IJCSN - International Journal of Computer Science and Network, Volume 6, Issue 3, June 2017
ISSN (Online) : 2277-5420
www.IJCSN.org
Impact Factor: 1.5

An Expert Technique for Content-Based Color Image Retrieval

Subha Vaiapury

Assistant Professor, CSE Department, RGCET,
Puducherry, India

Abstract: Recently, as web and various databases contain a large number of images, CBIR (content-based image retrieval) are greatly used. This paper proposes an image retrieval system using color-spatial information from the applications. First, we suggest two kinds of the indexing keys to prune away the irrelevant images to given query images using MCS (Major Color Set) and DBS (Distribution Block Signature). MCS's are related to color information while, DBS's are related to spatial information respectively. After successively applying these filters to a large database, we get only a small amount of high potential candidates that is somewhat similar to that of query images. We propose to use QM (quad modeling) method to set the initial weight of 2-dimensional cell in the query image according to each major color and retrieve more similar images through similarity association function associated with the weights. Finally, we evaluated the system's efficiency by statistically how many images were expected to be filtered out during the first and second filtering processes.

Keywords: *Content-Based Image Retrieval, Major Color Set, Global Color Signature, Distribution Block Signatures, Quad Modeling, Hue Saturation Value.*

1. Introduction

In the recent world with the advances in multimedia technologies (such as compression, display, and visualization technologies) and the increased emphasis on multimedia application, the production of image information has resulted in a large volume of images that need to be properly indexed for retrieval in the future

The digital image is becoming available at an everlasting rate and users navigate on this for their requirement. How to effectively retrieve the desired information is most crucial and challenging problem. To help users find and retrieve relevant information effectively, advanced technologies need to be developed for analyzing, representation and categorizing vast amount of images. CBIR (Content Based Image Retrieval) retrieval is the fundamental step for efficient retrieving, browsing and accessing. Literature reports various techniques for CBIR and the most commonly utilized features are color, shape and texture. Based on this observation, this work concentrates on image retrieval using color and spatial information.

2 Related Works

Image classification deals with the problem of identifying image in large Database. It is desirable to classify and categorize image content automatically so that end-users can search and choose or verify a desired image based on the color content. By using CBIR application, color content of an image is identified and different types of images are retrieved. In a recent approach, [1] Michel J.Swain demonstrates the color histogram of multicolored objects provide a robust, efficient use for indexing into a large database of models [2]. Anik-Jain used color and shape content in image for efficient retrieval of images from large databases. In [3], John. R. Smith proposed a technique by which the color content of images and videos automatically extracted to form a class of metadata that is easily indexed.

In [4], F. Liu presents an image model with a new set of features that address the challenge of perceptual similarity. The three resulting mutually orthogonal sub field have perceptual properties which can be desired as "periodicity", "directionality" and "randomness" approximating what is indicated to be the three most important dimensions of human texture perceptions. In [5], Kian-Lee Tan describes the color-spatial technique, the color information is integrated with the knowledge of the color spatial distribution to facilitate CBIR.

In [6], Hun Woo Yoo uses a CBIR method that can search large image database efficiently by color, texture and shape. In [7], John P. Eakins reviews recent trends in the field, distinguishing four separate types of activity: automatic seen analysis, model-based and statistical approach to object classification and adaptive learning from user feedback.

In [8], Hun Woo Yoo proposed a color image retrieval using indexing key and use relevance feedback mechanism. Deng Shang Zhang [9] proposed a region based image retrieval system with high level semantic learning, the system segments an image to different regions and extracts low level features for each region. Quynh Nguyen Huu1 , Ha Nguyen Thi Thu1 and Tao Ngo Quoc2[10] uses a region matching and clustering techniques.

3. Basics of Content- Based Image Retrieval

CBIR is the function of computer vision to the image retrieval crisis that is the problem of probing for digital images in huge databases. Query techniques and content comparison technique are the two types of technique. In content comparison techniques the common methods for extracting content from images are color, shape and texture. **COLOR**: Retrieving images based on color similarity is achieved by computing a color histogram for each image. It has the advantage of achieving fast results Early color-based retrieval systems have used the global RGB histogram information. These systems, however, do not capture the local color information within the image to overcome this problem color and spatial information has been exploited for more accurate retrieval.

3.1Color and Spatial Techniques

We utilize the HSV (Hue, Saturation, Value) color space. HSV is intuitive color space in the sense that each component contribute directly to visual perception.

The human visual system is, more sensitive to hue than saturation and value. So hue should be quantized more finely than saturation and value. In our experiments we uniformly quantized HSV space into 18 bins for hue, 3 bins for saturation and 3 bins for value and by adding 4 gray a total 166 colors are used to represent the images.

