# Flag Identification Using Color Feature Technique of Image Retrieval System

**San San Maw**
*Computer University(Monywa)*
*sounhninpwint@gmail.com*

## Abstract

*The growth of digital image and video archives is increasing the need for tools that effectively filter and efficiently search through large amounts of visual data. In content-based image retrieval systems(CBIR) the most efficient and simple searches are the color-based searches, which can be realized in several color spaces and by several color descriptors. This paper present an image classification system which is based on Dominant Color Descriptor (DCD) used in image retrieval system. First step is to calculate the DCD value of the input image. Second step performs the comparing process between the DCD value of input image and DCD values of the images of the system. This system uses flag images of the ASIA region. When a user input a flag image file and the system will classify the image is one of the flags of the ASIA region.*

*Keywords; CBIR ,DCD*

## 1. Introduction

In recent years, the rapid increase in the number of images has led to a parallel increase in the public demand of image retrieval systems having preferably content based search capabilities. These phenomena led to the implementation of many content based image retrieval systems. The CBIR system provides for the retrieval of images by color, texture and shape. Content-based image retrieval seeks techniques to re-encode contents of images and techniques to measure the similarities of two images based on the coding, in order to index and retrieve images which are semantically related to a visual information request without fully understanding the semantics carried by these images.

Color features are among the most important and extensively used low-level features in image retrieval. They are usually robust in noise, resolution, orientation and resizing. Due to their little semantic meaning and its compact representation, color features tend to be more domain independent compared to other features. The color has great importance in content-based image retrieval systems, which is stored in the intensity vectors of image pixels, and this information can be retrieved simply.

## 2. Related Work

Wang Surong, Chia Liang-Tien, and Deepu Rajan introduced a new method which is used to calculate the similarity of Dominant Color Descriptor is discussed. They expressed that better retrieval results can be obtained by using Earth Mover's Distance (EMD) from the original MPEG-7 reference software.[1]

Ot´avio Augusto Bizetto Penatti and Ricardo da Silva Torres presented a comparative study of color descriptors for content-based image retrieval on the Web. Several image descriptors were compared theoretically and the most relevant ones were implemented and tested in two different databases to find out the best descriptors for web image retrieval. Descriptors are compared according to the extraction and distance functions complexities, the compactness of feature vectors, and the ability to retrieve relevant images.[2]

Irena Valova and Boris Rachev expressed that color descriptor is to represent global color features of the images. Their system uses a color code book, created on the basis of vector quantization in RGB space with 9 common colors. The number of the color groups should be a reasonable amount that is easy for people in general to handle while using the system. Since there is evidence from psychology of vision that generally people are not able to identify many different colors, their system use nine color groups.[3]

## 3. Overview of the system

The system is an image-identification program. The system includes a folder containing the flags of the nations of ASIA-region. The system generates the DCD values of the flags of the ASIA region. After that the system accepts the user's input which is a filename of a flag-image-file. The DCD value of the input-flag-file is generated by the system. Then the system determines the input-flag-file is one of the flag files of the nations of the ASIA region by matching the DCD values. If the match is found the system response back the name of the corresponding nation. There are

nations in the ASIA region. The flag files of these nations are stored for the system.

## 4. Image processing step

Image processing steps are applied for color approximation and color mapping.

### 4.1 RGB color space

An RGB color image in an M × N × 3 arrays of color pixels, where each color pixel is a triplet corresponding to the red, green, and blue components of an RGB image at a specific spatial location (see Figure 1.). An RGB image may be viewed as a "stack" of three gray-scale images that, when fed into the red, green, and blue inputs of a color monitor, produce a color image on the screen. By convention, the three images forming an RGB color image are referred to as the red, green, and blue component images. The data class of the component images determines their range of values. The number of bits used to represent the pixel values of the component images determines the bit depth of an RGB image. For example, if each component image is an 8-bit image, the corresponding RGB image is said to be 24 bits deep. Generally, the number of bits in all component images is the same, In this case, the number of possible colors in an RGB images is $(2^b)^3$, where b is the number of bits in each component image. For the 8-bit case, the number is 16,777,216 colors.

