Image Retrieval Based on Color Feature Similarity

Bably Dolly¹, Deepa Raj²
¹,²Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow

Abstract: In the early days, many user interactive systems performed with basic concepts but such systems are not reaching the user specifications. A huge amount of interest had been taken in the CBIR system. In most recent years, a lot of work has been carried out in the content-based image retrieval system. There are three primitive visual features namely color, texture and shape of an image with reference to content-based image retrieval. Although visual features cannot be completely determined by semantic features, still semantic features are in use because they are easier to integrate into mathematical formulas. In this paper, we focus on the retrieval of images within an extensive image collection based on color projection. We applied a new approach for the retrieval of images within extensive image collections. Therefore good visual feature is one of the important tasks of representing image compactly. Among the visual features, color is the most vital, reliable and widely used features. In this paper RGB color combination considered for retrieval of images. This paper focuses on methods to extract color images quickly for CBIR. First, RGB was quantified in L*a*b and use with the new approach to finding out the most similar images out of given image collection data. The proposed approach shows an efficient image retrieval system which results in an improved system over the prior techniques.

Keywords: Image Retrieval, Image Histogram, Segmentation, Feature Extraction

1. Introduction

Image retrieval techniques are helpful in several image processing applications. Content-based image retrieval systems work with different image datasets analyzed by comparing the query image to the searched image. General techniques for image retrieval are color, texture and shape. These techniques are applied to search for an image from the image dataset [2]. These have not involved with the different image resolutions, size and special color distribution, thus these strategies do not seem to be applicable to the image retrieval system. Many other image retrieval systems use other approaches like color, size and texture. Thus we tend to offer the new read of image retrieval system exploitation both content and data.[7]

2. Color Image

A color image is a grouping of some fundamental colors (Red, Green and Blue). For different color band R, G & B different matrix is generated. An RGB image transformed into an indexed image which decreases the number of colors in the process [1][3]. This process consists of the following methods for approximating the colors in the original image:
2.1 **Color Histogram:** To view the image into different color channels, the Red, Green and Blue color channels are transformed back to 24-bit color channel by building remaining color channels as black. For that reason, each color channel has its individual histogram and the color image can be seen into different color channels separately, figure-1(A). Red, Green and Blue color histograms of the same image are also shown in figure 2(A).

![Fig 1(A) Original RGB Image and Red, Green and Blue color Component images](image1.png)

![Fig 2(A): Red, Green and Blue color Component image Histograms](image2.png)

2.2 **Quantization:** Quantization is the method of decreasing the amount of colors in an image.[5]. The image quantization is *RGB color cube*, which is often used. The RGB color cube is a three-dimensional (3D) array of all of the colors which are used to show for a particular data type, as RGB images can be of type uint8, uint16 or double.

3. **Proposed approach for CBIR system on the basis of color similarity:**

Proposed approach uses combinations of color based features to decrease the problem description. Proposed approach is implemented and concentrated on visual contents of an image specially color applying on the roses dataset. By using this approach, it selected image databases with common feature values. Proposed system extracted all images features separately R, G, B values for problem solving. Proposed approach implemented features like color histogram, color projections Mathematical approaches like mean, median and standard deviation are proposed for efficient retrieval and provides platform to extract images from the database using user query method.
Figure 3: Proposed approach of Content Based Image Retrieval System based on Color Features

Image retrieval is implemented in the following steps:

1. Input an query image. [Input Image indicates Query Image]
2. Convert RGB image to specified color model space.
3. Do Color Segmentation. In this Create Mask by finding Threshold RGB image. The colorspace and minimum/maximum values for each channel of the colorspace were set that result in a binary mask and a composite image which shows the original RGB image value.
4. Initialize output masked image based on input image.
5. Set background pixels, false to zero.
6. Convert masked image to indexed images using color quantization method.
7. Create Color look-up table. It returns the three-column matrix of RGB triplets defining the colormap for the current figure. Each axes within a figure can have a unique colormap.
8. Another feature etc. Taken is color histogram, color moment.
9. Retrieve image to be matched from dataset.
10. Repeat step 2 to step 9.
11. Calculate difference factor of Query image and selected dataset image.
12. Generate score on the basis of minimum difference factor.
13. Sort the image of training dataset on the basis of generated score of each image in it.
14. Show to best top four similar image to Query Image.

