# Content-Based Image Retrieval Incorporating the AHP Method

**WANG XIAOLING, XIE KANGLIN**

*Department of Computer Science & Engineering,*
*Shanghai JiaoTong University*

*Email:clisdy@126.com*

## Abstract

In this paper, a Content-Based Image Retrieval (CBIR) system is presented. In this system, a method called the Average Area Histogram (AAH) is proposed based on the area features of the regions formed by the pixels of each color, which retains the advantages of the conventional histogram. A key issue in CBIR is to develop a proper weight assignment method for combining various image features in retrieval. The Analytic Hierarchy Process (AHP) method is imported into this system for assessing the importance of various image features. Compared with the traditional experience-based linear assignment method, our assignment is more reasonable and easy to communicate. The experimental results demonstrate the feasibility and efficiency of our proposed scheme.

**Keyword**: Content-based Image Retrieval; Color Histogram; Spatial Information; Analytic Hierarchy Process; Weight Assignment

## 1. Introduction

Due to the rapid development of digital imaging and networking technologies, visual information has been widely used in many fields. How to retrieve this information efficiently has led to a rise of interest in techniques for retrieving images using information contained in image databases. In the past, various approaches to image retrieval have been proposed, most of which were Content-Based Image Retrieval (CBIR) that retrieves images by keywords or image contents such as color, texture, shape or any combination of these features. In recent years, many image retrieval systems, both commercial and research, have been developed such as QBIC, VisualSEEK and Blobworld, etc. QBIC is the first commercial content-based image retrieval system. It uses color, texture, shape, example images and sketches to retrieve images. In VisualSEEK, each image is automatically decomposed into regions by using regional colors and their relative locations. Query is based on the colors and locations of the regions provided by the user. Features used in Blobworld for querying are the color, texture, location, and shape of regions (blobs) and of the background.

Obviously, there is a need of weight assignment strategy in CBIR for combining various features in retrieval. Currently, the majority of weight assignment methods proposed in literature falls in one of the following approaches: the linear combination approach and the nonlinear approach. The coefficient of each feature used in the linear combination approach is either acquired by user assigning directly online or predetermined by the developers through iterative tests [1]. For example, in CIRES (http://amazon.ece.utexas.edu/~qasim/research.htm), users are required to give importance to perceptual grouping, color and texture. The default settings give equal weights to the three feature vectors. Users can set these to any value such that the total sums to 1. In the VIR Image Engine (http://www.virage.com), queries can be performed on various user-defined combinations of

primitives that denote a feature's type, computation and matching distance. However, it is not easy for a user especially an inexperienced user to decide proper weights of various features. Furthermore, the assignment results lack of theoretical instruction and it is difficult to communicate. The nonlinear approach (or the adaptive weight assignment) deals with the combination of various features in virtue of the Artifical Inteligence(AI) method such as neural networks [2], fuzzy logic [3] and genetic algorithm [4] etc. incorporating the Relevance Feedback (RF) technique from the Information Retrieval (IR) domain [5]. However, it is also difficult to explain where the weight results come from. Furthermore, RF has a drawback in terms of the heavy burden it imposes upon users in image retrieval [6]. In this paper, we study the linear assignment approach in CBIR. To facilitate the weight assignment and reflect the way people actually think in retrieval, we import the Analytic Hierarchy Process (AHP) method into image retrieval for assigning weights for various image features. AHP was developed in the 1970's by Dr. Thomas Saaty [7] and continued to be one of the most widely used decision-making theories. In virtue of the AHP method, users only need to conduct a pairwise comparison of image features in a user-friendly interface, whereby the corresponding weights can be calculated automatically. Compared with the traditional linear weight assignment method introduced above, the advantage of using AHP in image retrieval is that the weights results are more reasonable and easy to explain theirs origins.

Color is one of the most salient and commonly used features in image retrieval. There are three major types of color representations: color moments [8], the color histogram [9] and color sets [10]. Among these, the color histogram is most popular for its effectiveness and efficiency. The advantage of using the histogram in image retrieval is its robustness to rotation and scaling of the image content [11]. However, the color histogram is subject to spatial information loss and therefore incorrectly matches the images with same colors but different color distributions [12]. Many researchers have investigated this problem by integrating spatial information into the conventional color histogram. Pass and Zabih [13] propose a concept called the Color Coherence Vector (CCV) to split the histogram into two parts: a coherent one and a non-coherent one depending on the sizes of their connected components. Cinque et al. [14] present a spatial-chromatic histogram considering the position and variances of color blocks in an image. Rickman and John Stonham [15] define the color tuple histogram. They use a predefined equilateral triangle of a fixed length and then randomly move the triangle over an image to calculate the frequency of each tuple of triple pixels. Hsu et al. [16] modify the color histogram by first selecting a set of representative colors and then analyzing the spatial information of the selected colors using maximum entropy quantization with an event covering method. All the above methods attempt to improve the retrieval performance by integrating the spatial information into the color histogram. However, the way of extracting colors from the image spoils robustness to rotation and translation of the conventional histogram as seen with the position of color block adopted by Cinque et al. [14], and the shape and size of the predefined triangle used by Rick Rickman and John Stonham [15]. Therefore these improvement methods spoil the merit of the conventional histogram. For most of the images mismatched by the conventional histogram, we notice that the pixels of each color usually form several disconnected regions with different sizes, and this fact can be used as a key to distinguish these images. We believe that these regions could help to integrate spatial information into the conventional histogram. In this paper, we present an improvement on the traditional histogram called the Average Area Histogram (AAH) based on the area features of the regions formed by the pixels of each color, which retains the advantages of the conventional histogram.

