International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM)
ISSN: 2150-7988 Vol.1 (2009), pp.279-286
http://www.mirlabs.org/ijcisim

A Study on the Algorithm Based on Image Color Correlation Mining

Chen YongYue[1], Zhang HuiPing[2], Xia HuoSong[1]

[1]Department of Information Management and Information System
Wuhan University of Science & Engineering, Wuhan, China
chenyongyue_17@163.com

[2]School of Politic and Public Administration
University of Electronic Science and Technology of China, ChengDu, China

Abstract: Because of the semantic gap between low-level feature and high-level semantic feature of images, the results of the traditional color-based image retrieval can't meet users' needs. In order to eliminate interference factors in the image retrieval, use image semantic feature and improve the accuracy of image retrieval, the paper introduces an algorithm based on the color correlation mining. It regards the pixel rows as a transaction set, uses the Apriori algorithm to find out the rows by looking for the continual co-occurrence color in the transaction set. These rows are correlative with the semantic object of the image. Then it extracts the correlative color histogram of image form the correlative color set to realize the correlative color mining.

Keywords: image retrieve, Correcolor Mining, Apriori algorithm, Color spatial Quantization

1. Introduce

Since M.J.Swain and D.H.Ballard proposed using the color histogram to calculate the similarity of the image colors in 1991 [1], the color histogram has become the simplest and common means to express the color characteristics of the image. This traditional global color histogram is a simple and efficient method to express the color characteristics of an image, but its natural statistical feature determines it can only reflect the image visual information [2]. The best way to alleviate this problem is to add the information of color space distribution to the color histogram. Moreover, its another weakness is that the color histogram can't reflect the semantic feature of an image. In order to solve the above problems, the paper introduces a method about the color correlation mining. It regards an image as a transaction set, each row of the image as a transaction and pixels in rows as the items of the transaction. Then it discusses the algorithm of the correlative colors mining by the classic Frequent Itemsets Mining method based on Apriori[3][4][5].

It is one of the main tends in the development of the image retrieval to apply data mining in the image retrieval [6]. On the one hand, the traditional color histogram does not reflect the spatial distribution information of colors [7]; on the other hand, most of previous studies lie in the object identification in images, moreover, these methods are greatly complicated. Therefore, this paper discusses a kind of method about the correlative colors mining and mainly researches the following two questions.

(1) Color spatial quantization. Color spatial quantization is the basis of color feature expression. A good color quantization approach can finally determine whether the performance image retrieval algorithm is good or bad. In the paper, we discuss the color quantization methods about RGB and HSV.

(2) Feature extraction algorithm. Feature extraction algorithm is the core of the image retrieval based on the color characteristics. By the algorithm, the correlation between the characteristics of the colors in an image can be mined. It also researches the expression approach of color characteristics.

2. Color spatial Quantization

2.1 Color space model

The quantization of color space is the basis of color-based image retrieve and the choice of color space is the basis of the color space quantization. Different color space models require different quantitative methods. The quantization of different color spaces has a great effect on the performance of image retrieve. In accordance with the basic structure of color space, it can be divided into two broad categories: the color space of the base color and the color space whose color and bright are separated [8]. The former is a typical RGB color space. It also includes CMY, CMYK, CIE XYZ and so on. The latter include the HSV color space as well as YCC / YUV, CIE Lab and so on. In two different types of color space, due to the device-dependent of RGB color space and the visual consistency of HSV color space, they has been widely used in image retrieve.

2.1.1 RGB color space

RGB color space model is put forward by the U.S. National Television System Committee (NTSC) in order to display color images in CRT [9]. RGB space can be expressed as a cube in three-dimensional color space coordinate system. Its side length is 256 (Figure1). R, G and B are respectively the axis of the color space. In the cube, any point expresses a color. In each axis, the coordinates (r, g, b) of the point are quantified into an integer between 0 and 255. The eight vertices of the cube respectively correspond to the black, red, yellow, green, blue, purple, white and cyan.