3.2 Major Color Sets

The query image is partitioned into 32×32 cells of equal size in each cell, quantized joint HSV histogram is computed to extract the most frequent [highest peak] bin as the dominant color in the cell. In general many images tend to have a small object area [i.e. Region of interest] ROI and a large background. We extract two major colors MCb from the background area and two major colors MCc which are not duplicated from MCb from the ROI.
The four hues from four major colors are represented in bit stream called color signature it contains a one-dimensional array with 19 bits. The first 18 bits are used to represent four hues of the MCS and the final bit is used to represent gray information of the MCS. Initially, all 19 bits are set to 0 MCS color signatures obtained by assigning bits associated with the four hues of MCS to 1and is used in query side.

3.3 Global Color Signature

GCS color signature is used in database side and capture global color information in the image. It contains 19 bits like the MCS. The GCS signature constructed with all hues in 32×32 cells. Each bit of the GCS is assigned to 1 when the corresponding hue exists in the cells.

3.4 Distribution Block Signature

The second filtering process is performed to compute further similarity by considering color spatial information. We will call the bit stream used in this step the distribution block signature. First an image is partitioned into 4×4 blocks of equal size. Each block will contain 64 cells since we partitioned the image into 32×32 cells. Then, with respect to the first color in the MCS, the block is assigned to 1 when it contains same first color, or assigned to 0 when it does not. That produces a 16 bit stream for first major color. The same process is repeated for the rest of the MCS the result in a total of 4×16 bit streams.

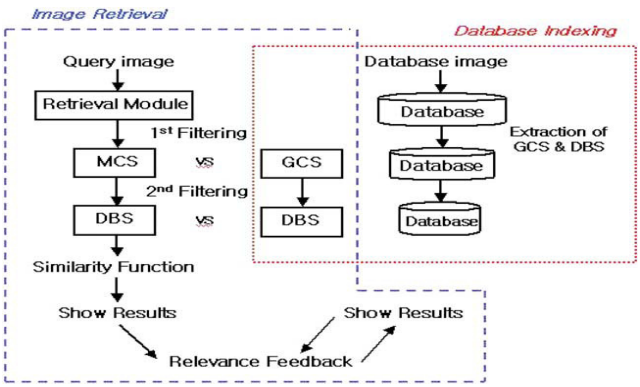

Figure 1: The Overall Retrieval Procedure of the System

4. Filtering and Quad Modeling Techniques

The images are stored in the database and two selective features such as GCS and DBS are extracted. During retrieval once query image is presented, the retrieval system extracts the MCS (MAJOR COLOR SETS) and the DBS of the query. In first filtering, MCS of the query and the GCS of database images are compared using a bitwise logical AND operation, database images, which do not contain all MCS of a query are filtered out and in second filtering database images that do not have a similar spatial

location to that of the query image are filtered out. In this step, a bitwise logical-XOR operation is applied. Note that the XOR operation is 0 only when the two compared bits are both 1s and 0s. Images with similar spatial distribution are chosen as potential candidates.

Among the candidate images a similarity measure is used to perform a final similarity computation for more retrieval a new feedback mechanism based on QM (Quad Modeling) is employed. In order to that, for each major color of a query image 32×32 weight matrix called quad modeling (QM) is computed. It's presented to capture the color, spatial information of a query. A total four 32×32 QM matrices are used to present the 4 major colors of a query. Similarly, in a database images four 32×32 cells are employed. For a color that is identical to a major color of a query, each cell has a 1 or 0 it does not.

4.1 Efficiency

Expected efficiency is calculated using the following equation

$$\left(\frac{{}_{B}C_{C}}{AC_{C}} \right) \times record\_number$$

Where, record number represents the total number of database images. After extracting the GCS bit stream that consists of 19 bits. We noticed that most of 54 images have 3-19 bits with a value of 1. In our database among the 19 bits, 12 bits on average have a value of 1. The number of different GCS: ${}_{A}C_{B} = {}_{19}C_{12}$ In case of the MCS bit stream, we noticed 3bits on average have 1s.The number of different MCS: ${}_{A}C_{C} = {}_{19}C_{3}$

In the first filtering process, from the above, we can filter about 77.3%of database images. Thus, statistically 12.26 images among 54 images can be chosen as candidate image to further similarity.

In the second filtering process we filtered out images with dissimilarity of 43 or higher, which leads to eliminate roughly 33% images from the candidate images remaining after the first filtering. Thus we expect that 8 images among all 54 images remain for final similarity retrieval. Statistically, this means an advantage of 84.75% search space reduction after the first and second filtering.

5. Experimental Results

In order to evaluate the proposed image retrieval system the programs have been implemented in Visual Basic.NET. In our experiments, 54 images were chosen consisting of 15 categories including airplanes, mountains, bears, cheetahs, dolphin, eagles, elephants, horse, scenery, polar bears, roses, sunset, sunrise, tiger, and zebra.