Figure 4: Query image (a) before segmentation (b) after segmentation
Input image is shown in figure 4 database of the image set are display in figure 6. Segmented image is shown in figure 5.

Figure 5: different Color Models for input image

Figure 6: Set of Images in database
4. Experimental result:

Methods about a different model like RGB, HSV, YCbCr and L*a*b for segmentation, masked image is taken and return its indexed value and return another color feature as color histogram, color moment.

Table 1: Comparative Study by using different color models in proposed work

| Sn. | Tested Images | Distance vector(d) from Query Image im6 |
|-----|---------------|----------------------------------------|
|     |               | YCbCr | HSV | L*a*b |
| 1   | im12.jpg'     | 0.2900 | 0.2525 | 0.3033 |
| 2   | im1.jpg'      | 0.2366 | 0.1836 | 0.2920 |
| 3   | im10.jpg'     | 0.2463 | 0.1745 | 0.2651 |
| 4   | im11.jpg'     | 0.1447 | 0.1339 | 0.1708 |
| 5   | im13.jpg'     | 0.2901 | 0.2762 | 0.2858 |
| 6   | im14.jpg'     | 0.2789 | 0.2606 | 0.2773 |
| 7   | im15.jpg'     | 0.1526 | 0.1353 | 0.1564 |
| 8   | im2.jpeg      | 0.1762 | 0.1458 | 0.2052 |
| 9   | im3.jpeg'     | 0.1575 | 0.1126 | 0.1897 |
| 10  | im4.jpg'      | 0.2467 | 0.1990 | 0.2570 |
| 11  | im5.jpg'      | 0.1302 | 0.1186 | 0.1207 |
| 12  | im6.jpg'      | 0.0000 | 0.0000 | 0.0000 |
| 13  | im7.jpg'      | 0.3090 | 0.2914 | 0.3052 |
| 14  | im8.jpg'      | 0.3086 | 0.2371 | 0.3259 |
| 15  | im9.jpg'      | 0.2016 | 0.1649 | 0.2185 |
| 16  | ye1.jpg'      | 0.1912 | 0.1917 | 0.3562 |
| 17  | ye0.jpg       | 0.1960 | 0.2269 | 0.3562 |
| 18  | ye11.jpg'     | 0.1692 | 0.2429 | 0.3242 |
| 19  | ye2.jpg'      | 0.2247 | 0.2622 | 0.3562 |
| 20  | ye3.jpg'      | 0.1911 | 0.1881 | 0.3562 |
| 21  | ye4.jpg'      | 0.1790 | 0.1972 | 0.3562 |
| 22  | ye5.jpg'      | 0.2655 | 0.2426 | 0.3562 |
| 23  | ye7.jpeg'     | 0.2013 | 0.1850 | 0.3562 |
| 24  | ye8.jpg'      | 0.3133 | 0.2999 | 0.3562 |
| 25  | ye9.jpg'      | 0.2883 | 0.2440 | 0.3562 |
Table 2: map values of R, G, B component of different images in dataset

| Tested Images | Map value |
|---------------|-----------|
|               | R         | G         | B         |
| QUERY IMAGE   | : im6     | 0.3451    | 0.0471    | 0.0745    |
| im12.jpg      | 0.0549    | 0.0000    | 0.0000    |
| im1.jpg       | 0.0667    | 0.0000    | 0.0000    |
| im10.jpg      | 0.0902    | 0.0039    | 0.0157    |
| im11.jpg      | 0.1922    | 0.0039    | 0.0118    |
| im13.jpg      | 0.0706    | 0.0039    | 0.0078    |
| im14.jpg      | 0.0784    | 0.0039    | 0.0118    |
| im15.jpg      | 0.1961    | 0.0275    | 0.0314    |
| im2.jpg       | 0.1529    | 0.0118    | 0.0118    |
| im3.jpg       | 0.1647    | 0.0118    | 0.0275    |
| im4.jpg       | 0.0980    | 0.0078    | 0.0157    |
| im5.jpg       | 0.2353    | 0.0157    | 0.0353    |
| im6.jpg       | 0.3451    | 0.0471    | 0.0745    |
| im7.jpg       | 0.0510    | 0.0000    | 0.0078    |
| im8.jpg       | 0.0314    | 0.0000    | 0.0000    |
| im9.jpg       | 0.1412    | 0.0000    | 0.0118    |
| ye1.jpg       | 0.0000    | 0.0000    | 0.0000    |
| ye10.jpg      | 0.0000    | 0.0000    | 0.0000    |
| ye11.jpg      | 0.0314    | 0.0000    | 0.0078    |
| ye2.jpg       | 0.0000    | 0.0000    | 0.0000    |
| ye3.jpg       | 0.0000    | 0.0000    | 0.0000    |
| ye4.jpg       | 0.0000    | 0.0000    | 0.0000    |
| ye5.jpg       | 0.0000    | 0.0000    | 0.0000    |
| ye7.jpg       | 0.0000    | 0.0000    | 0.0000    |
| ye8.jpg       | 0.0000    | 0.0000    | 0.0000    |
| ye9.jpg       | 0.0000    | 0.0000    | 0.0000    |

