A novel image retrieval method using color coherence quantity

Pei-Cheng Cheng¹, Been-Chian Chien², Hao-Ren Ke¹, and Wei-Pang Yang¹
cpc@cis.nctu.edu.tw, cbc@isu.edu.tw, claven@lib.nctu.edu.tw, wpyang@cis.nctu.edu.tw
¹Department of Computer & Information Science, ²Department of Information Engineering
National Chiao-Tung University, I Shou University
Hsinchu, Taiwan 30050, R.O.C. Kaohsiung, Taiwan 84041, R.O.C.

Abstract
Color histogram has been used for color image retrieval and is a well-known method. Although color histogram represents the color ratio of an image, it lacks the information about spatial relationship between pixels. In this paper we propose a new statistic method called CCQ (color coherence quantity) to improve color histogram. Based on this method we improve the color histogram measurement for image retrieval systems. The color coherence quantity is not influenced by image rotation and zooming. The experimental result shows that our method outperforms previous methods in retrieving similar images.

Keywords: Color image retrieval, Color histogram, spatial relationship.

1. Introduction
The research topics on multimedia information retrieval have been ardently discussed because massive amount of digital content is produced by digital cameras, scanners, and is available in the world wide web and digital museums. Image retrieval is one of the most significant research topics on multimedia retrieval. Most Images do not annotated by keywords; therefore, they cannot directly apply the traditional text-based retrieval methodology. However, images contain copious information like color, shape, and texture. Image data consists of two-dimensional pixels and is more complex than text data. Unlike textual data, image data cannot be retrieved by string matching; furthermore, the size of image data is extremely larger than text data. Many researchers are studying to provide tools and methods to efficiently manage pictorial digital libraries containing image content, and color is the most conspicuous feature for content-based image retrieval.

The same object in different images may have different size or face because images are shot in different places or lens’s zooming. However, object segmentation and identification is still an open issue. Computers still cannot automatically recognize which objects are contained in an image. So, image retrieval is unlike traditional text retrieval, which can extract exact keywords for finding documents. Image retrieval is more complex than text retrieval and need to be paid more attention.

In the past few years, many researchers strived for the development of novel representation of pictorial content. They were devoted to designing an image-retrieval method that can find similar image efficiency and effectively. The image-retrieval process can be divided in two steps: the first uses an efficient representation to denote extracted features and to enable indexing, and the second designs a metric to measure the similarity of two images. The features contain in an image include color, shape, texture, and spatial relationships. In general image-retrieval systems, color is the most outstanding feature and color histogram methods [1] have good performance to retrieval similar images.
Color histogram measures the proportion of each color in an image. It is fast to compute and is not influenced by rotation and small variation in camera viewpoints. Methods using color histogram can find similar images that contain the same color ratio; however, color histogram is not able to describe the spatial distribution of each color. As a result, an image containing a mass of yellow color may have the same quantity as an image containing many small spots of yellow color.

Many variants and extended histogram methods have been proposed to conquer the drawback of losing spatial relationship. In this paper we propose a new representation to describe the distribution of color. The proposed representation is very concise and the experiment shows that our method outperforms previous works.

In this paper, we review the related work in section 2. Section 3 illustrates and defines the new representation for comparison and proposes a similar metric corresponding to the new representation. In section 4, the experimental result shows that our proposed method has better performance than previous methods. The conclusion and further work is discussed in section 5.

2. Related work

In general, the process of image-retrieval systems can be separated into two steps: the first finds a good representation of original images for fast and easy comparison. The second designs a suitable metric to calculate the similarity between two images. The visual features of an image are color, shape, texture and spatial relationship. Color is the most conspicuous feature from the point of human’s visual perception. Color has been accredited as the most useful feature for general-purpose image searching, and color histogram has been widely used for image retrieval [4]. Each pixel of an image consists of many type of color in computer system. Efficient and effective computation of color indices usually involves reducing the color space by color space quantization.

Mitra [7] used the most significant bits of each of the R, G, B color channels to reduce the number of image colors. But the RGB color space does not design for human perception. Smith and Chang [8] partitioned the HSV color space into 166 bins; the HSV color space is more consistent with human perception than the RGB color space. There are many extended methods for color space quantization [11][9], by predefined color palette (static quantization), or by clustering and/or spatial segmentation (dynamic quantization). The common objective is to find a minimum subset that can present the distance between each color with respect to human perception. As the number of colors is significantly reduced, the computation time can be improved.