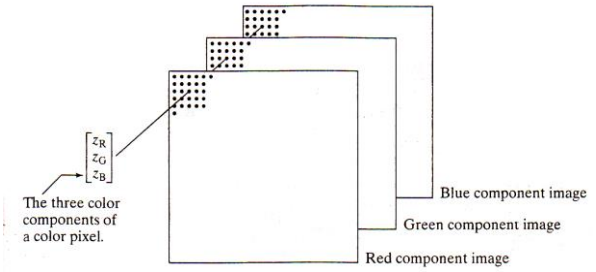



**Figure 1.Pixels of an RGB color image are formed corresponding pixels of the three component images**

### 4.2 Index color space

An indexed image has two components: a data matrix of integers, X, and a color-map matrix, map. Matrix map is an m × 3 arrays of class double containing floating-point values in the range [0, 1]. The length, m, of the map is equal to the number of colors it defines. Each row of map specifies the red, green, and blue components of a single color. An indexed image uses "direct mapping" of pixel intensity values to color-map values. The color of each pixel is determined by using the corresponding value of integer matrix X as a pointer into map. These concepts are illustrated in Figure 2.

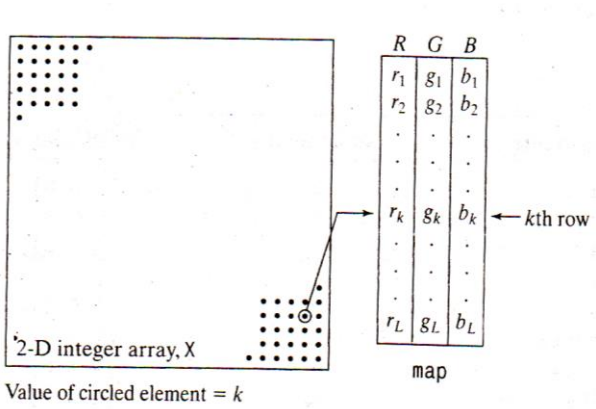



**Figure 2.Elements of an index image**

### 4.3 Color Approximation

To reduce the number of colors in an image, use the color reduction method. This method converts a true color image to an indexed image, reducing the number of colors in the process. This method provides the following ways for approximating the colors in the original image: Quantization Uniform quantization Minimum variance quantization Colormap mapping.

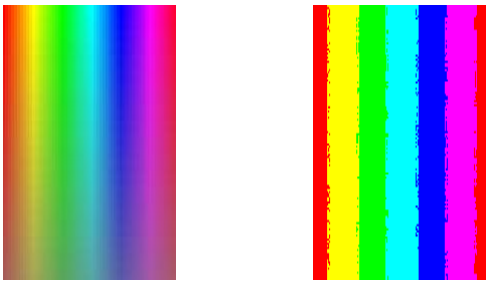


**Figure 3. Original and Reduced colors of a sample picture**

Any image, being processed by the system will be converted to eight basic colors for further use. The following table consists of eight basic color of the system.

**Table 1. Eight basic colors**

| | | | |
|---|---|---|---|
| | (0) Black | | (5) Red |
| | (1) Blue | | (6)Magenta |
| | (2) Green | | (7)Yellow |
| | (3) Cyan | | (8)White |

## 5. Dominant color descriptor(DCD)

In general, the color descriptor of a region is defined to be:

$$Color\ Descriptor=\{\{c_i,p_i\},i=1,\ldots,M\} \quad (1)$$

where,

$M$ = the total number of color clusters in the image.

$c_i$ = an Arabic numeral corresponding to the color

$p_i$ = is the percentage of $c_i$ color

So that $\sum p_i = 1$.

### 5.1 DCD for a region of the system

For the first, the reduced color image has only eight basic colors. The DCD for the region of the system include for the following format expressed in Figure 4. The system will calculate the percentage of the eight basic colors. These values in position become for the DCD value for a location.

