then $d(o_i, o_j) = 0$; if o_i and o_j are highly dissimilar objects, then $d(o_i, o_j) = 1$. These objects can be extended as points in the R_n space and $d(o_i, o_j)$ as the Euclidean distance between them.

In Lumer's algorithms, all data items are randomly scattered on the toroidal grid, each ant is placed at a random position and randomly picks up one data item. They proposed the following density measure for a data item o_i:

$$f(o_i) = \begin{cases} \dfrac{1}{s^2} \sum\limits_{O_j \in Neigh(r)} \left[1 - \dfrac{d(o_i, o_j)}{\alpha}\right] & if \quad f(o_i) > 0 \\ 0 & otherwise \end{cases} \qquad (3)$$

It means at time t, an ant at the site r finds an object o_i, $f(o_i)$ measures the mean similarity of the object o_i with the other objects that are in its neighborhood with perception area. The parameter α defines the dissimilarity scale, and the parameter s is the radius of the surrounding region. The sum in this expression extends over all surrounding objects o_j in the local area of element o_i, $d(o_i, o_j)$ measures the distance between a pair of data elements (o_i, o_j) scaled by α.

3 Dominant color features extraction

3.1 Extract dominant colors

Firstly an input image is transformed into *CIELAB* color space [17]. Then we get the training sequence consisting of M source vectors: $T = \{x_1, x_2, \ldots, x_M\}$. The source vector that is 3-dimensional consists of L, a, b value in CIELAB color space, for example, $x_m = \{L_m, a_m, b_m\}$, $m = 1, 2, \ldots, M$.

Then we utilize the ant colony clustering algorithm that is described in section 2 to extract the dominant colors from the training sequence T. The first step is to randomly project training sequence T onto a plane, then a few virtual ants are generated, randomly placed on the plane. The behavior of an individual ant is similar to that of the ant in the BM mentioned in section 2. Then the density measure of each ant is computed. Each ant acts according to its current state and corresponding probability. Finally several clustering centers are visually formed through the ants' collective actions.

The formula of computing density measure is the same as formula (3). In the following dominant colors clustering algorithm we proposed another two probability conversion functions for simplicity. The two modified versions of probability conversion functions are shown as formula (4) and (5).

Probability of an unloaded ant picking up an object is:

$$P_{pickup}(o_i) = \begin{cases} 1.0 & if \quad f(o_i) \le 1.0 \\ \dfrac{1}{f(o_i)^2} & else \end{cases} \quad (4)$$

Probability of a loaded ant dropping an object is:

$$P_{drop}(o_i) = \begin{cases} 1.0 & if \quad f(o_i) \ge 1.0 \\ f(o_i)^2 & else \end{cases} \quad (5)$$

The procedure of the algorithm is shown in the following.

Step 1: Initialize parameters, set Cycle_Count as i , Max_Cycles as *N*.

Step 2: Project training sequence T onto a plane, give each ant an initial object, initial state of each ant is randomly set.

Step 3: For all the ants, begin clustering actions such as picking up or dropping objects.

Step 4: Compute density measure $f(o_i)$ of each ant by formula (3).

Step 5: **(1)** If the ant is loaded, compute the dropping probability P_{drop} by formula (5). Compare P_{drop} with a random generated probability P_{random} . If $P_{drop} > P_{random}$, then the ant drops the object and the coordinates of the ant are given to the object, the state of the ant becomes unloaded; else a pair of randomly generated coordinates is given to the ant.

(2) If the ant is unloaded, compute the pick-up probability P_{pickup} by formula (4). Compare P_{pickup} with a random generated probability P_{random}. If $P_{pickup} < P_{random}$, the ant doesn't pick up the data object, another data object is randomly given to the ant; else the ant picks up the data object, the state of the ant becomes loaded, and a new pair of randomly generated coordinates is given to the ant.

Step 6: The Cycle_Count $i = i+1$, if $i > N$ or no object is dropped or picked-up during Step5, the algorithm is ended with a few clustering dominant colors generated, else go to Step3 and the algorithm continues.