Another problem is how to represent the content of an image. Color histogram [13] is a basic method and has good performance for representing image content. Color histogram methods gather the statistics about the proportion of each color as the signature of an image. The color histogram \( H_i \) of image \( I \) is a vector \((h_1, h_2, ..., h_j, ..., h_n)\), in which each bucket \( h_j \) counts the ratio of pixels of color \( c_j \) appeared in the image. The similarity of two images \( I \) and \( I' \) can be measured as the distance of their corresponding color histograms. The distance of two color histograms are computed by using standard methods (such as the \( L_1 \) distance or \( L_2 \) distance) for comparing histogram buckets. Then, a query image \( I \) can find the most similar images by evaluating the smallest distance from the image database.

The color histogram may misjudge similar image cause by color shift. There are other researchers discussing the problem of how to reduce the influent of color shift. Stricker and Orengo [11] proposed cumulative histograms and Hafner [4] used a weighted distance between histograms that count the “cross-talk” between colors.

The indexing method described above, does not consider the spatial information of image content. Thus, they cannot distinct the difference between a mass of yellow or
many spot of yellow which have equal proportion in statistic. The simplest way to provide spatial information is to divide the image into sub-image, and then index each of these [2][8][12], but it causes the another problem that position shift will be misjudge non-similar.

Pass, zabih and miller [15] present a histogram-based for comparing images called CCV (color coherence vector) that incorporates spatial information without influenced by position shit. They classify each pixel into two groups, coherent or incoherent, depending upon if it is belonged a large uniform color region. And, calculate the color histogram of them individually. For each color \( c_i \) was divided into two parts the number of coherent pixels, denote as \( \alpha_i \), and the number of non-coherent pixel, denote as \( \beta_i \). The sum \( \alpha_i + \beta_i \) is equal the number of pixels of color \( c_i \) present in the image; the sum of each set for \( i=1..n \) will correspond the color histogram. The coherence vector of an image \( I \) is defined as:

\[
CCV(I)=\{(\alpha_1, \beta_1), \ldots, (\alpha_i, \beta_i), \ldots, (\alpha_n, \beta_n)\}
\]

Two image \( I \) and \( I' \) been presented by two CCV can use the L1 distance to compare the distance.

\[
\text{Distance}(I, I') = \sum_{i=1}^{n} (|\alpha_i - \tilde{\alpha}_i| + |\beta_i - \tilde{\beta}_i|)
\]

Mitra [7] have proposed another new color features for image indexing called color correlograms. The color correlogram is a set of values \( \gamma_{c_i,c_j}^{(k)} \) that denotes the probability that a pixel of color \( c_i \) have distance \( k \) far from a pixel of color \( c_j \), Computing the correlogram for \( n \) distances and \( c \) colors of quantization, the whole feature size is \( c^2 n \), which is rather high. To reduce the feature size, a revised feature, called auto correlogram, only records the distance \( k \) that each pair contains the same color. The auto-correlogram is a set of values \( \gamma_{c_i,c_j}^{(k)} \) that count the probability that a pixel of color \( c_i \) has distance \( k \) from another pixel of the same color \( c_j \). Then, the feature size will reduce to \( cn \).

The represent of image content is designed for human perceptual. There will generate a balance problem. If the representation is too rough then it can cause misjudge that find out too many irrelevant images. For example, the color histogram cannot describe the spatial information of image content, so color histogram have more misjudge than CCV. On the other side if the representation is too strict then it will cause false drop that relevant images may become large distance. If the objects’ position of an image had been changed, in human’s perceptiveness these two images still very close. But, the auto-correlogram will magnify the distance between these two images.

In this paper we propose a new feature representation that more correspond to human’s thinking point. And, by our experimental result also show that the new representation have better performance than previous work.

3. Color coherence quantity

The color histogram evaluates two
images by color ratio, if two images have the same color ratio then they are most similar. However, as figure 1, a red mass and many red spot may have the same red ratio and color histogram method cannot distinguish that they are different. The CCV method classify they into two groups will improve the performance than color histogram, but will falls into the same problem. Due to the CCV method only divided the color histogram into two classes. The uniform color region whether is below or up a threshold. If all uniform color regions are greater than the threshold then it will falls into the same problem with color histogram.

We propose a novel method to describe the coherence of color more detail and the feature space is concise. The main idea is that we first gather the pixels which have similar color and are neighborhood as a group and we describe there are how many group of a specify color. We use a value call coherence quantity to denote the divergent degree of an uniform color. If the group number of a color is small then the coherence quantity is small, otherwise the coherence quantity is large. Thus, we can use the coherence quantity to measure the different between two images although they consist of the same ratio of colors.

