

A study on the Image Retrieval Technology Based on Color Feature Extraction

Wenlie Zhu^a, Jing Chang^{*b}, and Zilan Hu^c

South China Business College Guangdong University of Foreign Studies, Guangzhou
510545, China

^a970931754@qq.com, ^{*}Corresponding author
Email: ^b78977216@qq.com, ^clan742044787@qq.com

Keywords: Image retrieval, Color features, Color histogram; HSV color space

Abstract. The text-based image retrieval technology is sufficiently mature now, but it still fails to be accurate. It is urgent to further investigate into the content-based image retrieval technology which is quite new and widely applied to a variety of fields. As color is one of the fundamental features of image, the retrieval based on the color features of image can effectively improve the efficient. In this paper, we analyzed and studied the color-based image retrieval and verified the universality of CBIR system in application with nighttime license plate identification case. To sum up, CBIR has a promising future in application. With the future development, it is believed to have higher retrieval efficiency and similarity when meeting the demand of people for image retrieval so that the users can rapidly and accurately locate the image resources they want against a sea of information and better help can be provided for the image classification.

1. Introduction

Being more vivid than text, images are more likely to be accepted by people. As the quantity of digital images grows at a rocketing speed, how to retrieve the data and information demanded by the clients from a sea of image databases remains as an important issue in the present image application field. The development process of image retrieval technology can be roughly divided into two stages: text-based image retrieval and content-based image retrieval.

The text-based image retrieval, namely KBIR (Keywords-Based Image Retrieval), is already quite mature in research and influential in application. The fundamental principle is to design some properties to the image, such as text information, then link them in the relationship database to create formatted data, and finally develop some indexes with the image file name and keyword properties related to the image. As shown in figure 1.

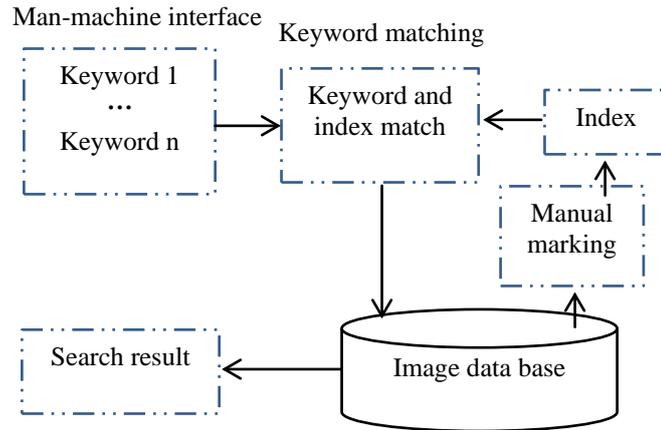


Figure 1 Text-based retrieval flowchart

The content-based image retrieval or CBIR is what we call “image-based image retrieval”, which stays as an important research direction in this field. The user can search the target image by input a related one. In such case, the system will extract the features of the retrieved image and compute the similarity between those features with that in the feature library before outputting the retrieval result. Therefore, CBIR conquers the problem concerning “semantic gap” on behalf of people so that they can find out the target image with higher efficiency.

This technology features not only rapid process, simple operation and spared artificial annotation but also gradually improved retrieval accuracy, so it gets widely applied in different image retrieval systems. To establish a CBIR system, it demands interdisciplinary research and integrated employment of the theories and methods about statistical learning and machine vision so that such system can both establish effective connection with the upper semantic information with lower visual features of images and retrieves the images demanded by the users.

2. Color feature extraction methods

Color features are the most widely applied visual features in image retrieval also the most important features perceived by people in image identification. They are also one of the most distinct characteristics of yarn dyed fabric. The major reason is that color is usually much related to the scene and object contained in the image. In the natural world, objects of same type usually share same color features, whereas different objects may have distinct color features. Color features are usually viewed as the easiest and most effective ones in distinguishing objects and scenes. Compared with other visual features of images, color features depend scarcely on the direction, dimensions and perspective of images and have higher robustness.

2.1 Color Histogram

Color histogram is the method that is most widely used to express the color features. It is first proposed by Swain and Ballard in 1990, which divides first the color space into several fixed subspaces and then figures out the number of pixels in each subspace of each image. What it sets to depict is the proportions of different colors in the whole image. This method enjoys such strengths as simple computation and insensibility to translation and rotation, but it can’t describe the object in the image due to lack of spatial location information about the colors.