Section 2 introduces the details of the Average Area Histogram (AAH). Section 3 and section 4 describe the shape and texture representations respectively. In section 5, the application of AHP in

image retrieval is presented. The normalization of features is introduced in Section 6 and section 7 gives the image matching method. Experiments are shown and discussed in Section 8 followed the conclusion made in section 9.

## 2. Color Representation: the Average Area Histogram (AAH)

Given a color space $\boldsymbol{C}$, the conventional color histogram $\boldsymbol{H}$ of image $\boldsymbol{I}$ is defined

$$\boldsymbol{H}_C(\boldsymbol{I}) = \{N(\boldsymbol{I}, \boldsymbol{C}_i) \mid i \in [1,...,n]\}, \quad (1)$$

Where $N(\boldsymbol{I}_i, \boldsymbol{C}_i)$ is the number of pixels of image $\boldsymbol{I}$ that fall into cell $\boldsymbol{C}_i$ and $i$ indicates the color levels of color space $\boldsymbol{C}$. $\boldsymbol{H}_C(\boldsymbol{I})$ shows the proportion of pixels of each color within the image. According to the definition of the conventional color histogram, it is only dependent on summations of pixel of each color and ignores the spatial distributions of colors completely. However, there exist different spatial distributions for each color as shown in Fig.1.

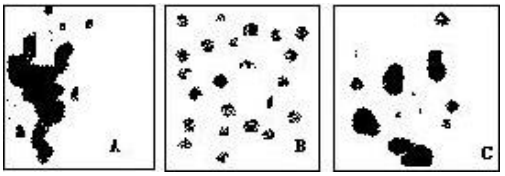



**Fig. 1 Regions formed by pixels with identical color**

Fig.1 illustrates three common cases of identical color where the pixels form connected and disconnected regions. For the three cases, suppose that the total number of the pixels is equal or close. Evidently, color distributions in the three cases are very distinctive. However, their corresponding color buckets in the traditional color histogram are equal or close. We notice that the sizes of the disconnected regions in the three cases are different and can be used as a key to distinguish the three cases correctly. We propose the Average Area Histogram (AAH) based on this idea. In the AAH, we integrate the area features of regions formed by each color into the conventional histogram.

Let $D$ be the number of disconnected regions formed by the pixels in cell $\boldsymbol{C}_i$ of image $\boldsymbol{I}$. We define

$$D(\boldsymbol{I}, \boldsymbol{C}_i) = \{N(\boldsymbol{C}_i, m) \mid i \in [1,.., n]\}, \quad (2)$$

Where $N(\boldsymbol{C}_i, m)$ counts the number of disconnected regions formed by the pixels in each cell $\boldsymbol{C}_i$ through an m-connectivity operation over image $\boldsymbol{I}$. Either an 8-connectivity operation or a 4-connectivity operation is optional. In this paper, we adopt the former. The Average Area Histogram $\boldsymbol{H}^*$ is then defined

$$\boldsymbol{H}^*_C(\boldsymbol{I}) = \{\frac{N(\boldsymbol{I}, \boldsymbol{C}_i)}{D(\boldsymbol{I}, \boldsymbol{C}_i)} \mid i \in [1,.., n]\}, \quad (3)$$

Where $N(\boldsymbol{I}, \boldsymbol{C}_i)$ counts the total number of pixels in cell $\boldsymbol{C}_i$ of image $\boldsymbol{I}$. Then $\boldsymbol{H}^*$ represents the average area of the disconnected regions formed by the pixels in each cell $\boldsymbol{C}_i$.

According to the average areas of the disconnected regions in cases Fig.1 A, Fig.1B and Fig.1C, the AAH can distinguish the three cases correctly. However, since the conventional histogram only calculates the summations of the pixels (summations of the areas of all the disconnected regions), the three cases will be regarded as identical and be matched incorrectly. In most of the mismatching cases

by the traditional histogram, it fails to distinguish case Fig.1A with case Fig.1B while the AAH outperforms the traditional histogram for distinguishing Fig.1A with Fig.1B and Fig.1A with Fig.1C.

Fig.2 shows two images (a) and (b) with substantially different content, and Fig.3 shows the traditional histogram and the AAH of images of Fig.2 in the gray-level space. Fig.3 (1) and Fig.3 (2) are the respective conventional histograms of image (a) and (b), and Fig.3 (3) and Fig.3 (4) are their respective AAH. Evidently, the AAHs of image (a) and (b) are more dissimilar compared with their traditional histograms, which is consistent with the different contents of image (a) and (b). In other words, the AAH is more suitable and effective for reflecting the real color distributions of images than the traditional histogram.