In this paper, the major consideration of us is how to describe the spatial information for the representation. And the representation must be concise for efficient computational time. The distribution of HSV color space is more corresponded to human’s expectancy. So we use HSV color space to evaluate our systems.

The color histogram $H_I$ of image $I$ is a vector $(h_1, h_2, \ldots, h_n)$, in which each bucket $h_j$ counts the ratio of pixels of color $c_j$ in the image. The coherence quantity $D_I$ of image $I$ is a vector $(d_1, d_2, \ldots, d_n)$, in which each bucket $d_j$ record the divergent degree of color $c_j$. First, we cluster the neighbor (up, down, right, left) pixels that have the same color after quantization. There are $k$ group of color $c_p$. $m_i$ denote the ratio of group $i$ divided by total number of color $c_j$. Each $d_j$ is defined by formula (1).

$$d_j = \frac{1}{\sum_i (m_i)^2}$$

For example, figure 1(a) the coherence quantity of color red is 1, and figure 1(b) the coherence quantity of color red is 5. We can easy to describe that two images are distinct although two images have the same color ratio.

Our feature representation contains two vectors $H_I$ and $D_I$. Color histogram $H_I$ describes each different color’s ratio and lack spatial information to describe each color’s distribution. Coherence quantity $D_I$ describes the divergent degree of each color that contain spatial information can assist the lack of color histogram and don’t be influent by image size change, rotation or moving.

The distance between two image $I$ and $I'$ based on our new representation is defined as follows:

$$\text{Distance}(I, I') = \sum_{i=1}^{n} \left( |h_i - h_i'| + \lambda |h_i - h_i'| \times |d_i - d_i'| \right)$$

Where $h_i$ is the ratio of color and $d_i$ is the divergent of color, respect. $\lambda$ is the weight of spatial relationship. In our experiment, $\lambda=0.4$ had better result.

4. Experimental results

In this section, we show some experimental results and evaluate the performance of our proposed techniques. The image database used in our experiment contains 9916 images download from [6]. We use the HSV color space. When computing histograms in the HSV space, the hue color component is quantized uniformly into 16 levels and the saturation and value components are each quantized uniformly into 4 levels, i.e. the number of histogram bins in HSV space is 256. And we use the same test data pair as[3] to evaluate the performance. The test data contains 52 query-and-answer pairs. The pairs were selected to represent various situations such as different views of the same scene, small
scale’s change, spatial translation, and small change of lighting. Each query Q at least contains one correct answer. Figure 2 is an example of query-and-answer pair.

To evaluate retrieval performance, we adopted the R-measure and P-measure to compare the performance of different methods[15]. The R-measure is the average rank of the correct answer over all queries. The small the value of R-measure the better the system performance is. The P-measure is the average precision at recall equal to 1. The bigger the value of P-measure the better the system precision is. The formulas as follow are R-measure and P-measure, respective.

\[
\text{R-measure} = \text{average}(\text{rank}(q_i)).
\]
\[
\text{P-measure} = \text{average}(1/\text{rank}(q_i)).
\]

In Table 1, we report the average rank and precision, averaged over all queries, for the conventional histogram, the CCV histogram and we proposed CCQ histogram, in HSV color space, L1 distance measure. It shows that the retrieval effectiveness of the CCQ histogram is better than that of the histogram and CCV. The CCQ histogram descriptor improved both R-measure and P-measure at the same time. Thus, the CCQ method we proposed contains more information than CCV method and uses the same space to record the information.

<table>
<thead>
<tr>
<th></th>
<th>Histogram</th>
<th>CCV</th>
<th>CCQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-measure</td>
<td>15</td>
<td>7.7</td>
<td>6.36</td>
</tr>
<tr>
<td>P-measure</td>
<td>0.708</td>
<td>0.746</td>
<td>0.752</td>
</tr>
</tbody>
</table>

Table 1: Retrieval performance

5. Conclusion

In this paper we proposed a novel color image representation. It is easy to compute and has concise feature space. We do not want to substitute any other method, but by experiment result show that proposed method can improve the performance. Our representation can easy to add any other method to improve the spatial information of images.

During the experiment, we find that some images are similar to the query image, but they have low rank. The problem is some color had been quantized into different buckets but in human’s perceptual they are the same color. And, some color in human’s perceptual is different but they had been quantized into the same buckets.

How to quantize the color space to correspond human’s perceptual is a complex problem. The colors exist the how similar problem. In the future, we propose to conquer this problem by fuzzy model. One color of image may be quantized into several buckets simultaneously to improve our system.

References