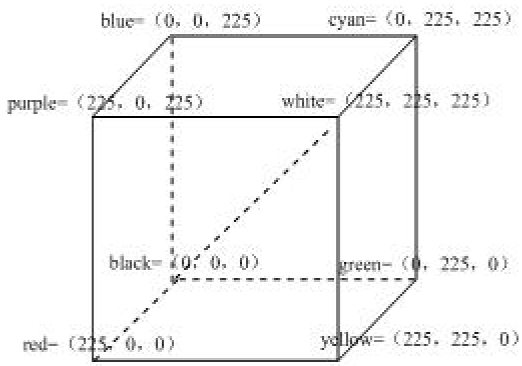


Figure1. RGB color space

At present, RGB space has become one of the most important color space and most of the color digital images stored in computer are expressed by RGB space. R, G and B are also called as the three channels of the color. 24bit color image which we commonly refer to is that a true-color pixel respectively uses 8bit (1 byte) to express three color components. Therefore, it requires 24bit (3 bytes) to represent the three color of RGB in all and it can express $2^{(8+8+8)}$ = 16,777,216 colors.

RGB space is characterized by the visual inconsistencies and the strong correlation among components. However, it is directly defined according to the image-forming principle of the camera and expresses the raw data by which the digital images are displayed on a computer monitor without the conversion, moreover it is easy to understand, so the quantization of RGB color space are still worth researching.

2.1.2 HSV color space

HSV space directly corresponds to the three elements of color vision characteristics, namely, hue H, saturation S and Value V which are independent each other. H expresses the color tone of light, such as red, green, blue, orange and so on. S refers to the depth or shades of color. The depth of saturation is related to the ratio about color and white. If the ratio of white is 0, the saturation is 100%. If there is only white, the saturation is 0. V is the light and shade.

HSV space can be expressed by an inverted cone (Figure 2), the long axis of the cone expresses Value (V) which is measured from black (0%) to white (100%) by percentage so as to reflect the light and shade. The distance with the long axis shows Saturation (S) which expresses the uniform purity of the color. It is also expressed by the percentage. The angle around the long axis expresses Hue (H) which is measured from 0 to 360 degrees.

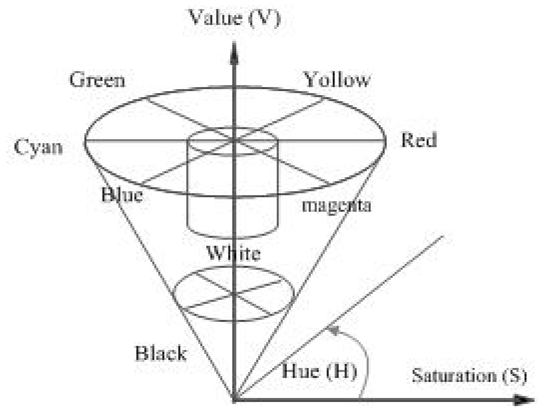


Figure 2. HSV color space

HSV space is characterized by visual consistency. The visual consistency indicates that the Coordinate distance in the color space is similar to the distance in the observer's Color perception space.

It contains two meanings:

- The distribution of the color pixels in the color space is consistent, uniform;
- The transition between the different colors and the color space is smooth from the perspective of human visual perception.

Studies have shown [10] that the visual consistence of HSV color space is better than the traditional RGB space and other visually consistent color spaces. HSV color model in the CBIR application is more suited to the judgment of the user's eye and can get closer to the retrieval results the human can observe.

2.2Color spatial Quantization

Color space quantization is closely linked to color feature representation. The result of color space quantitative is namely the color feature representation. At present, almost all of the color retrieval algorithms regard the color histogram as the representation of the color characteristics. Color histogram describes the statistical distribution of the image color. It is characterized by the translation, scaling, rotation invariance, so it has been widely used in the content-based image retrieval system at home and abroad.

There are many ways using color histogram to retrieve. They mainly include general histogram method [11], cumulative histogram method [12] and so on. Since the overall distribution of the chrominance signal about HSV color space does not fully meet the applied premise of cumulative Histogram from the visual sense, we still use the general statistical histogram method. Moreover, RGB space doesn't also meet the conditions of the cumulative histogram. Therefore, under RGB space, we are still using the general statistical histogram method.

Generally speaking, a lot of colors can be identified by a computer, but these colors which can be named by people are very few. In nature, the name and classification of colors usually include red, yellow, blue, green, and cyan, purple, black and white. We regard the eight hues as the basic categories and all colors belong to these categories. The color histogram used for image matching is the statistical information of occurrence frequency about eight colors.