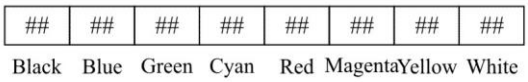



**Figure 4.Color position in a DCD for a region**

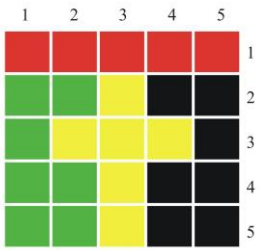



**Figure 5. A region of an image having 5-pixel-width and 5-pixel-height**

**Table 2. Sample calculation**

| Item | No | Percentage(%) |
|---|---|---|
| Red | 5 | 20% |
| Green | 7 | 28% |
| Black | 7 | 28% |
| Yellow | 6 | 24% |
| Total | 25 | 100% |

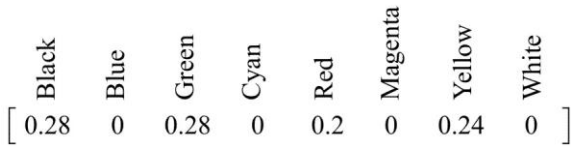



**Figure 6. DCD for a region of the system of Figure5.**

### 5.2 DCD for an image of the system

In the previous section, the color descriptor for the region is mentioned. In this section the color descriptor for the image of the system will be explained. The image is divided to four regions of same size as shown in figure 7.

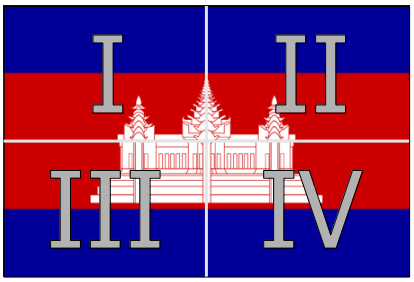



**Figure 7.Four region of an image**

After that, the DCD value of the whole region of the image will be calculated. And the DCD for every quadrant region will be calculated. By placing the DCD of the whole region, DCD of the quadrant I region, DCD of the quadrant II region, and DCD of the quadrant IV region in order, the system will generate the DCD for the image of the system. As one DCD for a region holds eight floating point numbers, there are forty floating point numbers for the DCD value of the image of the system.

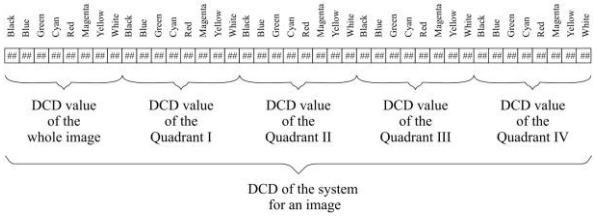



**Figure 8.DCD for an image of the system**

### 5.3 DCD values matching

The matching will be performed by comparing the DCD values for two files to identify the two images are identical or not.

## 6. Overview of the System

The proposed system consists of images. The images are of two kinds. They are the input image and the existing image of the system. The system will accept an image from the user by accepting the file name and path. The system will perform the classification between the input image from the user and the images.

The proposed system uses the color extraction methods of image retrieval system to extract colors from the image. The every image of the system is retrieved to get the color information of the image. The extracted color information is encoded by DCD method to gain DCD value for the image. So there is a DCD value for every image. This DCD value will be used later.

The system will perform the classification of the input image is one of the images of the existing images. The image accepted from the user is used to generate DCD value for the input image. And the proposed system performs the match the DCD values of the pre-existing flag-images' encoded DCD. If the match is found the name of the nation is responded back to the user.

The proposed system consists of two different kinds of processes. The first process is pre-processing steps before the identification of the image is performed. The second process is image identification process to identify the input image is one the pre-existing images.

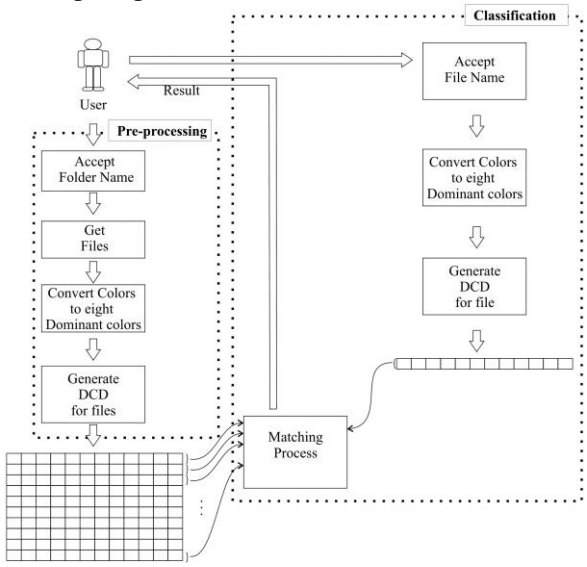



**Figure 9.Overview of the System**

### 6.1 Pre-processing steps

Actually pre-processing is a sub system which has to be performed before the identification is performed. These steps are needed to be performed at the first time or when the user has added images to the image folder of the system. In this process the following steps will be performed.