For the purpose of testing of proposed method three categories of databases of specific color images have been taken (see table 4) to compare with the proposed method as well as famous precision and recall measures have been calculated (can be seen in Table 5) for given pre-categorised databases. Both precision and recall are strong evaluation measure. Most of each experiment, image retrieval performed for given every query image. Some of similar images retrieved as per given query image based on proposed system and on the basis of these best measurement calculation have been performed for all images for its performance measurement. The precision of the system significantly shows as in following formula. In this paper, for given Query image Q, if total N number of images has been retrieved so for this, calculation of Precision and Recall can be given below:
\[ P(Q, N) = \frac{\text{no. of relevant images retrieved}}{N} \]
\[ R(Q, N) = \frac{\text{no. of relevant images retrieved}}{N r d} \]

Table 3: Proposed Techniques using given color model

| Proposed Techniques using given color model | HSV       | YCbCr    | L*a*b    |
|-------------------------------------------|-----------|-----------|-----------|
| Input Images                              | Im6       | m6        | m6        |
| Recall                                    | 0.1600    | 0.2400    | 0.4000    |
| Precision                                 | 0.4000    | 0.6000    | 1.0000    |

Table 4: Analysis of Proposed Techniques using L*a*b in different image categories

|          | L*a*b    |          | L*a*b    |          |
|----------|-----------|----------|----------|----------|
| lemon    | lem1      | leaf     | lem9      | leaf9    |
|          | 0.000000  | 0.000000 | 0.003922  | 0.008769 |
|          | lem9      | leaf8    | lem9      | leaf9    |
|          | 0.003922  | 0.008769 | 0.017971  | 0.016169 |
|          | lem14     | leaf5    | lem14     | leaf5    |
|          | 0.017971  | 0.016169 | 0.031859  | 0.033962 |
|          | lem10     | leaf11   | lem10     | leaf11   |
|          | 0.146417  | 0.042418 | 0.150151  | 0.043844 |
|          | lem13     | leaf16   | lem13     | leaf16   |
|          | 0.150151  | 0.043844 | 0.166609  | 0.045733 |
|          | lem12     | leaf7    | lem12     | leaf7    |
|          | 0.166609  | 0.045733 | 0.176253  | 0.071669 |
|          | lem19     | leaf17   | lem19     | leaf17   |
|          | 0.176253  | 0.071669 | 0.205649  | 0.074407 |
|          | lem21     | leaf13   | lem21     | leaf13   |
|          | 0.205649  | 0.074407 | 0.227721  | 0.074407 |
|          | lem8      | leaf14   | lem8      | leaf14   |
|          | 0.227721  | 0.074407 | 0.256735  | 0.083189 |
|          | lem18     | leaf6    | lem18     | leaf6    |
|          | 0.256735  | 0.083189 | 0.384254  | 0.106966 |
|          | lem4      | leaf9    | lem4      | leaf9    |
|          | 0.384254  | 0.106966 | 0.434109  | 0.123140 |
|          | lem7      | leaf9    | lem7      | leaf9    |
|          | 0.434109  | 0.123140 | 0.434498  | 0.140412 |
|          | lem5      | leaf15   | lem5      | leaf15   |
|          | 0.529092  | 0.146364 | 0.529092  | 0.147568 |
|          | lem10     | leaf10   | lem10     | leaf10   |
|          | 0.529092  | 0.147568 | 0.529092  | 0.158181 |
|          | lem11     | leaf2    | lem11     | leaf2    |
|          | 0.529092  | 0.158181 | 0.529092  | 0.162497 |
|          | lem12     | leaf3    | lem12     | leaf3    |
|          | 0.529092  | 0.162497 | 0.529092  | 0.162497 |
|          | lem5      | leaf6    | lem5      | leaf6    |
|          | 0.529092  | 0.162497 | 0.529092  | 0.162497 |
|          | lem6      | leaf8    | lem6      | leaf8    |
|          | 0.529092  | 0.162497 | 0.529092  | 0.162497 |
|          | lem8      | leaf12   | lem8      | leaf12   |
|          | 0.529092  | 0.308710 | 0.529092  | 0.308710 |

