

New Approach for Content based Image Retrieval System using Texture and Color Features

Alka Choudhary

Assistant Professor

Department of Computer Engineering, YMCA
University of Science & Technology, Faridabad,
Haryana, India

Savita Choudhary

Department of Computer Science and Engineering,
MVN University,
Palwal, Haryana, India

ABSTRACT

Now-a-days Content based Image Retrieval (CBIR) techniques are in great demand in every field. In this paper a new approach is proposed for retrieving images. Color Features is extracted from HSV image using color histogram. The texture feature gets extracted from the RGB image by applying GLDM technique. Feature vectors of query image and database image are compared using similarity measures. After comparisons distance vectors are added to get a resultant vector. The images are ranked according to the resultant vector and retrieved. This approach is implemented using MATLAB8.0. The results and conclusions are shown in the paper.

Keywords

Content-based Information Retrieval, color model, HSV, Euclidean distance, texture, GLDM.

1. INTRODUCTION

Information retrieval is a field dealing with how the information is represented, stored and how it is accessed. Broadly, we can retrieve the information by two methods, the first one is Texted based Information Retrieval System(TBIR) in which the information is retrieved in the form of text and the another one was content based information retrieval in that information is extracted on the basis of visual content [1]. Content Based Information Retrieval System is our main focus of the paper.

2. CONTENT BASED INFORMATION RETRIEVAL

Content Based Information retrieval (CBIR) is a technique which retrieves the image based on visual content. It retrieves the image on the basis of feature extraction like color, texture and shape etc. CBIR is used for retrieving similar image from database. In order to understand it more closely CBIR System is a set of modules that interact with each other to retrieve the image from database as per the user query. The introduction of the Internet with its massive amount of information has lead to a high focus on the science of information retrieval as a mean to organize, store, representing and accessing these information items. Fig.1 shows the basic architecture of CBIR system. The main modules in a CBIR system are -feature Extraction, similarity comparisons, indexing and Retrieval. In a CBIR system first the features of input image and database image are extracted. These features are compared using similarity measures. Finally, on the basis of comparisons the images will be indexed and retrieved.

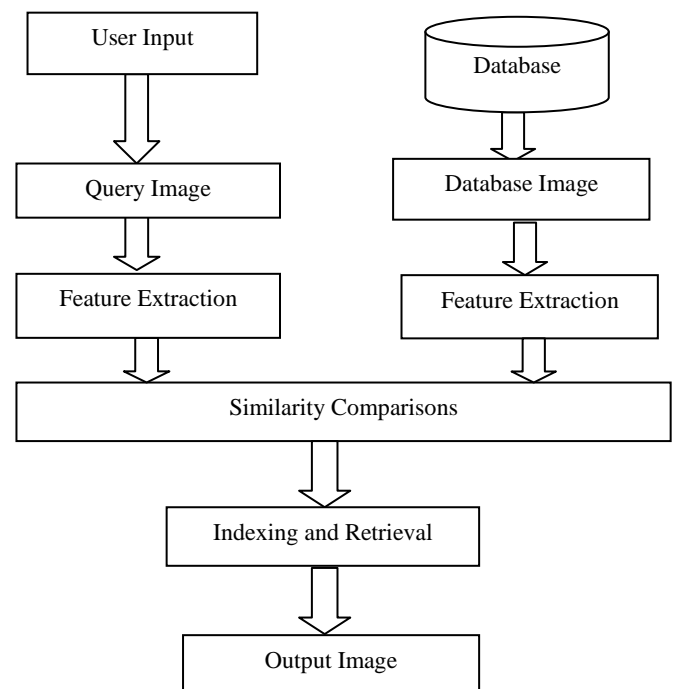


Fig 1: Block diagram of CBIR System

3. FEATURE EXTRACTION

Feature extraction is the process of creating or transforming from the original data within the visual feature system, feature can be further classified as general feature and domain specific features. Feature extraction is the basis of any content –based image retrieval technique. In a broad sense, features may include both text-based features and visual features like color, texture, shape, spatial layout. Within the visual feature scope, the features can be further classified as low-level features and high-level features. The selection of the features to represent an image is one of the keys of a CBIR system. Because of perception subjectivity and the complex composition of visual data, there does not exist a single best way to represent a given visual feature. Image feature can be either global or local. A global feature uses the visual features of the whole image, whereas a local descriptor uses feature of the regions or objects to describe the image content. To obtain the local feature, an image is often divided into parts first. The simplest way of dividing an image is to use a partition, which cuts the image into tiles of equal size and shape. A simple partition does not generate perceptually meaningful regions but is a way of representing the global features of the image at

a finer resolution. A better method is to divide the image into homogenous regions according to some criterion using region segmentation algorithms that have been extensively investigated in computer vision.