3.2 The clustering result of dominant colors

After using the ant colony clustering algorithm, we extract the dominant colors codebook denoted as $C = \{c_0, c_1, \cdots c_{N-1}\}$, where N is the number of dominant colors of the image and each dominant color $c_i = \{L_i, a_i, b_i\}$ is a 3-dimensional *LAB* color vector, which is the centroid of the colors belonged to c_i . In our experiment N is set as 8, which is sufficient for image matching compared with other histogram based method. This indicates that our method largely reduces the feature file sizes needed for image storage and indexing. Finally the percentage of each dominant color is calculated and the dominant color feature vectors are denoted as:

$$f_c = \{(I_j, P_j) \mid 1 \le I_j \le N, 0 \le P_j \le 1\} \quad (6)$$

where I_j is the index in the dominant colors codebook C , P_j is the corresponding area percentage with $0 \le P_j \le 1, \sum_{1 \le j \le N} P_j = 1$.

4 Color similarity measure

4.1 Distance measure

The dominant color $\{c_i\}$ and the corresponding area percentage described above allow the definition of a color similarity measure closely match the human perception. The basic principle is that similarity between two images in terms of color composition should be measured by a combination of color and area differences [6]. Based on human visual system, for two images to be considered similar in terms of color features, two conditions need to be satisfied. First, the colors of dominant color elements of the two images need to be similar. Second, the color components with similar colors need to have similar area percentage.

Given two images, a query image A and a target image B with N_A and N_B dominant colors respectively, feature vectors of the two images are denoted as $f_c(a) = \{(I_a, P_a) \mid 1 \le a \le N_A\}$ and $f_c(b) = \{(I_b, P_b) \mid 1 \le b \le N_B\}$. Then we define the Euclidean distance measure between any two dominant colors from the dominant colors codebook C as:

$$W(I_a, I_b) = \left\| C_{I_a} - C_{I_b} \right\| = \sqrt{(L_{I_a} - L_{I_b})^2 + (a_{I_a} - a_{I_b})^2 + (b_{I_a} - b_{I_b})^2} \quad (7)$$

Use the formula (8) as the distance measure of color features.

$$D((I_a, P_a), (I_b, P_b)) = | P_a - P_b | + W(I_a, I_b) \quad (8)$$

Note that the distance between the dominant color feature is defined as the sum of the distance in terms of the area percentage and the distance in the Lab color space.

The $D((I_a, P_a), (I_b, P_b))$ can be pre-computed and stored in a table. Supposing that (I, P) is one of the dominant color elements of image A , we need to identify the best matching color $k(I, B)$ from the dominant color feature vectors $\{(I_b, P_b) \mid \forall b \in [1, N_b]\}$ of image B.

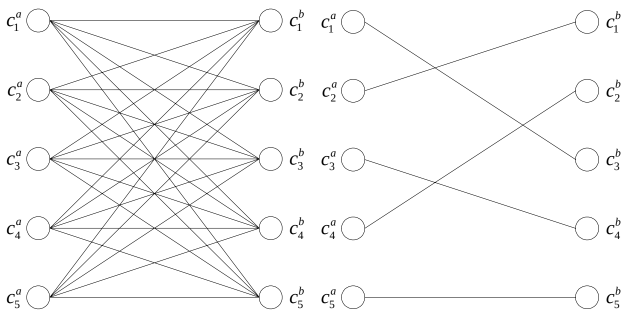
$D((I_a, P_a),(I_b, P_b))$ is a normalized value so that the value of similarity between c_i and c_j can be defined as:

$$Sim((I_a, P_a),(I_b, P_b)) = 1 - D((I_a, P_a),(I_b, P_b)) \quad (9)$$

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 [20] to solve the problem. We construct the bipartite graph as $G = \{C^a, C^b, E\}$, where C^a and C^b are dominant colors sets of the 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 [20] can be used to solve the optimal matching problem. This algorithm has been applied in some research such as content-based video retrieval [21] and document similarity search [22]. 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 can be ranked according to the value of similarity.



(a) (b)

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

5 Experimental results and analysis

We have developed a content-based image retrieval system called PKUQBIC to validate the efficiency of proposed algorithms and techniques in this paper. Our experimental image database consists of 4000 natural images, distributed into 28 different categories. We present the retrieval examples of our scheme and compare it with two previously proposed algorithms in [3] and [4], shown in Fig. 2. From Fig. 2(a) we can see the histogram based method [3] retrieves some dissimilar images because they neglect similarity of colors near quantizing boundary, thus similar colors are quantized into different bins that may lead to false retrieval. GLA quantization algorithm [4] is sensitive to the initial clustering centers, and may get clustering into local optimality, just as Fig. 2(b) shows. The ant based clustering technique of extracting perceptually dominant colors is self-organizing, and provides a very compact color feature representation of image. From the retrieving results Fig. 2(c) we can see that our scheme is well defined, and achieves much better retrieving results than the other two methods.