2.2.1 RGB spatial quantization

In the cube of RGB space, the eight colors express the 8 vertexes of the cube; the similarity of two colors can be expressed by the distance between them in the RGB color space. This distance is often called the color distance. In RGB color space, the color distance ΔC between (r_1, b_1, g_1) and (r_2, b_2, g_2) can be defined as:

$$\Delta C= (r_1 - r_2)^2+ (g_1 - g_2)^2+ (b_1 - b_2)^2$$

Since the sensitive degree of the human eye to red, green, and blue is different. Therefore, every color component should be added a certain weight. The calculation formula of the experience color distance is:

$$\Delta C=3(r_1 - r_2)^2+4(g_1 - g_2)^2+2(b_1 - b_2)^2$$

Extract the RGB component value of any one pixel of an image and use the above formula to calculate the distance between the pixels and the eight vertexes. Then, regard the color of the nearest vertex as the color of the pixel.

2.2.2 HSV spatial quantization

In HSV space, we first used two components (S and V), to classify gray and similar gray into black or white, then and classify the remaining colors into the remaining 6 kinds of colors.

$$\text{color}=\begin{cases}\text{white } ((S<0.1)\&\&(V\geqq 0.8))\|\ ((S<0.1) \\ \qquad\qquad \&\&(0.5<V<0.8)) \\ \text{black } (V<0.1)\|\ ((S<0.1)\&\&(0.1\leqq V\leqq 0.5)) \\ \text{other} \qquad \text{other}\end{cases}$$

Since a color is codetermined by H, S and V, when S is given, the boundary between various similar colors is not only related to the value of H, but also the value of V. Therefore, it is not convenient for the division of similar colors. In order to make the division of similar colors approximate invariability to V, Liu Weizhong etc brought forward the interval transformation method of similar color in 1998[14]. The paper uses the method to transform H. Then according the value of H and its interval, we classify all colors into red, yellow, blue, green, and cyan, purple. The specific method is as follows:

If(color==other)
{
H_{old}->H_{new};

$$\text{color}=\begin{cases}\text{red} & (H_{new}\leqq 30^o)\|\ (H_{new}\leqq 330^o) \\ \text{yellow} & 30^o< H_{new}<90^o \\ \text{green} & 90^o\leqq H_{new}\leq 150^o \\ \text{cyan} & 150^o< H_{new}<210^o \\ \text{blue} & 210^o\leqq H_{new}\leqq 270^o \\ \text{purple} & 270^o< H_{new}<330^o\end{cases}$$

}

3. An algorithm based on the correlation mining of colors

3.1 The definition of some concepts

Generally, not all colors in an image are related to the query object described by the image. Some colors are irrelevant to the query object. Therefore, the statistics on the colors of these pixels will affect the retrieval effect [13]. In order to reduce this kind of error, during the retrieval, the color of the object described by the image must be mined so as to search.

The retrieval object which users are interested in is very intensive in images. Moreover, the pixels of the object and the background are often frequent co-occurrence in some rows. According to this characteristic, in order to find the rows related to the object in an image, we should find out the rows in which the background color and the object color are often frequent co-occurrence (Figure 3).

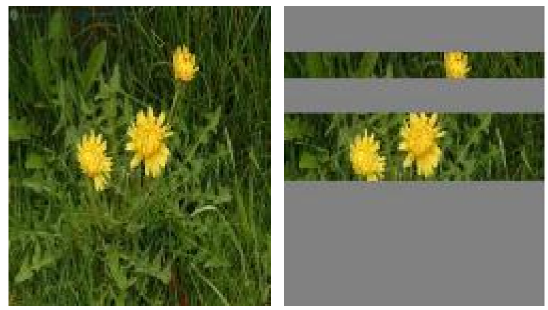


Figure3. The image and the retrieval object

This search form is similar to mining association rules. Therefore, we can regard the image as a pixel matrix. That is to say, a row of pixels can be regarded as a transaction and each pixel can be regarded as an item of the transaction. Since the pixels are the representative of the quantized colors, a quantized color can be regarded as an item. In order to discover the correlation of the color, several concepts are defined as follows:

Correlation: suppose that C= (c_1, c_2 ,..., c_m) is a quantized color set; image I is a pixel row set; each pixel row R is a quantized color set, and $R\subseteq I$. Suppose that A is a quantized color set; the pixel row R contains A and if and only $A\subseteq R$. In the image I, the percentage of the pixel rows containing A in all pixel rows is called the correlation of the quantized colors in A.

Correcolor: the set which contains k quantized color sets is called k-color set. If a minimum correlation (min_corr) is given, the quantized color sets, whose correlation are equal or bigger than min_corr are called correcolor set. The correcolor set which contains k quantized colors is called k-correcolor set, recorded as L_k. In correcolor set, all quantized colors are called the correcolor of this image. In k-correcolor set, quantized colors are called k-correcolor.