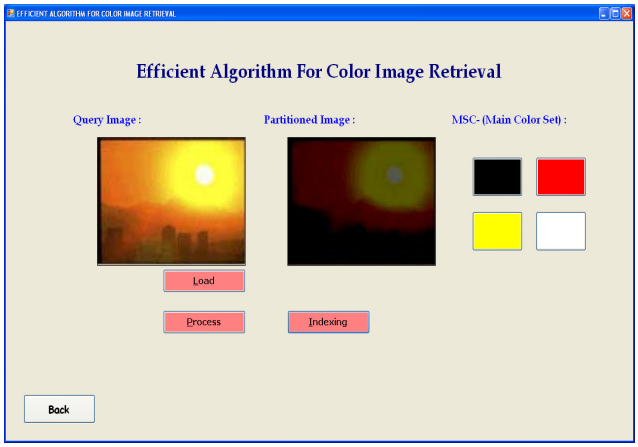

Figure 2: First Filtering

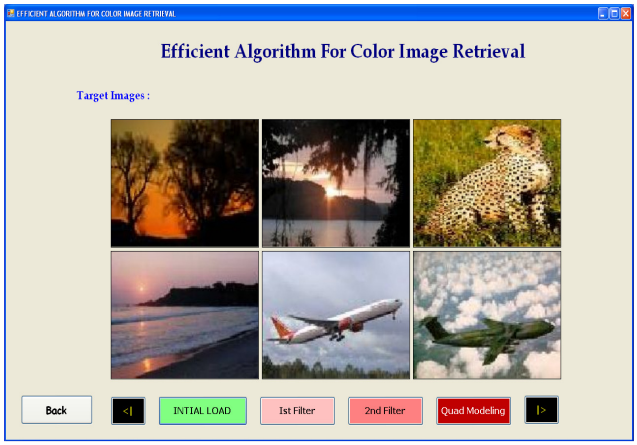

Figure 3: Image Set

Figure 2 shows the process of how a query image is loaded, transformation of RGB to HSV is done. The four major colors are obtained from the partitioned image. In Figure 3, initial load of different category of database images for example, in this case 54 is shown.

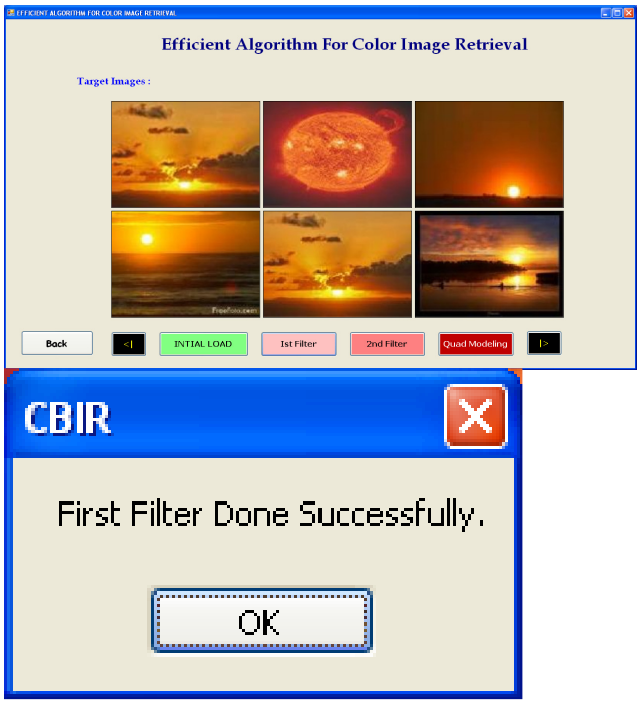

Figure 4: First Filter Result

In the Figure 4, first filtering process that is AND operation between MCS of a query image with a GCS of database images is shown.

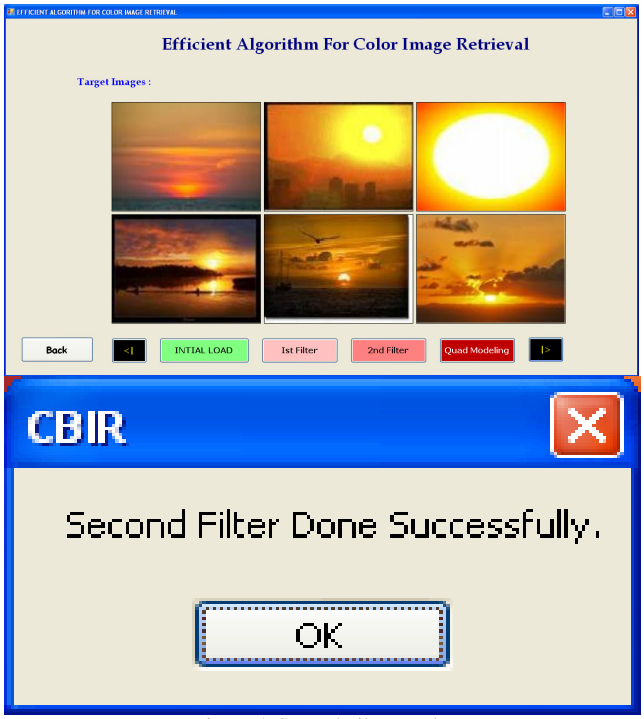

Figure 5: Second Filter result

In Figure 5, the second filtering process that is XOR operation between DBS of query image with DBS of database images is shown.