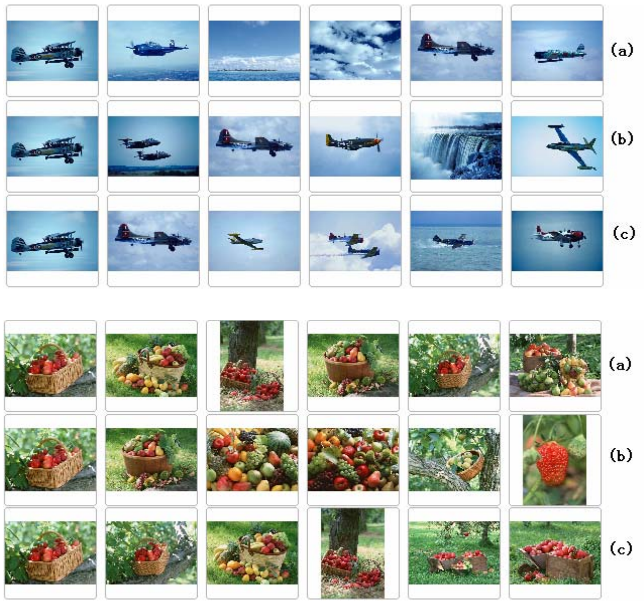


Fig.2. Retrieving examples of three methods, the query sample is the first image of each group. (a) The proposed method in [3], (b) The proposed method using GLA algorithm in [4], (c) The proposed method in this paper

We also use average retrieval rate (ARR) and average normalized modified retrieval rank (ANMRR) [23] to evaluate the performance of our proposed algorithm in the 4000-image database of PKUQBIC. ARR and ANMRR are the evaluation criterions used in all of the MPEG-7 color core experiments [23]. ANMRR measure coincides linearly with the results of subjective evaluation about retrieval accuracy. We also give the ARR and ANMRR evaluation of the two previously proposed methods in order to compare with our scheme, shown in Fig.3. According to the definition of ARR and ANMRR, the ARR should be larger and the ANMRR should be smaller in order to get better performance. Fig.3 shows that the proposed method in this paper gets a significant improvement in retrieval performance compared with the other two methods. The horizontal axes in Fig.3 denote corresponding image category, listed as: 1-cup, 2-building, 3-fruit, 4-face, 5-car, 6-hill, 7-fire, 8-bird, 9-dog, 10-sea.

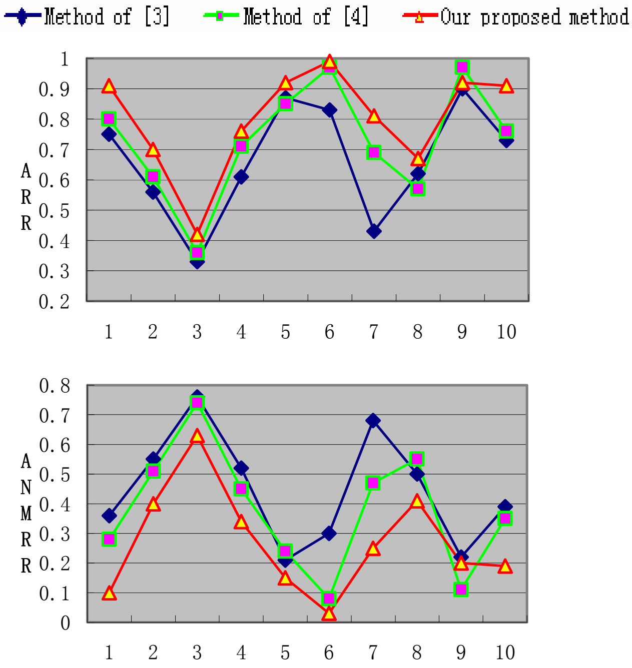


Fig.3. ARR performance (up) and ANMRR performance (down) of the three methods

6 Conclusions

Based on the fact that visual features have a significant correlation with semantic information of image, this paper propose an ant colony clustering scheme, which has the eminent property of self-organizing, to extract the dominant color features that are well matching human perception of images. Besides we develop a perceptually based image similarity metric based on dominant colors matching. The optimal matching algorithm is used to search the optimal matching results of the dominant colors sets of any two images. The proposed method in this paper also has some important advantages in image retrieval system, such as great reduction of image feature file sizes needed for storage and indexing. Because the perceptually dominant colors are well clustered, the proposed techniques can be easily extended to include the texture features or spatial information to measure similarity, which is our research direction in the near future.