3.2 Correcolor Mining

Because correcolor set is usedto express the relationships between colors and correcolor set is similar to frequent itemsets, using the method of the modified frequent itemsets mining can discover the relationships between the colors of an image. Apriori algorithm is one of the most classical frequent itemsets mining algorithms. Its efficiency is not high, but Apriori algorithm has little or no effect on the retrieval efficiency under the situation of only eight kinds of quantized colors. Moreover, Apriori algorithm is simple, effective and easy to realize.

3.2.1 Image pretreatment

For the realization of the algorithm, the concept of quantized colors should correspond to frequent itemsets, which can be obtained by using the frequent itemsets mining algorithm. However, the color of the pixel is not the original color in actual operation. It must be the quantified color, so when the image is scaned each time, it must be re-quantified. In addition, after all colors of image are quantified, the cost of scanning the image one time is still large. However, the frequent itemsets mining algorithm usually needs to scan the image many times. As a result, before the correcolor mining, it is necessary to use the color quantization method to process the image, eliminate repetition of the quantized colors in the whole image and store these colors to a simplified image logic model in rows, which is called Row Color Set. During the process of the correcolor mining, it only needs to scan Row Color Set. The row color set is much smaller than the original image. So the pretreatment process improves the efficiency of correcolor extraction and reduces the time of image retrieval. The pretreatment algorithm is as follows:

Algorithm: pretreatment (input: image, output: RowColors)
Algorithm: pretreatment (input: image output: RowColors)
Input: image // image
Output: RowColors // Row Color Set
1) for all rows do begin
2) for all columns do begin
3) get_rgb (pixel); // get the rgb value of pixels
4) pixel (r, g, b) → color; // quantify rgb value for color value
5) nrow [color] + +;
6) end; //Apriori algorithm treats transaction set to find out frequent itemsets.
7) for (color = 0; color <c; color + +) do begin
8) if nrow [color] / width> = value do begin // the percentage of color in the current row is higher than the threshold
9) RowColors [row] [color]. Exsit = 1;
10) RowColors [row] [color]. Count = nrow [color]; // store the appearance frequency of the color in the current row.
11) nrow [color] = 0;
12) end;
13) end;
14) return RowColors;

The input image in algorithm is the original image and is regarded as the pixel set in the algorithm, which is expressed by the RGB color value. Its structure is the same to sPixel. The output RowColors in algorithm is the Row Color Set and can be regarded as the frequency record of quantized colors in each row. Its structure is the same to sQuantiColor, which is a two-dimensional array. Its size is r × c (r is the number of image; c is the number of the color categories). The value of the array is initialized to 0.

```
struct sPixel
{
int x;// The x-coordinate of the pixels in the image
int y;// The vertical coordinates of the pixels in the image
sRGB rgb;// The value of RGB about the pixels
}

struct sQuantiColor
{
int exist;// It expresses whether the quantitative color exists or not in the current row; 1 expresses existence and 0 expresses inexistence.
int count;// The frequency of the color in rows
}
```

After the pretreatment algorithm to scan an image on time, it can obtain the simple logic expression of the image: Row Color Set. By using Row Color Set as the expression of an image, it can make the correcolor mining simpler and easier so as to reduce the time of the image scan and improve the efficiency of the image retrieval algorithm.

3.2.2 The correcolor mining Based on Apriori algorithm

From the above analysis, we can know the concept of correcolor set is similar to frequent itemsets. Therefore, we can use the modified frequent itemsets to realize the mining of an image pixel row set so as to get the relationships between colors expressed by correcolor set.

There are many frequent itemsets mining algorithm. However, Apriori algorithm is one of the most classic algorithms. It utilizes the occurrence frequency statistics of the transaction items in a transaction set and an iterative tier-by-tier search method to realize the mining. Therefore, when inputting parameters during the process of mining, we can change the transaction set into the row color set so as to get a correcolor mining algorithm which is called I-Apriori. The efficiency of I-Apriori algorithm is not high, but it has no great effect on the image retrieval time under the eight kinds of quantized colors. The algorithm is as follows:

Algorithm: I-Apriori (input: RowColors, output: L)
Input: RowColors // row color set
Output: L // correcolor set
1) L_1 = (large 1-itemsets);
2) for (k = 2; $L_{k-1} \neq \Phi$; k + +) do begin
3) C_k = Cand-gen (L_{k-1}); //find k-candidate correcolor set through new L_{k-1}
4) for all Color Row rc in RowColors do begin //scan row color set to get the count of candidate correcolor set
5) C_{rc} = subset (C_k, rc); // the candidate correcolor set in the color row

```
        for all candidates c ∈ C_rc do
            c.count + +;
        end
        L_k = (c ∈ C_k | c.count ≥ min_corr);
    end
    return L = ∪_k L_k;
```