The color histogram is also a probabilistic method that computes the number of pixels of each color and expresses it in the form of histogram. The specific formula is as follows:

$$H(P)=[h(x_1), h(x_2)] \quad (1)$$

This color histogram featuring stability in spite of rotation, dimensions and translation and simplicity in computation is widely used in image retrieval. Nevertheless, it also has its own defects, because it contains no information about the spatial locations of different colors and it may be same or similar for different images although it is an overall way of image expression.

2.2 Color correlogram

Color correlogram is another way for expressing the color distribution of image. It not only describes the proportions of each color's pixel quantity in the whole image but also reflects the spatial correlation among different colors. Experiments reveal that compared with color histogram and color coherence vector (CCV), color correlogram generates higher retrieval efficiency, especially for those images with consistent spatial relationship.

One simplified variant is color auto-correlogram which examines only the spatial relationship of the pixels of same colors and thus reduces the spatial complexity to $O(Nd)$. Suppose an image is marked as $I(x, y)$ in which x and y are its spatial coordinates and it contains following colors: $C_1, C_2, C_3, \dots, C_n$. The distance between two colors is expressed as d . Then a histogram will be generated, in which the number of bins will be the square result of n (number of color combinations) and the size of each Bin is: $\text{Bin}(C_i, C_j) = \sum_{x, y} \{ \| I(x, y, C_i) - I(x, y, C_j) \| = d \}$ Where $\|*\|$ indicates the spatial distance between two pixels with values being C_i, C_j . The number of such pixels is to be figured out. As a result, when different distances are set: $d_1, d_2, d_3, \dots, d_m$ (D in total), the dimensions of Bin will be $(n*n*D)$. To go further, when we consider only the spatial relationship between same colors, color auto-correlogram will be produced in which the dimensions of Bin will be $(n*D)$. Color correlogram can be viewed as a table of colors indexed to $\langle x, y \rangle$, in which no. k component of $\langle x, y \rangle$ indicates the possibility that distance between the colors with pixels being $c(x)$ and $c(y)$, respectively, is below k . If correlation between any colors is considered, the color correlogram will be extremely complicated and tremendous (with space complexity being $O(N^2d)$).

2.3 Color moment

A color moment is an easy but effective method of color feature representation, which is first proposed by Stricker and Orengo. Its mathematic basis is that the distribution of any color in an image can be expressed with its moments. Besides, since the color distribution information mainly exists in the low-order moments, the color distribution of an image can be expressed with the moments of three orders: mean, variance and skewness. Compared with color histogram, this method also stands out with its sparing of feature vectorization. In actual application, color moments are usually used for filtration and scope narrowing before other features are used to retrieval the image.

2.4 Color coherence vectors (CCV)

Color coherence vector is a color feature that contains the spatial information about color distribution. It overcomes the failure of color histogram in expressing the spatial locations of image colors. It is actually a kind of more complicated color histogram that categorizes each pixel into coherence pixel or incoherence pixel: the former means the pixel belongs to a large connected area while the latter means the pixel is in a small connected area. The connected area usually accounts for 1% of the pixels in the whole image.

3. Color identification algorithm in HSV color space

3.1 Conversion from RGB to HSV

In image processing, the most frequently used color space is RGB model. This model is usually applied to color display and image processing with its easily comprehensible 3D coordinates. By contrast, HSV model is a color mode that targets users' visual effect and focuses on color

expression such as color, brightness, and darkness. In image retrieval, HSV plays a substantial role in the segmentation of specific colors.

(1) Optical analysis of RGB color space

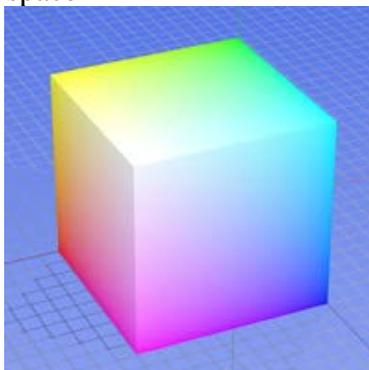


Figure2 RGB 3D coordinate model diagram

In the 3D coordinate model of RGB, the central axis from point of origin to the white apex is the greyscale line, r, g and b are equal, and intensity can be expressed with the vectors of three components. Thus, following conclusions can be drawn:

- Change color: the line from max component apex of three coordinate axes RGB to the apex of yellow-purple-green YMC;
- Change in shade: the distance from RGB apex and CMY apex to the point of origin and central axis of white apex;
- Change in lightness: the location of point on central axis which is dark when approaching the point of origin and bright when approaching the white apex.