(1) The every existing images which will be used to be identified as a standard must be stored in a folder. At this step, the user will be prompted to enter the name of the folder which includes the pre-existing images. The images of the folder become the fundamental elements for the whole pre-processing process. These are the fundamental images to be used as a standard for identification.

(2) The images in the folder which are chosen at the previous step are taken in part. For each and every image, the colors in the image files are accessed and converted the colors of that image to eight dominant colors which are the fundamental colors of the RGB color system.

(3) By counting the eight dominant colors in an image, the DCD values for each image are calculated and the system will store all the DCD values of images for further use. These values will be used at the identification steps of the system.

### 6.2 Image Identification steps

Actually, this process is the key process of the whole system. In this process, the following steps will be preformed to identify the image is one of the pre-existing images.

(1) The system prompts to the user to enter a file name as an input file for the system to be classified. This input file will be identified that the image is one of the pre-existing images of the system. In other word, the system will identify that the input image is one of the flags of ASIA continent.

(2) The image file is accessed and converted the colors in that image to eight dominant colors which are the fundamental colors of the RGB color system. The system is aimed to use with eight dominant colors.

(3) By counting the eight dominant colors in the image and placing these values in order, the DCD value for the input image is calculated.

(4) Identification is performed by matching the DCD value of the input file along with each of the DCD value of the pre-existing image files. These values are stored at the pre-processing steps.

(5) This step displays the result of the previous steps. If the match is found, the system displays the corresponding nation's name of the flag-image which is having the same value of DCD with the DCD of the input image. If no-match is found, then the system will display "No-match is found".

## 7. System flow diagram

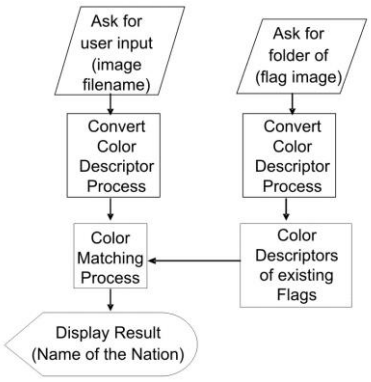



**Figure 10.System Flow Diagram**

The above system flow diagram shows how the system works. The system consists of a folder which is holding the flag-images of ASIA continent, the program and an input image. When the user wants to know that the input image is the one of the flags, he then input the input image to the system. Then the system will response back to the user the name of the nation or no match found message.

The user can easily add new flag image that the system uses. Just copy the new image to the folder which is holding flag-images. After it, user must ask the system to do pre-processing steps in order to get correct answer. This former step is only needed to do when adding new image to that folder. The system keeps the results of the pre-processing steps for further use when the system works next time.

## 8. Data Flow Diagram

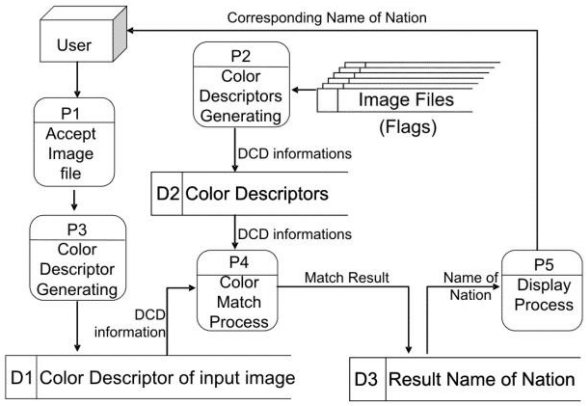



**Figure 11.Data Flow Diagram**

The five processes in the system are

1. Accept Image File Process
2. Accept Folder Process
3. Color Descriptor Generating Process
4. Color Matching Process
5. Display Process.