I-Apriori first finds out 1-correcolor set L_1 in row color set and uses L1 to find out 2-candidate correcolor set C_2. By scanning the row color set, it can count the minimum correlation (min_corr) to these elements of C2. Then it removes these elements, which don't meet with the min_corr, to obtain 2-correcolor set L_2. Similarly, according to the above operation method, it uses L_2 to find L_3. Then use k-1-correcolor set L_{k-1} to get L_k until it can't find out a higher correcolor set. The final k-correcolor set forms the largest schema of image correcolor, which is called L_{Lk}. When each L_k is found, it must scan the row color set one time. Algorithm finally outputs all modes of the correcolor set.

In the algorithm, there is a minimum correlation (min_corr), whose value effects correcolor set. If its value is too high, it is too difficult to get correcolor set. If its value is too small, the correcolor set gotten by the algorithm is no significance. Generally, its value is set to 0.3. Its reasons are as follows: If about a third of image shows the correlative color, the correcolor set must be meaningful. In addition, since Golden Section is 0.37, choosing Golden Section point is in line with the aesthetic habits of people. However, considering the robustness of retrieval result, we set min_corr lower so as to improve the recall and precision of image retrieval. In order to test min_corr=0.3 rationality, we randomly choose ten piece of images and use the algorithm and the method of RGB spatial quantization to search them. To the same image, we choose min_corr from [0.25, 0.35], retrieval it ten time and add 0.01 to min_corr each time. According to the results of the experiment, the 80% results show that the effect is the best when min_corr is set to 0.3.

4. Experimental Analysis

To examine the performance of the algorithm, in the paper the experiment of the retrieval performance is done to compare the global-color histogram with the correcolor histogram. We call the image retrieval based on the correcolor histogram 8-correcolor-histogram-xxx and the traditional global color histogram retrieval method 8-color-histogram-xxx. Xxx is regarded as RGB or HSV. They respectively express the eight-core-color quantitative method under RGB space and the eight-color quantitative method under HSV space.

4.1 Evaluation criterion

There are many methods of the performance assessment about Image Retrieval, but they generally are based on the number of the similar images and the rank in the retrieval result. At present, comparing the Precision and the Recall about the result is the most currently used evaluation methods. Recall expresses the ratio of the number on the retrieved relevant images and the actual number on the relevant images to the objective image in the database during a query process. Precision is the ratio of the number on the retrieved relevant images and the number of all retrieved images in a query process. By the analysis of the retrieval results, four kinds of conditions can be obtained (table 1). They correspond to four basic parameters of A, B, C, D. They respectively represent the number of retrieved relevant images, the number of retrieved irrelevant images, the number of the missed relevant images and the number of the correct rejection to the images.

	Judgment	
	Correlation	Non-Correlation
Retrieved	A(Correct Result)	B(False Retrieve)
Not Retrieved	C(Missed Retrieve)	D(Correct Rejection)

Table1. The basic performance parameters of the image retrieve

With the basic parameters in Table 1, Recall and Precision can be defined as follows:

Recall = (the correlative correct retrieval results) / (all the correlative results) = A / (A+C)
Precision = (the correlative correct retrieval results) / (all the retrieval results) = A / (A+B)

Before the calculation on Recall and Precision of Image retrieval algorithm, we should know the correct query results to the retrieved images. In other words, when we calculate the recall rate and precision rate, the image library images, we must know the number of the correct retrieval result associated with the retrieved image.

4.2 Experimental results and analysis

In order to obtain the experimental data, we randomly select 150 images as the experimental image database from the natural scenery image database including 2000 pieces of images. In order to comprehensively examine the retrieval performance of the image retrieve algorithm, we design two experiments. They respectively examine the precision rate and recall rate of the image retrieve algorithm to the randomly selected images, and the trend of the precision rate based on the recall rate.