According to an optical analysis, the mixing of three primary colors RGB can form other colors. It doesn't mean the light of other colors results from the mixing of the light of three primary colors. As a matter of fact, every monochromatic light owns its unique spectrum. For instance, yellow light is a monochromatic light, but mixing of red with green can produce yellow. This phenomenon is caused by the sensory system of people and related to their physiological system. We can only say that "when being compounded at different ratios, three primary colors can produce equivalent color vision as the visible light of different frequencies for human eyes."

(2) Optical analysis of HSV color space

It can be seen from the HSV 3D coordinate model that HSV color space appears to be in the shape of an inverted cone. HSV is a more visual color model and gets widely applied in many image editing tools.

HSV model corresponds to a conical subset in the cylindrical coordinate system while the top plane of the cone to $V=1$. It includes three planes: $R=1$, $G=1$ and $B=1$ in the RGB model, and the colors represented by it have higher lightness. Hue (H) is given by the rotation angle around axis V. Red corresponds to angle 0° , green to 120° and blue to 240° . In HSV color model, the difference between each color and its complementary color is 180° , and the value of saturation is from 0 to 1, so the radius of cone's top plane will be 1.

The color domain represented by HSV color model is a subset in CIE chromaticity diagram. In this model, the color with 100% saturation is usually not 100% pure. At the apex of the cone (namely the point of origin), $V=0$ and H and S are not defined, so it means the apex is black. At the center of cone's top plane, $S=0$, $V=1$ and H is not defined, so it means the center is white. From this point to the point of origin, the grey scale declines all the way. For all those points, $S=0$ and H remains undefined.

(3) Connection and conversion between RGB and HSV

When understood from a more visual way, HSV conical model can be created after an RGB 3D coordinates have the axle wire erected and then undergo flattening processing. Nevertheless, V is

not directly related to the intensity, because it selects only one largest component of RGB, while RGB can reflect the change in intensity (or greyscale) of the illumination.

Suppose (r, g, b) are the coordinates of red, green and blue of a color and their values are real numbers ranging from 0 to 1, max is equal to the highest one among r, g and b , and min is equal to the lowest one among them. To find out the values of (h, s, l) in HSL space in which $h \in [0, 360)$ refers to hue angle while s mean the saturation and l the lightness, the computational formula is as follows:

$$h = \begin{cases} 0^\circ & (if\ max =\ min) \\ 60^\circ \times \frac{g - b}{max - min} + 0^\circ & (if\ max =\ r\ and\ g \geq b) \\ 60^\circ \times \frac{g - b}{max - min} + 360^\circ & (if\ max =\ r\ and\ g < b) \\ 60^\circ \times \frac{g - b}{max - min} + 120^\circ & (if\ max =\ g) \\ 60^\circ \times \frac{g - b}{max - min} + 240^\circ & (if\ max =\ b) \end{cases} \quad (2)$$

$$s = \begin{cases} 0 & if\ max =\ 0 \\ \frac{max - min}{max} = 1 - \frac{min}{max} & otherwise \end{cases} \quad (3)$$

$$v = max \quad (4)$$

3.2 Application of HSV in image processing

As a more visualized color model, HSV is widely applied in many image-editing tools by indicating the color distance with components H and S. The so-called color distance means the difference between two colors in value. Androutsos et al roughly divided the HSV color space through verification: the bright color area with lightness over 75% and saturation over 20%, black area with lightness below 25%, white area with lightness over 75% but saturation below 20%, and the color area constituting the rest. For areas of different colors, mixing the H with S and setting thresholds can realize simple segmentation.

3.3 Realization of HSV color identification algorithm

HSV color model is proposed by Munseu. Defining the color space from the perspective of human perception, HSV color space have three properties, namely hue (or tinge), saturation (purity of color) and value (indicating the color shade). The color with higher value is closer to white and the color with lower value is close to grey or black. HSV color space is natural and quite close to the visual perception of human beings. It reflects the way who people observe colors, so it is helpful for image retrieval.