Acknowledgement

The research was supported by No. 2004CB719403 from The National Basic Research Program of China (973 Program), No. 60473100_No, 60573151 from National Natural ScienceFoundation of China. 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 this research.

7 References

[1] M.Swain and D.Ballard, “Color indexing,” IEEE *Int. J. Comput. Vision*, vol. 7, pp. 11-32, 1991.

[2] J.R. Smith and S.F. Chang, “A fully automated content-based image query system,” *Proceedings of ACM Multimedia*, pp. 87—98, 1996.

[3] L. Zhang, F. Lin, and B. Zhang, “A CBIR method based on color-spatial feature,” *IEEE Regin10 Annual International Conference*, pp. 166—169, 1999.

[4] 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.

[5] 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.

[6] A.Mojsilovic, J.Kovacevic, J.Hu, R.J.Safranek, and K.Ganapathy, “Matching and retrieval based on the vocabulary and grammar of color patterns,” *IEEE Trans. Image Processing*, vol. 9, pp. 38–54, Jan. 2000.

[7] X. Wan and C.-C. J. Kuo, “A new approach to image retrieval with hierarchical color clustering,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 628–643, Sept. 1998.

[8] B. Rogowitz, T. Frese, J. Smith, C. A. Bouman, “Perceptual image similarity experiments,” *Proc. SPIE*, 1997.

[9] L. Kokare, P.K. Biswas, and B.N. Chatterji, “Rotation invariant texture features using rotated complex wavelet for content based image retrieval,” *Int. Conf. on Image Processing*, vol. 1, pp. 393-396, Oct. 2004.

[10] O. El Badawy and M. Kamel , “Shape retrieval using concavity trees,” *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 3, pp. 111-114, 2004.

[11] J.L. Deneubourg, S. Goss, and N. Frank, “The dynamics of collective sorting: robot-like ants and ant-like robots,” *Proc. Of the 1st International Conference on Simulation of Adaptive Behavior*: *From Animals to Animats*, MIT Press/ Bradford Books, Cambridge, MA, pp.356-363, 1991.

[12] 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.

[13] M. Dorigo, A. Colorni, and V. Maniezzo, “The Ant System: optimization by a colony of cooperating agents,” *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, vol. 26, pp. 29-41, 1996.

[14] R. S. Parpinelli, H. S. Lopes, and A. A. Freitas, “Data mining with an ant colony optimization algorithm,” *IEEE Transactions on Evolutionary Computing*, vol. 6,

pp. 321–332, Aug. 2002.
[15] T. Stützle and M. Dorigo, "ACO algorithms for the traveling salesman problem," In: *Evolutionary Algorithms in Engineering and Computer Science*, K. Miettinen, M. Makela, P. Neittaanmaki and J. Periaux (Eds.), New York, pp. 163-183, 1999.
[16] E. Bonabeau, M. Dorigo, and G. Theraulaz, "Swarm Intelligence – from natural to artificial systems," Oxford University Press, New York, NY, 1999.
[17] Michael W. Schwarz**,** William B. Cowan**,** and John C. Beatty, "An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models," *ACM Transactions on Graphics*, vol. 6, pp. 123-158, Apr. 1987.
[18] Wei Jiang, Guihua Er, and Qionghai Dai, "Multilayer semantic representation learning for image retrieval**,"** *International Conference on Image Processing*, vol. 4, pp. 2215-2218, 2004.
[19] Feng Jing, Mingjing Li, Hong-Jiang Zhang, and Bo Zhang, "A unified framework for image retrieval using keyword and visual features**,"** *IEEE Transactions on Image Processing*, vol. 14, pp. 979-989, 2005.
[20] L. Lovász and M. D. Plummer, Matching Theory. Amsterdam, The Netherlands: North Holland, 1986.
[21] 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.
[22] 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.
[23] MPEG video group. Description of Core Experiments for MPEG-7 Color / Texture Descriptors, ISO/MPEGJTC1/SC29/WG11 MPEG98/M2819, July. 1999.