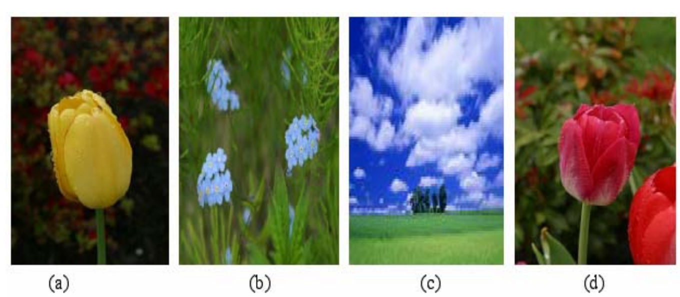


Figure 4. The image cases in the experiment

We randomly select four images (Figure 4) as the retrieval image cases. The four image cases were respectively used the traditional histogram-RGB algorithm, the histogram-HSV algorithm, and the cumulate-HSV algorithm, the 8-color-histogram-RGB algorithm and the eight 8-correcolor-histogram-RGB algorithm, as well as 8-color-histogram-HSV algorithm to retrieve. Based on the retrieval result of each algorithm, we compare these color-based image retrieval algorithms.

4.2.1 The experiment of Recall and Precision

By the analysis of the image database, the number of these similar images to the four image cases in the database is respectively 22, 10, 17 and 35. Therefore, we respectively choose the top 22, 10, 17 and 35 images to analyze the result. To each image, the recall rate and precision rate of each image retrieval algorithm are respectively shown in Table2 and Table 3.

	image（a）	image（b）	image（c）	image（d）	average
histogram-RGB	0.36	0.6	0.94	0.5	0.6
histogram-HSV	0.18	0.6	0.47	0.53	0.45
Cumulate-HSV	0.32	0.8	1	0.38	0.63
8-color-histogram-RGB	0.32	0.4	0.65	0.29	0.41
8-correcolor-histogram-RGB	0.77	0.8	1	0.79	0.84
8-color-histogram-HSV	0.64	0.6	1	0.62	0.71
8-correcolor-histogram-HSV	0.64	0.7	0.94	0.68	0.74

*Table 2.*The comparison of the recall rate about different algorithm

	image（a）	image（b）	image（c）	image（d）	average
histogram-RGB	0.4	0.6	0.8	0.49	0.57
histogram-HSV	0.2	0.6	0.4	0.51	0.43
Cumulate-HSV	0.35	0.8	0.85	0.37	0.59
8-color-histogram-RGB	0.35	0.4	0.55	0.29	0.40
8-correcolor-histogram-RGB	0.85	0.8	0.85	0.77	0.82
8-color-histogram-HSV	0.7	0.6	0.85	0.6	0.69
8-correcolor-histogram-HSV	0.7	0.7	0.8	0.66	0.71

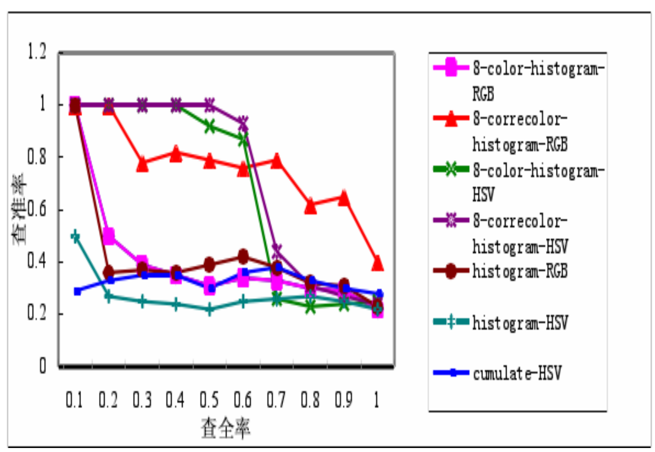
*Table 3.*The comparison of the precision rate about different algorithm

As can be seen from Table 2, for all image cases, the performances of the correcolor histogram algorithms under RGB space and HSV space are generally the best of them. Their average recall rate is higher than the other retrieval algorithm much better. By comparing the performance of the correcolor histogram algorithms under different spaces, we found that HSV is less than RGB. The reason may be that the whole space is be evenly spilt so that some similar colors can not be classified as the dissimilar class. It is easy to produce the correcolor patterns so as that the retrieval effects are better. If HSV can be closely split, the retrieval effects under HSV space are better than under RGB space.