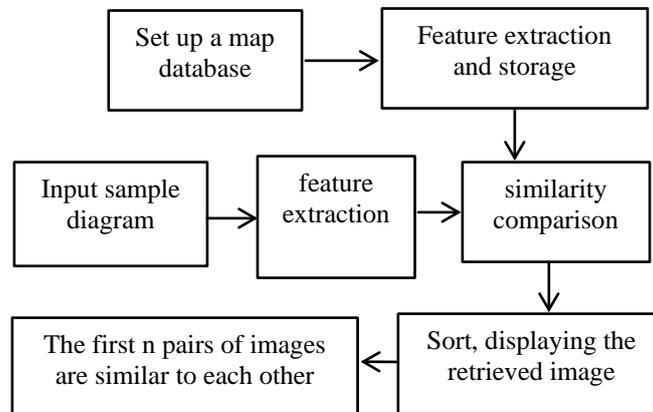


Figure 3 HSV algorithm

3.4 License plate identification during nighttime with HSV

HSV color model can be used to distinguish the license plate area from a number of noises such as the headlamps and their beams and lane lines on the roads as per the image features of vehicle in nighttime. It is able to rule out the areas with high saturation but low consistence with color features of license plate. Improvements are suggested in terms of max/min entropy optimization threshold segmentation method adopted in character segmentation as well as the foreground and background distribution in the character area. Experimental results suggest this method is simple and convenient. The positioning method remains effective for the nighttime images containing head light and taillight, which can effectively reduce the adhesion and isolated noise points concerning single character. The system enjoys high accuracy of identification.

In an age with increasingly busier traffic and profound traffic research, license plate identification is attached with high importance and attention. The key techniques involved in the identification of license plate include image acquisition, license plate area locating character segmentation and identification. The license plate identification is harder when license plate information become missing due to switching of highlight during nighttime. Therefore, targeting the insufficiently illuminated license plate images affected by light and noise, this paper offers a way of image processing that can improve the license plate locating and identification accuracy during nighttime. Some frequently adopted algorithms now include:

- Use Sobel boundary operator to detect the vertical boundary and locate the license plate;
- Employ mathematical morphology to perform differential operation on two adjacent images in order to eliminate the effect of headlight;
- Utilize the methods that locate license plate through such information as color and motion.

As license plate is usually made of some special material, the image can be easily identified when the license plate reflects light. Here we mainly consider the license plate reflecting not light during the nighttime and proposes an effective technique of identification. Among the aforesaid methods, Method (1) puts higher requirements for external hardware setting, which makes the implementation become quite difficult, while Methods (2) and (3) are proper for video images and their target original image of locating and identification is a single feature image of the vehicle and they are not able to get rid of the effects of different noises. Based on those considerations, this paper selects HSV color model that can locate the license plate area as per the image features of vehicle at night. The saturation component can effectively eliminate the noises caused resulting from headlight and highlight the license plate area; while hue component H can mark the areas related to the background colors of license plate so as to accurately locate the license plate target.

When being illuminated with automobile lighting at night, the license plate area features high saturation and constant colors. When being presented in HSV color model, it has high S component. H component can extract the areas related to license plate together. Those two features of the target area are ANDed to obtain the lowest result of blending, and then small areas are eliminated with mathematical morphology [3] so as to get the final target—license plate.

Now there are four kinds of licenses plates in China: white characters on blue background, black characters on yellow background, black characters on white background, and white characters on black background. The colors with highest frequencies are blue, yellow, white and black in turn. If any other color appears on the image, it must have nothing to do with the license plate. When the areas of aforesaid four colors are extracted from the original complete image, it can be inferred that the license plate must be in the image. This area can be called area of interest (AOI). Now color quantification is used to express the AOI in the image with same greyscale. The process is expression is as follows:

Suppose $C(x, y)$ indicate the H, S and V color components in HSV color model of (x, y) in original image X , $C_i(x, y)(i \in R, i \in [0, 4])$ are the H, S and V values of selected color component on

(x, y) , the selected colors include blue ($i=4$), yellow ($i=3$), white ($i=2$), black ($i=1$), and others ($i=0$). When $C(x, y) \in X$, then $C_r(x, y) \in C(x, y)$.

Where T means the tolerance of selected color, M is the span of greyscale quantification, and the result range $g(x, y) \in [0, 255]$.

In the license plate locating stage, H and S components are blended to extract the license plate area and AOI. Saturation component can eliminate the effect of extensive noises caused by automobile light on the license plate image at night, whereas the hue component can rule out the area with high saturation but inconsistency with the background color of license plate. An application of this method to practical nighttime image test indicates it is highly practical in engineering field. What's more, the character segmentation is combined with foreground and background distribution of license plate area and max/min segmentation resulting from the max entropy discrimination segmentation is further revised so that the method can better suit the binarization of license plate area and effectively reduces the effects of character adhesion and isolated point.

4. Improved method of color feature extraction

In view of the missing spatial location information of color in color histogram, a method this divides the image with quadtree to perform clustering segmentation on the color is proposed. In order to highlight the major body (the most interesting part) of an image, an image blocking method containing color optimization is designed and content similarity computing method is provided according to the color features of the image created with quantified color information. Related experiments indicate this method is both simple and effective, for it can utilize the spatial distribution pattern of color to more flexible and accurately represent the characteristics of major content in image and thus improve the efficiency of image retrieval.