**Fig. 2 image a and b**

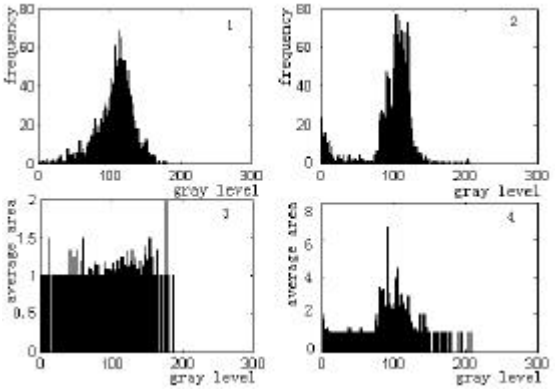



**Fig. 3 Traditional histogram and AAH of Fig.2a and Fig.2b**

According to the definition of AAH, we know that the average area is very sensitive to the number of connected regions. However, the very small regions which are composed of 1, 2, or 3 pixels etc., may not be significant in cell $\boldsymbol{C}_i$ and should be ignored. Suppose $A(\boldsymbol{C}_i, \boldsymbol{I}) = \{a_j \mid j = 1,...,k\}$ represents the areas of the disconnected regions by an m-connectivity operation over image $\boldsymbol{I}$, $k$ is the number of these regions. We define

$$\alpha = \frac{a}{A_{\max} - A_{\min}} \tag{4}$$

where $A_{\max}$, $A_{\min}$ are the respective maximum and minimum area of the disconnected regions in cell $\boldsymbol{C}_i$. Here, we assume that $A_{\max} \neq A_{\min}$. $\alpha$ reflects the contibution of a region in cell $\boldsymbol{C}_i$. In the experiment, we find $\alpha > 0.05$ is proper to filter those relative unimportant regions in each cell $\boldsymbol{C}_i$.

One of the most commonly used matching techniques for histogram is the Euclidian distance:

$$d(\boldsymbol{H}(\boldsymbol{I}), \boldsymbol{H}(\boldsymbol{Q})) = [\sum_{i=1}^{n} | \boldsymbol{H}(\boldsymbol{I}_i) - \boldsymbol{H}(\boldsymbol{Q}_i) |^2]^{1/2} \tag{5}$$

The histogram intersection was proposed by Swain and Ballard. Their objective was to find known objects within images using color histograms [9].

$$d(\boldsymbol{H}(\boldsymbol{I}), \boldsymbol{H}(\boldsymbol{Q})) = \sum_{i=1}^{n} \min(\boldsymbol{H}(\boldsymbol{I}_i), \boldsymbol{H}(\boldsymbol{Q}_i)) \tag{6}$$

Either the Euclidian distance or histogram intersection is optional for AAH because AAH reflects the real color distribution of an image. AAHs of two different images are distinctive enough themselves (see Fig.3) and therefore the distance metric has not vital influence on the measurement result of AAH. In this study, we select the Euclidian distance for AAH.

## 3. Shape Representation: Moment Invariants

Shape is one of the key visual features used by human for distinguishing image. In general, two of the most common approaches of shape representations are boundary-based and region-based [17]. The former employs shape boundary information and captures shape boundary features while the latter makes use of all the pixel information across the entire region. Fourier descriptors and moment invariants [18] are the respective typical approaches of the two categories.

Moment invariants are properties of connected regions in binary images that are invariant to translation, rotation and scale. In this paper, we adopt the moment invariants to describe shape features. Let $f(x,y)$ be a binary image of an object, we define its (p+q) $^{th}$ order moment:

$$m_{pq} = \sum_{x=1}^{M}\sum_{y=1}^{N} x^p y^q f(x,y) \tag{7}$$

Central moments are defined

$$\mu_{pq} = \sum_{x=1}^{M}\sum_{y=1}^{N} (x-\mu_x)^p (y-\mu_y)^q f(x,y) \tag{8}$$

Where $\mu_x = \frac{m_{10}}{m_{00}}$, $\mu_y = \frac{m_{01}}{m_{00}}$, p, q=0,1,2….

Central moments can be normalized as follows and become scale invariant as well as translation invariant.

$$\eta_{pq} = \frac{m_{pq}}{m^r_{00}}, \quad r = \frac{p+q+2}{2}, p+q = 2,3,\ldots, \tag{9}$$

Based on the second and third order central moments, seven invariants called the Moment Invariants can be used to represent shapes and objects. For the second and third order moments, we have the following six absolute orthogonal invariants:

(1) $\varphi_1 = \eta_{20} + \eta_{02}$

(2) $\varphi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta^2_{11}$

(3) $\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$

(4) $\varphi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$

(5) $\varphi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$
$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

(6) $\varphi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$

The seventh moments is a skew orthogonal invariants, it is useful in distinguishing mirror images.