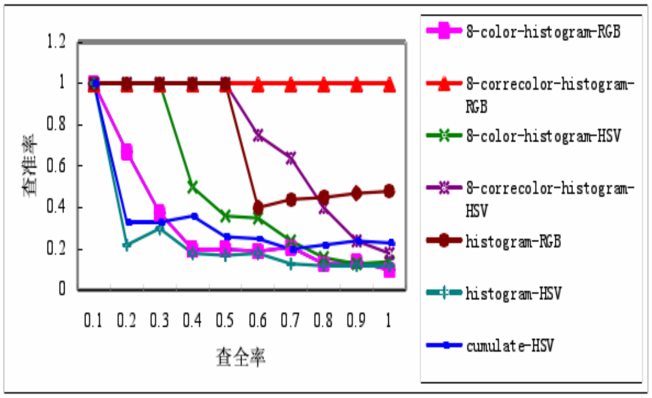
4.2.2 The tendency of Recall and Precision

The calculation of the above examine is based on a certain number of search results. However, if the results of the recall rate or the precision rate are equal on the above examine and we reduce the number of the analytical results, the different algorithms may be obtain the different Recall and Precision. So it is necessary for the different algorithms to calculate the precision rate under the different recall rate. The specific method is: Let a given recall rate is n ($n \leq x$). x is the total number of the similar images to the retrieval image in the database. According to the Similarity sorting, obtain the number y which includes the n*x similar images. So the precision rate is x/y. we choose four images as the cases. Then we use every algorithm to calculate the precision rate in [0.1, 1.0] of the recall rate and in each calculation the recall rate is an increase of 0.1. The retrieval results are showed in Figure 5.

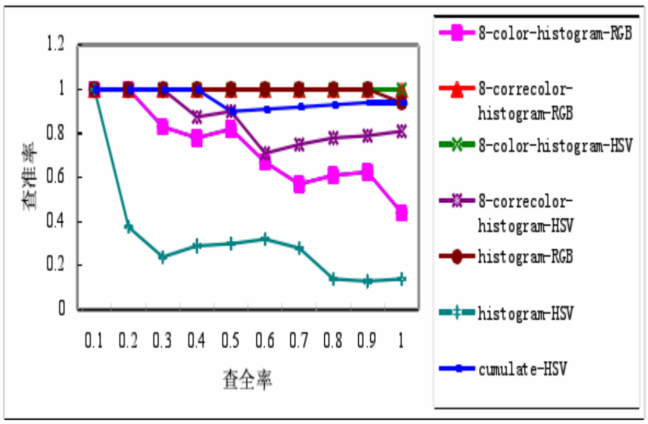
In general, a good image retrieval system requires both Recall and Precision high. Whichever is low, it can not be considered a good system. From the tendency of the precision rate and the recall rate (Figure5 (b)), we can see the precision rate of every algorithm can achieve the maximum value 1 when the recall rate is 0.1. However, once the requirement of the recall rate is heightened, the precision rate of every algorithm will be decreased in varying degrees. The degree of decline shows the performance of the algorithm. Of course, the too high recall rate and the too low precision rate can not be regarded as the evaluation standard of the performance.



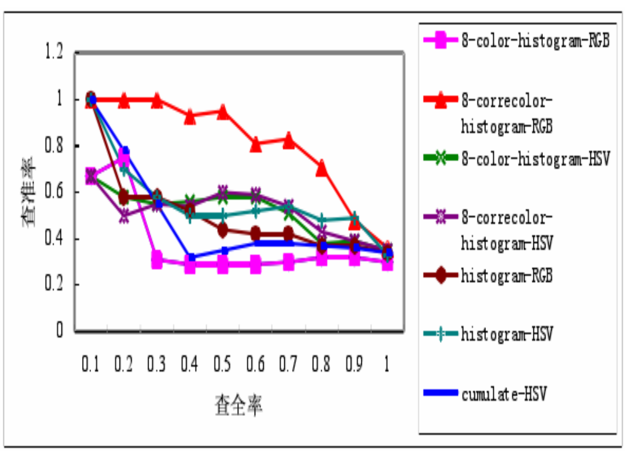
(a) The tendency of Recall and Precision about figure 4 (a)



(b) The tendency of Recall and Precision about figure 4 (b)



(c) The tendency of Recall and Precision about figure 4 (c)



(d) The tendency of Recall and Precision about figure 4 (d)

Figure 5. The tendency of Recall and Precision about the different image cases

As can be seen from Figure5, according to the Comparative experiment of the different algorithms, the recall rate and the precision of the image retrieval algorithm based on the correcolor-histogram is better than the traditional global color histogram. The effect is especially distinct to the images that the distinction between the object described in the image and the background of the image is rather clear. In figure5 (c), the advantage of the image retrieval algorithm based on the correcolor-histogram is not particularly obvious. It is because the target described in the image occupies a larger range of the image. Therefore, the number of the statistic rows in the traditional global algorithm and the algorithm is not distinct so that the result of them is similar. By the above experiments, it shows that our algorithm is effective.