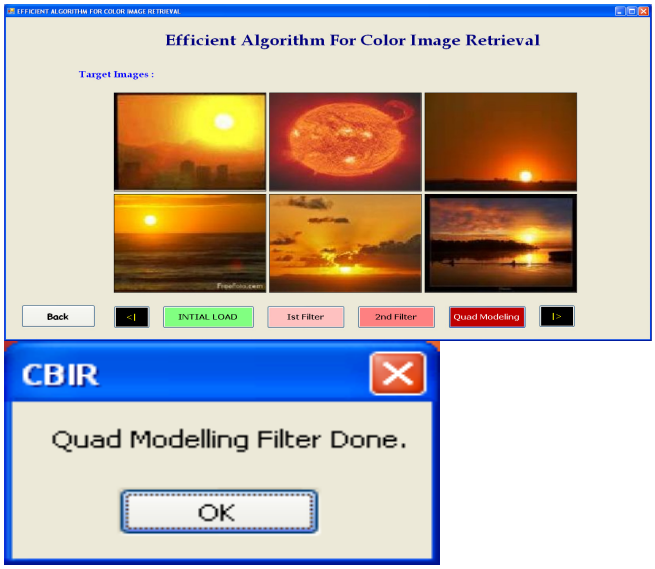

Figure 6: Quad Modeling result

Finally, in Figure 6, the output of quad modeling technique is shown where a query image is retrieved from database images by weighted quad modeling matrix.

6. Conclusion

In this work, we used signature based on the color, spatial approach for addressing the content based image retrieval problem. The chosen MCS, GCS and DBS color signature in the approach is used for efficient retrieval of an image. Further, in this work, we used an HSV color model to represent an image. We have studied and implemented the system up to the first step of QM modeling in VB.Net Environment. In future, similarity measure for QM matrix can be calculated to better capture the amount of overlap between query and database images. Further for accurate retrieval, the relevance feedback approach can be incorporated and performance of the system using precision and recall measures can be used for evaluation.

References

1. Swain .M & Ballard “Color indexing”, International journal of computer vision, 7(1), 11-32, (1991).
2. Anil k. Jain “Image retrieval using color and shape”, Pattern recognition society, 1233-1244, (1995).
3. Liu.Picard “Periodicity, Directionality, and Randomness: World Features For Image Modeling And Retrieval”. IEEE Transaction On Pattern Analysis And Machine Intelligence, 18(7), 722-733, (1996).

4. Smith J. R, "Integrated spatial and feature image query", Pattern Analysis and Multimedia system, 7(2), 129-140, (1999).
5. Kian-Lee, C.S.,Ma "Information Embedding Base On Users Relevance Feedback Technique in CBIR". Pattern Analysis and Application, 3486, 294-304, (1999).
6. John p. Eakins, "Towards Intelligent Image Retrieval", Pattern recognition, 35, 3-14, (2001).
7. Hun Woo Yoo, "Visual Information Retrieval System Via Content-Based Approach" Pattern Recognition, 35(3), 749-769 (2002).
8. Hun-woo, Yoo, "Expert system for color image retrieval." Expert System With Application, 28, 347-357 (2004).
9. Dengsheng Zhang, "Region Based Image Retrieval with High –Level Semantics Using Decision Tree Learning", Pattern recognition, (2008)
10. Quynh Nguyen Huu1 , Ha Nguyen Thi Thu1 and Tao Ngo Quoc2 "An Efficient Content Based Image Retrieval Method For Retrieving Images"ICIC International, 1349-4198, 2823-2836(2012)
11. Priyanka Shinde, Anushka Sinkar , Mugdha Toro , Prof.Shrinivas Halhalli "Image Retrieval Based on its Contents Using Features Extraction "International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 03 Issue: 01 Jan (2016)
12. Mohammad Sadeq Navabi* Zahra Ahmadi Brooghani "Content-Based Image Retrieval through Combined Data of Color Moment and Texture Mohammad" IJCSNS International Journal of Computer Science and Network Security, VOL.17 No.2, February (2017).

Author -

V.Subha is currently working as an Assistant professor in RGCET in department of computer science, puducherry, India. Her research interests include image, video processing, computer vision, big data and networks.