(7) $\varphi_7 = (3\eta_{12} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$
$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

The Euclidean distance has been widely used to measure the similarity of shapes represented by moment invariants [19][20]. Given a vector of a binary image composed of the seven invariant moments $M = (\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6, \varphi_7)$, two shapes of query image $\boldsymbol{Q}$ and image $\boldsymbol{I}$ of the image database can be matched by the Euclidean distance.

$$d(\boldsymbol{M}(\boldsymbol{I}),\boldsymbol{M}(\boldsymbol{Q}))=[\sum_{i=1}^{7}(\varphi_{Ii}-\varphi_{Qi})^2]^{1/2}. \quad (10)$$

## 4. Texture Representation: Gray Level Co-occurrence Matrix (GLCM)

In computer vision, texture is defined as all what is left after color and local shape have been considered or it is defined by such terms as structure and randomness. A variety of approaches has been proposed for measuring texture similarity. Most approaches rely on comparing values of what are known as second-order statistics calculated from query and stored images. These methods calculate measures of image texture such as the degree of contrast, coarseness, directionality and regularity etc. Alternative methods of texture analysis for image retrieval include the use of Gabor filters (or Gabor wavelet) [21] and fractals [22].

Gray Level Co-occurrence Matrix (GLCM) is a popular statistical technique for extracting texture features from different types of images. It has been used in the analysis and classification of various types of images [23]. In this paper, we use the GLCM to represent texture features. For a gray-level image $f(x,y)$, the GLCM calculates how often a pixel with the intensity (gray-level) value $i$ occurs in a specific spatial relationship to a pixel with the value $j$ which locate $(Dx,Dy)$ apart. Fix a distance $\ddot{a}$ and an angle $\grave{e}$, its mathematical definition is:

$$p(i,j,\ddot{a},\grave{e})=\{(x,y)\mid f(x,y)=i, f(x+Dx,y+Dy)=j; x,y=0,1,2,...,L-1\} \quad (11)$$

Where $i,j=0,1,2,...,L-1$. L is the number of gray-levels of image; $\boldsymbol{d}$ is the distance between a pair of pixels. Given this matrix, there are a large number of features that can be deduced including energy, entropy, contrast and relativity etc. The calculation details can be found in [24].

In this study, we build GLCM for three directions: $\theta=0,45\ and\ 90$. In order to eliminate the influence of different directions, we need to compute the mean and standard deviation of the four above features for the three directions, and construct a vector $\mathbf{T}=(m_{\mathrm{E}},\boldsymbol{s}_{\mathrm{E}},m_{\mathrm{H}},\boldsymbol{s}_{\mathrm{H}},m_{\mathrm{I}},\boldsymbol{s}_{\mathrm{I}},m_{\mathrm{C}},\boldsymbol{s}_{\mathrm{C}})$ to represent the texture features of an image. The texture similarity of the query image $\boldsymbol{Q}$ and image $\boldsymbol{I}$ from the image database is:

$$\begin{aligned}d(\mathbf{T}(\mathbf{I},\boldsymbol{d},\boldsymbol{q}),\mathbf{T}(\mathbf{Q},\boldsymbol{d},\boldsymbol{q}))=\{&(m_{BI}-m_{BQ})^2+(\boldsymbol{s}_{BI}-\boldsymbol{s}_{BQ})^2+(m_{HI}-m_{HQ})^2+(\boldsymbol{s}_{HI}-\boldsymbol{s}_{HQ})^2+(m_{II}-m_{IQ})^2\\&+(\boldsymbol{s}_{II}-\boldsymbol{s}_{IQ})^2+(m_{CI}-m_{CQ})^2+(\boldsymbol{s}_{CI}-\boldsymbol{s}_{CQ})^2\}^{1/2}\end{aligned} \quad (12)$$

Where $m_E,\boldsymbol{s}_E$ are the mean and standard deviation of energy feature derived from the GLCM for three directions of query image $\boldsymbol{Q}$ and image $\boldsymbol{I}$ from the image database respectively, $m_H,\boldsymbol{s}_H$ the entropy feature, $m_I,\boldsymbol{s}_I$ the contrast feature and $m_C,\boldsymbol{s}_C$ the relativity feature.

## 5. Analytic Hierarchy Process (AHP)-Integration of Various Features of Image

Designed to reflect the way people actually think, Analytic Hierarchy Process (AHP) was developed in the 1970's by Dr. Thomas Saaty. It is suitable for complex decisions that involve the comparison of decision elements, which are difficult to quantify. AHP uses a multi-level hierarchical structure of objectives, criteria and alternatives. The pertinent data are derived by using a set of pairwise comparisons. These comparisons are used to obtain the weights of importance of the decision criteria and the alternatives in terms of each individual decision criterion [25]. Relative importance of two factors i and j with respect to criteria U is determined with the value $a_{ij}$ illustrated in table 1:

**Table 1 Pairwise comparison of factor relative importance**

| **Intensity of Importance ($a_{ij}$)** | **Definition** |
|---|---|
| 1 | i has the same importance as j with respect to U |
| 3 | i has slightly more importance than j with respect to U. |
| 5 | i has more importance than j with respect to U |
| 7 | i has a lot more importance than j with respect to U |
| 9 | i totally dominates j with respect to U |
| 2,4,6,8 | Intermediate values between the two adjacent judgments |

The judgment matrix $\boldsymbol{A}_{n\times n}$ is satisfied with the following natures:

(1) $a_{ii} = 1$

(2) $a_{ij} = (a_{ji})^{-1}$

(3) $a_{ij} = a_{ik} a_{kj}$

Psychological experiments have shown that individuals cannot simultaneously compare more than $7 \pm 2$ objects [26]. This is the main reason used by Saaty to establish 9 as the upper limit of his scale, 1 as the lower limit and a unit difference between successive scale values.

The concrete computation of the weights is an eigenvalue method (mathematical approach used by AHP)[6] as follows:

1) Compute the product of elements in each row of matrix $\boldsymbol{A}$: $M_i = \prod_{j=1}^{n} a_{ij} (i = 1,2,...n)$, where $n$ is the rank of the matrix $\boldsymbol{A}$.

2) Compute $V_i = \sqrt[n]{M_i}$, where $n$ is the rank of the matrix $\boldsymbol{A}$.

3) Normalize $v_i$ so as to obtain weight $\boldsymbol{W} = (w_1, w_2, ..., w_n)^T$, where $w_i = v_i / \sum_{i=1}^{n} v_i$.

We realize a CBIR system involved in color, shape and texture features incorporating the AHP method for determining weights of these features. We build a 3-level hierarchical structure in this way: the objective is to retrieve image effectively, the criteria is the content features, and the alternatives are color, shape and texture features. Since we just have one criterion, we only need to conduct a pairwise comparison of the importance of the three features for retrieving the desired image. Note the phrases to compare the features are suitable for users to describe their perceptions of image features. According to the comparison results, we obtain a judgment matrix $\boldsymbol{A}$.

$$\boldsymbol{A} = (a_{ij})_{n\times n} \tag{13}$$

Since $\boldsymbol{A}_{n\times n}$ is orthogonal and we only need to compare $n$ factors for $n$ ($n$-1)/2 times.

## 6. Normalization of Image Features

Elements in the shape and texture features have different physical meaning, in order to put an equal emphasis on each feature element within a feature vector, these features should be transformed into interval [-1,1] with the Gaussian normalization method as Eq.(14).

$$F' = \frac{F - m}{3\sigma} \tag{14}$$

Where $m$ and $\sigma$ are the mean value and the standard deviation of feature $F$ respectively. The $3\sigma$-rule ensures approximately 99% of the elements in feature $F$ can be transformed into interval [0,1]. For the values out of the range, convert them into -1(if less than -1) and 1(if greater than 1). The distance metrics of the three features return different values, for example the color difference computed by the histogram intersection is in the range of [0,1]. While the texture feature computed by the Euclidean distance may be not. Therefore all the content features should be mapped into range [0,1] so as to be equally important in the linear combination, and the conversion details can be found in [27].

## 7. Image Matching

By far, we have acquired the color, shape and texture representations of an image and their corresponding weights generated from AHP method. Two images can be matched in Eq. (15):

$$S(\boldsymbol{Q}, \boldsymbol{I}) = \omega_c d(\boldsymbol{H}^*(\boldsymbol{Q}), \boldsymbol{H}^*(\boldsymbol{I})) + \omega_s d(\boldsymbol{M}(\boldsymbol{Q}), \boldsymbol{M}(\boldsymbol{I})) + \omega_t d(\boldsymbol{T}(\boldsymbol{Q}, \ddot{a}, \grave{e}), \boldsymbol{T}(\boldsymbol{I}, \ddot{a}, \grave{e})) \tag{15}$$

Where $w_c, w_s, w_t$ are the weights computed by the AHP method for color, shape and texture features respectively.

## 8. Experiments

### 8.1 Image Preprocessing

(1) A color space is perceptually uniform if a small perturbation to a component value is approximately equally perceptible across the range of that value. The RGB color space does not exhibit perceptual uniformity. To make sense of the color comparison in AHP, we convert images from the RGB-space into the HSV-space, which is an approximately uniform color space. The details of conversion form RGB to HSV can be found in [28].
(2) In order to extract the shape features, region of an object is needed. Firstly, detect the edges of object with Sobel operator, then dilate the image with line structure element, finally fill the hole inside the object.
(3) To extract texture features, images are transformed into 8-gray level images. Four features are extracted from the moment matrix of images for $\theta = 0, 45$ *and* $90$ every 2 pixels ($\delta = 2$).