Table 5: Calculation of Precision and Recall of Proposed Approach using given L*a*b in different image categories

| Proposed approach using given L*a*b in given dataset | Lemons | Leafs | Flowers |
|------------------------------------------------------|--------|-------|--------|
| Query Image(n1)                                      | Lem1   | Leaf1 | Im6    |
| Number of images in database(n2)                     | 25     | 25    | 25     |
| Recall                                               | 0.16   | 0.24  | 0.40   |
| Precision                                            | 0.40   | 0.60  | 1.00   |
It can be seen from the above experimental result (shown in table 4) that in above three color models HSV, YCbCr and L*a*b. Here L*a*b gives the better result rather than both HSV and YCbCr on the basis of recall and precision. To test it further, here taken the three categorized image dataset of leaf, lemon and flower by using the L*a*b color model with our proposed approach (as in Table 5), where it gives the best result on all these taken dataset.

5. Conclusion

The Content-Based Image Retrieval (CBIR) systems is used to working on different set of image dataset and do not address challenges of large specialized image collections and topics such as efficient image retrieval by image content. In our research work we concentrated on large set of images with dissimilar categories. This work is helping towards set of image retrieval applications. In the proposed approach, by considering the dataset we can choose suitable methods and implement good approaches in large set of images.

6. References:

1. R. Venkata Ramana Chary1, Dr. D. Rajya Lakshmi2 et.al “Feature Extraction Methods For Color Image Similarity”, Advanced Computing: An International Journal, Vol. 3, No. 2, Mar 2012, ISSN 2277-2685
2. Ashraf Al Zudool, Saleh Khawatreh, et.al “Efficient Methods used to Extract Color Image Features Mohammed”, International Journal of Computer Science and Mobile Computing A Monthly Journal of Computer Science and Information Technology, ISSN 2320–088X impact factor 6.017, Vol.6, Issue.12, Dec 2017, pg.7 – 14
3. N. Kim, N. Kehtarnavaz, “DWT-based scene-adaptive color quantization” Real-Time Imaging - Special issue on multi-dimensional image processing, Vol. 11 Issue 5-6, Oct-Dec 2005, Pages 443-453
4. http://matlab.izmiran.ru/help/toolbox/images/color4.html
5. https://www.pantechnsolutions.net/matlab-code-for-image-retrieval
6. Xiang-Yang Wang et. “An effective image retrieval scheme using color, texture and shape features”, Computer Standards & Interfaces, 33 (2011) pg no 59-68
7. Jun You et. El “Content-based image retrieval using color and texture fused feature” Mathematical and Computer Modelling, 54 (2011), 1121-1127
8. Yogita Mistry et-al “Content based image retrieval using hybrid features and various distance metric”, Journal of Electrical Systems and Information Technology, 2017
9. Bably Dolly, Deepa Raj, “Color Based Image Retrieval by Combining Various Features”, ‘International Journal of Engineering and Advanced Technology’, ISSN: 2249-8958 (Online), Volume-9, Issue-2, December 2019, Page No. 454-460
10. Bably Dolly, Deepa Raj, “Various Methods of Enhancement in Colored Images: A Review”, ‘International Journal of Computer Sciences and Engineering’, Volume-6, Issue-7, Jul 2018