Accept image file process accept an filename of an input image from the user to be identified the system.

Accept folder process accept the folder which is holding the flag images of ASIA continent.

Color descriptor generating process generates the DCD values of the files.

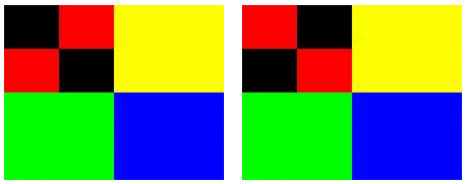
Color matching process compares the two DCD values of the input image and of the one of the flag file of ASIA region.

Display process will generate output to the user.

## 9. Conclusion

This system performs the matching by comparing the two DCD values of the two images. This calculating of DCD and comparing of DCD values are suitable for this system.

There is a limitation in this process while using with other logos and images. If the colors containing in one region is the same as the colors of the same region of the other image, but the locations of colors are not the same, the false result can be occurred.



(a) (b)

**Figure 12.Images having same DCD.**

For the further process, the number of the regions divided must be added to large number and overlapping area between these regions must be included. See figure 13.

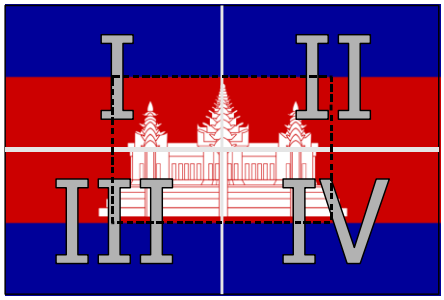



**Figure 13.Updated five regions of an image.**

DCD must include

1. DCD for the whole image
2. DCD for the Quadrant I
3. DCD for the Quadrant II

4. DCD for the Quadrant III
5. DCD for the Quadrant IV
6. DCD for the Quadrant V which is a dash-line region.

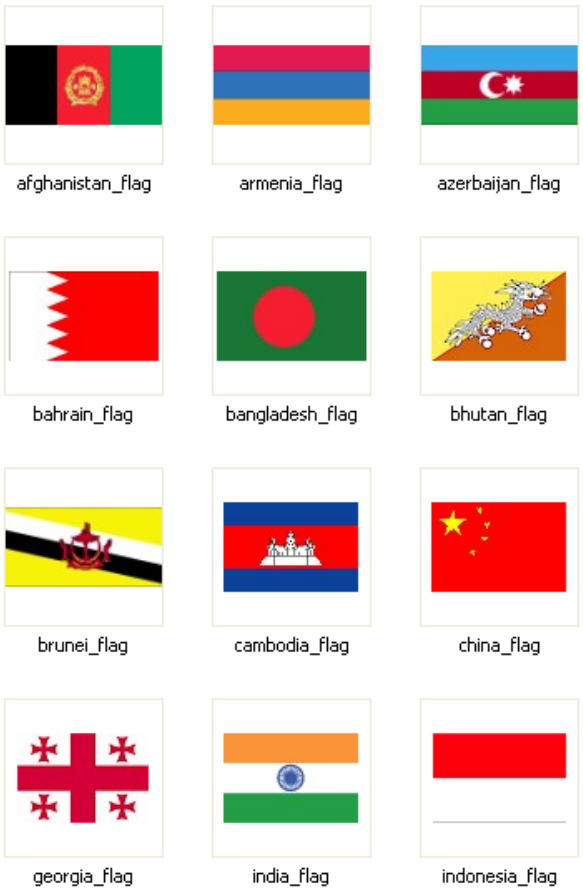
## 10. Result

If the two images compared by the system is having 90% match in DCD values, the system will assume that these two images are equal.

### References


[1] Wang Surong, Chia Liang-Tien, and Deepu Rajan "Image Retrieval Using Dominant Color Descriptor"
[2] Ot´avio Augusto Bizetto Penatti and Ricardo da Silva Torres "Color descriptors for Web image retrieval: a comparative study"
[3] Irena Valova , Boris Rachev "Retrieval by Color Features in Image Databases"


### Appendix(A) : Some Flags of ASIA Continent




### Appendix (B): Sample calculation of Flag of Georgia Nation.