### 8.2 Experiment Results

The image database consists of 5180 color JPEG images. This includes 900 car images, 200 face images, 150 stamp image, 500 fruit images, 2000 animal images, 100 rock image, 150 leaf images, 450 flower images, 140 gun images, 50 mobile phone, 140 desk images and 400 plane images downloaded from the website: http://photo.ayinfo.ha.cn/lingleijp/index.htm and http://image.yisou.com/. Our query set consists of 68 query images involved in 17 categories (face, eagle, butterfly, dog, dolphin, goldfish, orangutan, parrot, sea-gull, peony, water lily, leaf, rhododendra, orange, gun, plane, and car). Each of the query images definitely has a group of correct answers based on both our judgments of these images and the website classification. The queries are specially chosen to so as to include images of different views of the same scene, large changes in appearance, etc.

### 8.2.1 Performance Evaluation

There are two principals to evaluate a retrieval system: the precision and the recall measures. Suppose $R(q)$ is the set of images relevant to the query image $q$ and $A(q)$ is the set of retrieved images. Precision is the fraction of retrieved images that are truly relevant to the query image:

$$P = \frac{|A(q)| \cap |R(q)|}{|A(q)|} \tag{16}$$

Recall is the fraction of relevant images that are actually retrieved:

$$R = \frac{|A(q)| \cap |R(q)|}{|R(q)|} \tag{17}$$

### 8.2.2 Experiments on the AAH

In the HSV-space, we compute the traditional histogram, AAH, and CCV using 64 color bins [13] for each image in the database. We adopt Eq.(5) as the distance metric for the traditional histogram, Eq.(6) for AAH and the city-block metric for CCV. Fig. 4 shows a comparison of the average precision-recall of the three histograms on our database. "TH" means the traditional histogram.

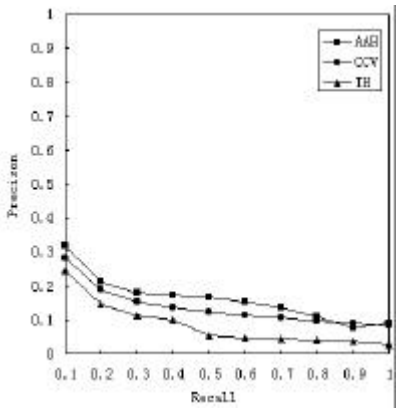



**Fig.4 P**erformance comparison of the three histograms

As shown in Fig.4, AAH outperforms both the conventional histogram and the CCV. Compared with the traditional histogram, the precision improvements of AAH at same recall rate are shown in table2.

**Table 2 precision improvements of AAH vs. TH at same recall rate**

| Recall | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|
| Precision improved (%) | 33 | 51.7 | 73.3 | 95.5 | 234 | 285 | 263 | 200 | 193 | 295 |

### 8.2.3 Experiments of the retrieval system incorporating the AHP

We extract color, shape and texture features for each image in the image database. For a specific query, according to his interpretation of the contribution of the three features for retrieving the query image, a user performs a pairwise comparison of these features by the interface designed in Fig.5. For three features, we only need compare them 3 times as mentioned in section 5. However, sometimes it is inconvenient for a user to compare feature A with B with the importance grade provided in table1,

but it may be easy to compare feature B with A. so in the interface, we design 6 pairwise comparisons for users. Then the corresponding feature weights can be calculated and the final matching can be performed through Eq.(15).

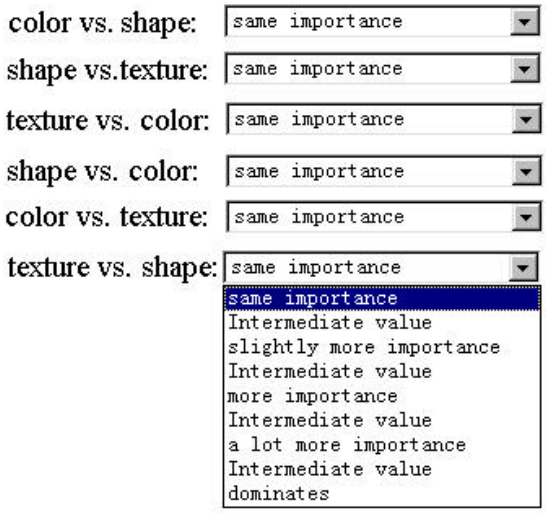



**Table 2 the comparison of features of a rhododendra image**

| | Texture | Color | Shape |
|---|---|---|---|
| **Texture** | 1 | 1/5 | 3 |
| **Color** | 5 | 1 | 7 |
| **Shape** | 1/3 | 1/7 | 1 |

**Fig. 5 Feature importance comparisons**

Next, we give one query example. For retrieving a rhododendra image at the upper-left corner of Fig.6, a user needs to complete the importance comparison of the three features through the user interface designed in Fig.5. One user may regard that color is the most important feature among the three features; texture the second important and shape the third important; color is more important than texture; color is a lot more important than shape; texture is slightly more important than shape. According to table 1, the corresponding importance matrix is shown in table 2 and the respective weights for texture, color and shape are 0.22, 0.73 and 0.05. To verify the efficiency of our improvement on the traditional histogram, we compare the retrieval performance of the three histograms with same weights for retrieving a same image. The reason why we choose a rhododendra image is that color is of the most importance among the three features.