5. CONCLUSION

It is difficult for the low-level feature of images to embody the semantic feature of images so that it is very difficult for the traditional color-based image to meet users' needs [15]. In order to utilize image semantic feature and improve the accuracy of image retrieval, the paper introduces an algorithm based on the correlation mining of colors. It regards the pixel rows as a transaction set, uses the Apriori algorithm to find the colors which appear frequently in the same transaction, gets the rows which are correlative with the image's semantic object, then extracts the correlative color histogram of image from the correlative color sets to achieve the goal of the correlative color mining.

Acknowledgment

This research was supported by China Social Science Fund (07BTQ010), HuBei Province Statistical research project (HB092-33), HuBei Province Department of Education Social Science(2009q069) and HuBei Province Social Science Fund of China (2007097).

Reference

[1] M.J.Swain, D.H.Ballard. Color Indexing. International Journal of Computer Vision, 1991(7): 11~32

[2] R.Schettini, G.Ciocca, S.Zuffi. A Survey on methods for color image indexing and retrieval in image databases in: R.Luo,L.MacDonald eds.. Color Imaging Science: Exploiting Digital Media. John Wiley,2001.183~211

[3] Gu QiHong. The research and application about Association Mining Apriori Algorithm http://www.ahcit.com/lanmuyd.asp? id=3231

[4] MEI Cheng, ZHOU Xing-bin. An Improvement of Apriori Algorithm Based on Matrix. Computer and Modernization, 2008 (12): 5-7

[5] Agrawal, R,Imielinski, T,Swani A. Mining Association Rules between Sets of Items in Large Databases //Proc. ACMSIGMOD Conf. Management of data, May 1993: 207-216.

[6] Wu Nan, Song Fang-min. An Image Retrieval Method Based on High-level Image Semantic Information. Journal of Image and Graphics, 2006(12): 1774-1780

[7] Liu Fang, Liu Tao. Image Retrieval Based on Color-Space 2D Histogram. Computer Engineering and Applications, 2002(12): 36-39

[8] Lin FuZhong. The Basic of Multimedia Technology. TsingHua University Press, 2009(9): 182-196

[9] A. Hamrapur, A. Gupta, B. Horowitz, et al. Virage Video Engine. in: Ishwar K. Sethi,Ramesh Jain eds.. Proceedings of SPIE Conference of Storage and Retrieval for Image and Video Databases V. San Jose, CA. Feb. 1997. 188~197

[10] B. SManjunath, W. Y. Ma. Texture Features for Browsing and Retrieval Image Data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(8): 837~842
[11] Li YongGang. A Method for Image Color Quantization Based on Histogram. Computer Systems Applications, 2008(7): 35-39
[12] Bob Johnson. Your Friend, the Histogram, Earth- bound Light (earthboundlight.com), 2004(11).
[13] Liu Weizhong, Zhang Yujin. Color Image Retrieval Based on Color and Shape Features. JOURNAL OF IMAGE AND GRAPHICS , 1998(7): 332-336
[14] AN Ying. Mining Interest Set-Weighted Association Rules Based on Algorithm Apriori. Journal of Beijing Union University(Natural Sciences), 2008(04):44-47
[15] A.W.M.Smeulders, M.Worring, S.Santini etal. Content-Based Image Retrieval at the End of the Early Years. IEEE Trans.Pattern Analysis and Machine Intelligence, 2000, 22(12): 1349~1380

Author Biographies

Chen YongYue Born in YueYang, HuNan, China in October, 1973. I received a bachelor degree in 1998. From 2002 to 2008, I studied information management and information system and received a master degree and a doctor degree in Information Science. In July 2008, I became a lecturer in Wuhan University of Science and Engineering and mainly studied information management, knowledge management, Electronic Commerce and so on.

Zhang HuiPing Born in HeNan, China in 1982. I received a bachelor degree from University of Electronic Science and Technology of China in 1998. From 2003 to 2008, I studied Management Science and Engineering and received a master degree and a doctor degree in Information Science. Moreover, since 2003, I have studied in information management and knowledge management in University of Electronic Science and Technology of China.

Xia HuoSong Born in HuangGang, HuBei China in 1964. I received a doctor degree of Management Science and Engineering from HuaZhong University od Science and Technology in 2003. Since then, I have studied information management and information system, Knowledge Management and Supply Chain Management. At present, as a professor, I act as the dean of School of Economics and Management in WuHan University of Science and Engineering.