The reason why color histogram is not selected as the basis of color feature acquisition is because it lacks spatial information about the image. In order to find out the eigenvalues in the histogram, quadtree may be used to segment the image. Then, the colors in the image are corrected as per the image information and further discretized to produce independent and distinct color features.

The segmentation process with quadtree is as follows: set the number of iteration layers k first to segment the image for k times at most; then establish two queues: one to be combined and one to be split; and next put the original image as a window into the queue to be split, and finally repeat the steps above until the iteration depth of layer K is reached.

The queues to be split is detected while the selected window is divided (with initial window being the original image) to segment the image into four equal shares and mark them as the children of present window. When the colors in one area are found to be consistent in any sub-window, the sub-window will be determined as a target area and will undergo no more segmentation but put into the combination queue for next round and endowed the characteristic color of the parent area as its characteristic color. Other sub-windows are to be put into the splitting queue of next round.

When a vacancy is detected in the combination queue, it means there is a target area. One object may be divided into two windows under the previous window quartering method, so it is necessary check if two adjacent target areas meet the condition for combination. When two areas are similar to each other in term of characteristic color, they can be combined to form a new confirmed area. If the target area is a very small one and there is a larger target area next to it, root nodes between them should be detected during the combination operation. To achieve automatic combination of those areas belonging to the same object into the parent area and alternately process those two queues can generate the regional division of the whole image.

The adoption of quadtree structure greatly cuts down the time complexity, especially for those pure-color images with fewer colors which can be divided more quickly and have much noise

eliminated. The final result of segmentation will be an image composed of color blocks. The color information demanded undergo statistical operation already in the segmentation, and a color distribution histogram measured as per regional area will finally be generated.

5. Conclusion

With the development of modern network technology, image retrieval has become one of the hottest topics in current research. In real life, the number of image libraries keeps growing, error exists in the image perception by human beings' visual system, and users put forward higher demands concerning image retrieval technology, the studies on image retrieval are gaining more and more attention. In this paper, we analyzed and studied the color-based image retrieval and verified the universality of CBIR system in application with nighttime license plate identification case. To sum up, CBIR has a promising future in application. With the future development, it is believed to have higher retrieval efficiency and similarity when meeting the demand of people for image retrieval so that the users can rapidly and accurately locate the image resources they want against a sea of information and better help can be provided for the image classification.

Acknowledgements

"Realization of Image content based Retrieval system",Guangdong University climbing Program 2018(pdjhb0666).“Research on content- based image retrieval technology” South China Business College of Guangdong University of Foreign Studies. “Research on content- based image retrieval technology” South China Business College of Guangdong University of Foreign Studies.

References

- [1] Jitesh Pradhan,Sumit Kumar,Arup Kumar Pal,Haider Banka. A hierarchical CBIR framework using adaptive tetrolet transform and novel histograms from color and shape features[J]. *Digital Signal Processing*,2018.
- [2] Wajdi Besbas. Content Based Image Retrieval(CBIR) of Face Sketch Images Using WHT Transform Domain[A]. *CBEES.Proceedings of 2014 3rd International Conference on Informatics,Environment,Energy and Applications[C].CBEES:,2014:5.*
- [3] Nishant Shrivastava,Vipin Tyagi. An efficient technique for retrieval of color images in large databases[J]. *Computers and Electrical Engineering*,2015,46.
- [4] Xiang-Yang Wang,Hong-Ying Yang,Dong-Ming Li. A new content-based image retrieval technique using color and texture information[J]. *Computers and Electrical Engineering*,2013,39(3).
- [5] Khadem B, Farahzadeh E, Rajan D, et al. Embedding Visual Words into Concept Space for Action and Scene Recognition[J]. 2010.
- [6] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. *Journal of Machine Learning Research*, 2014, 15(1):1929-1958.
- [7] Nair V, Hinton G E. Rectified linear units improve restricted boltzmann machines[C]// *International Conference on International Conference on Machine Learning. Omnipress, 2010:807-814.*
- [8] Zaharia M, Das T, Li H, et al. Discretized streams: fault-tolerant streaming computation at scale[J]. In *Proceedings of the 24th ACM Symposium on Operating Systems Principles (SOSP, 2013:423-438.*
- [9] Raju U S N, George S, Praneeth V S, et al. Content Based Image Retrieval on Hadoop Framework[C]// *IEEE International Congress on Big Data. IEEE, 2015:661-664.*