**Fig. 6 Retrieval results using the traditional histogram (a), CCV (b), and the AAH (c)**

Fig. 6 shows the retrieval results of the traditional histogram method (Fig. 6a), CCV method (Fig. 6b), and the AAH method (Fig. 6c). The top 40 retrieved images are arranged from left to right and top to

bottom in order of decreasing similarity.

As shown in Fig.6, we may conclude that 1) AAH outperforms both the CCV and the traditional histogram; There are 100 relevant rhododendra images in image database. We have retrieved 17 of them in the top 40 images using the AAH, 16 using the CCV, and 12 using the traditional histogram method (we have illustrated them with a rectangle in Fig.6). Note the ranks of the retrieved relevant images using AAH are lower than those using the CCV method. 2) It verifies the idea of AAH we explained in section 2. The query image has the same color distribution as shown in Fig.1B. Most of the rhododendra images retrieved by the AAH have smaller and scattered petals as the query image. While most of the images retrieved by the traditional histogram incorrectly have a mass of large petals, which have similar distributions as shown in Fig.1A. Evidently, AAH outperforms the traditional histogram for distinguishing Fig.1A with Fig.1B. 3) The weights generated from the AHP method are effective for retrieving images. Furthermore, compared with the traditional linear method which requires users to supply weights of various features directly, our weight assignment is easier to perform and it provides a clear explanation that why it was that. The superiority gets more evident in the occasion where more than three features are involved in retrieval. On the other hand, we can also compute feature weights for a category of object offline according to the general accepted human perceptions of it and adopt them directly for image retrieval. Fig.7 illustrates the average precision-recall of our proposed scheme.

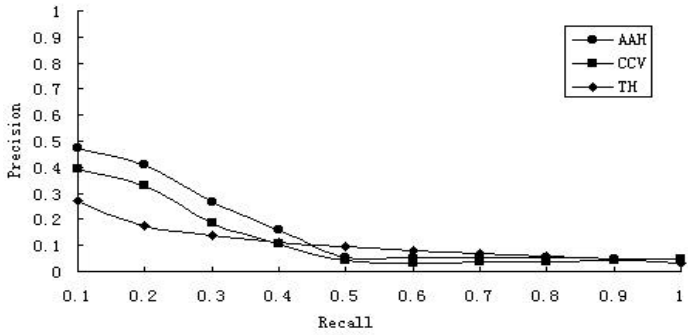



**Fig. 7 Precision-recall comparison of AAH, CCV and the traditional histogram methods**

## 9. Conclusion

In this paper, we have presented and discussed a retrieval system integrating the AHP method based on texture, shape and color features. An improvement on the traditional histogram called the Average Area Histogram is presented based on the sizes of the regions in each color bucket. Experimental results have proven the effectiveness of applying the AHP method for weights assignments in image retrieval. In the future, we will study new image content features such as the spatial information in retrieval and novel application of AHP in image retrieval without having to indicate the 6 choices, which makes it easier for users.

## 10. References

[1] Bao, P.; Xianjun Zhang, "Image Retrieval Based on Multi-scale Edge Model," in *proceedings of International Conference on Multimedia and Expo*, Lausanne, Switzerland, 2002, pp.417-420.

[2] Joo-Hwee Lim, Jian Kang, Wu., Singh, S. and Narasimhalu, D., "Learning Similarity Matching in

Multimedia Content-based Retrieval," in *IEEE Transactions on Knowledge and Data Engineering*, vol.13, no.5, 2001, pp.846-850.

[3] Sung-Bae Cho and Joo-Young Lee, "Human-oriented Image Retrieval System Using Interactive Genetic Algorithm," in *IEEE Transactions on Systems, Man and Cybernetics*, vol.32, no.3, 2002, pp.452-458.

[4] Filho, A., Mota,G.L.A., Vellasco, M.M.B.R. and Pacheco, M.A.C., "Query by image similarity using a fuzzy logic approach," in *proceedings of Fourth International Conference on Computational Intelligence and Multimedia Applications*, Yokusike, Japan, 2001, pp.389-394.

[5] J. J. Rocchio, *Relevance Feedback in Information Retrieval*. In *The SMART Retrieval System: Experiments in Automatic Document Processing*, G. Salton, Ed. Prentice-Hall, Englewood Cliffs, NJ, 1971, pp.313-323.

[6] Apostol (Paul) Natsev, John R. Smith, "Active Selection for Muti-Example Querying by Content," ICME-2003, Baltimore, USA, 2003, pp. 445-448.

[7] Saaty, T.L. *The Analytic Hierarchy Process*, McGraw-Hill, 1980.

[8] Mostafa, T., Abbas, H.M., Wahdan, A.A, "On the Use of Hierarchical Color Moments for Image Indexing and Retrieval," in *proceedings of IEEE International Conference on Systems, Man and Cybernetics*. Hammamet, Tunisia, vol.7, no. (6-9), 2002, pp. 6.

[9] Swain, M.J., Ballard, D.H., "Color Indexing," in *International Journal of Computer Vision*, vol.7, no.1, 1991, pp.11-32.

[10] Mlsna, P.A, Rodriguez, J.J, Efficient indexing of multi-color sets for content-based image retrieval. *In proceedings of the 4th IEEE Southwest Symposium of Image Analysis and Interpretation*. Texas, USA, 2000, pp.116-120.

[11] Colombo, C., Del Bimbo, A. and Genovesi, I., "Interactive image retrieval by color distributions," in *proceedings of IEEE International Conference on Multimedia Computing and Systems*, Texas, USA, 1998, pp.255-258.

[12] Androutsos, D., Plataniotiss, K.N.Venetsanopoulos, N., "Distance measures for color image retrieval," in *proceedings of ICIP' 98*, Chicago, USA, 1998, pp. 770-774.

[13] Pass, G. and Zabih, R, "Histogram refinement for content-based image retrieval," in *proceedings of WACV '96*, Sarasoto, FL, 1996, pp.96-102.

[14] Cinque, L., Levialdi, S., Olsen, K.A., Pellicano, A., "Color-based image retrieval using spatial-chromatic histograms," in *proceedings of IEEE International Conference on Multimedia Computing and Systems*, Florence, Italy, 1999, pp. 969-973.

[15] Rick Rickman and John Stonham, "Content-based image retrieval using color tuple histograms," *in proceedings of the SPIE-The International Society for Optical Engineering*, 1996, Vol.2670, pp. 2-7.

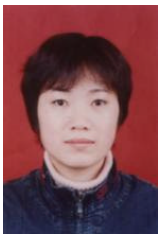
[16] Hsu, wynne, Chua, T.S., Pung, H.K., "Integrated color-spatial approach to content-based image retrieval," in *proceedings of the ACM International Multimedia Conference & Exhibition*, 1995, pp. 305-313.

[17] Safar, M. Shahabi, C. and Sun, X., "Image retrieval by shape: a comparative study," in *proceedings of International Conference on Multimedia and Expo*, New York, USA, 2000, pp.141-144.

[18] K. Hu. "Visual pattern recognition by moment invariants," in I*RE Transactions on Information Theory*, vol. IT-8, 1962, pp. 179-187.

[19] B.F. Jones1, G. Schaefer and S.Y. Zhu, "content-based image retrieval for medical infrared images," in *proceedings of the 26th Annual International Conference of the IEEE EMBS*, San Francisco, CA, USA, 2004.

[20] Flickner et al. "Query by Image and Video Content: The QBIC System", in *Computer*, vol.28, no. 9, 1995, pp. 23-32.
[21] B. S. Manjunath and W. Y. Ma. "Texture features for browsing and retrieval of large image data," in *IEEE Transactions on Pattern Analysis and Machine Intelligence, (Special Issue on Digital Libraries)*, vol. 18 (8), 1996, pp. 837-842.
[22] Kaplan, L M et al. "Fast texture database retrieval using extended fractal features," in *proceedings of Storage and Retrieval for Image and Video Databases VI*, SPIE 3312, 1998, pp.162-173.
[23] Robert M. Haralick, K. Shanmugam, and Its' hak Dinstein, "Texture features for image classification," in *IEEE Trans. On Sys, Man, and Cyb*, SMC-3 (6), pp.610-621, 1973
[24] Baraldi A, Parmiggian F., "An Investigation of the Texture Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters," in *IEEE Transaction on Geoscience and Remote Sensing*, vol.33, no.2, 1995,pp.293-303.
[25] Evangelos Triantaphyllou and Stuart H. Mann," Using the Analytic Hierachy Process for Decision Making in Engineering Applications: Some Challengies," in *International Journal of Industrial Engineering: Applications and Practice*, vol. 2, no. 1, 1995, pp. 35-44.
[26] Miller, D.W., and Starr, M.K, *Executive Decisions and Operations Research*. Prentice-Hall, Inc., Englewood Cliffs, NJ, U.S.A, 1969.
[27] Michael Ortega, Yong Rui, Kaushik Chakrabarti, Sharad Mehrotra, Thomas S. Huang, "Supporting Similarity Queries in MARS", in *proceedings of ACM Multimedia 97-Electronic Proceedings*, Crowne Plaza Hotel, Seattle, USA, 1997.
[28] Ardizzone, M. Cascia, "Automatic video database indexing and retrieval," in *Multimedia Tools and Applications*, 1997, vol.4, pp.29-56.



**Wang xiaoling** received the B.S. degree from Taiyuan Technology University, Taiyuan, Shanxi, China, in 1997. She is currently working toward the Ph.D. degree in computer science and engineering in Shanghai Jiatong University, Shanghai, China. Her research interests include content-based image retrieval, multimedia mining and pattern recognition

**Xie kanglin** graduated from Shanghai Jiaotong University, Shanghai, China, in 1964. He is now a professor of computer science and engineering in Shanghai Jiatong University, Shanghai, China. His main research interests include fuzzy logic and embedded system, data mining and decision supporting system. He has published about 100 papers and 6 books in these fields.