3.1 Color

Color is a perception that depends on the response of the visual system to light and the interaction of light with object [3]. Color feature generally adopts histogram to describe it. Color study consists of the physical properties of light which give rise to color, the nature of the human eye and the ways in which it detects color, the nature of the human vision centre in the brain, and the ways in which messages from the eye are perceived as color [4]. Color histograms methods have the advantage of speediness, requires low memory space.

3.1.1 RGB color model

The RGB color model has three basic primary colors: red, green and blue. All other colors are obtained by combining them. This model can be thought as a cube, where three non-adjacent and perpendicular corners are R, G and B. As can be seen, RGB is an additive color model, since the combination of green, red and blue gives white. This is the color model that is most commonly used in computer graphics, since it matches the way the color is stored in video memory. In general application, we require 8 bits for presenting a color. For RGB, we need $= 8 \times 3 = 24$ bits, so total number of colors $= 2^{24} = 16,777,216$. But this large number of colors are practically not usable that why we use only 256 colors. The set of colors that can be used instead of whole colors is called subset of colors. This is also called safe RGB colors.

3.1.2 HSV color model

The HSV stands for the Hue, Saturation, and Value. The hue-saturation-value space is a cone for a given point, h and s v are the angular and radial coordinates of the point on a disk of radius v at height v ; all coordinates range from 0 to 1 [6]. Hue is defined as an angle in the range from 0 to 360 degrees, the corresponding colors vary from red through yellow, green, cyan, blue, magenta, and black to red. S , is the depth or purity of color and is measured as a radial distance from the central axis with value between 0 at the center to 1 at the outer surface value, or brightness, varies from 0 to 1.0, the corresponding colors become increasingly brighter.

3.2 Texture

Texture is the innate property of all the surfaces that describes visual patterns and that contain important information about the structural arrangement of the surface and its relationship to the surrounding environment. A texture element or texture pixel is the fundamental unit of texture space textures are represented by Texel which are then placed into number of sets, depending on how many textures are detected in the image. These sets not only define the texture but also where it is located [5]. Here we have used grey level difference matrix by calculating four coordinates. GLDM (Gray Level Difference Matrix) calculates the gray level difference matrix and the probability density function for a given image. This technique is used for extracting statistical texture feature of digital image.

3.3 Shape

Shape is the characteristic surface configuration that outlines an object giving it a definite distinctive form. shape does not refer to the shape of an image but is the particular regions that is being targeted. Shape can be determined by segmenting the image or by edge detection of an image. Shape filters can also be used for detecting the shapes in the image.

4. SIMILARITY MEASURES

As there are many methods to find similarity measures between images based on some feature set so it is a difficult task to find one good similarity measure. On the one hand, the ultimate goal is to define similarity functions that match with human perception, but how humans judge the similarity between images is again a typical task. The Euclidian Distance between a database image and query image has been used in this paper. The other similarity measure techniques are Minnow ski distance, manhattans distance and mahalonobis distance.

5. PROPOSED WORK

In this paper we have proposed the new algorithm using the combination of color and texture features extraction techniques. Here we have used color histogram, Euclidean distance, HSV (Hue Saturation Value) color model and GLDM (Grey level dependence Matrix) texture technique to retrieve the similar image from the database Image collection.

5.1 Proposed Algorithm

The proposed algorithm has following steps:

1. Read the database image DI and query image QI. both images are RGB color images.
2. Convert DI and QI image into HSV image as DI' and QI' respectively.
3. Find the color histogram Hist1 and Hist2 of DI' and QI' respectively.
4. Find the Euclidean distance ED1 between Hist1 and Hist2.
5. Calculate GQI and GDI Gray Level Difference Matrix (GLDM) of QI image and DI image.
6. Find the Euclidean distance ED2 of GQI image and GDI image.
7. Add ED1 and ED2 ($ED = ED1 + ED2$) and Save the results based on ED. (Image having minimum ED will be ranked 1st).
8. Repeat step 1 to 7 till all the pictures in the database has been covered.
9. Show the results which is retrieved images corresponding to input or query image according to their rank based on ED.

5.2 Flowchart for Proposed Algorithm

Flowchart of proposed algorithm is given in fig. 2

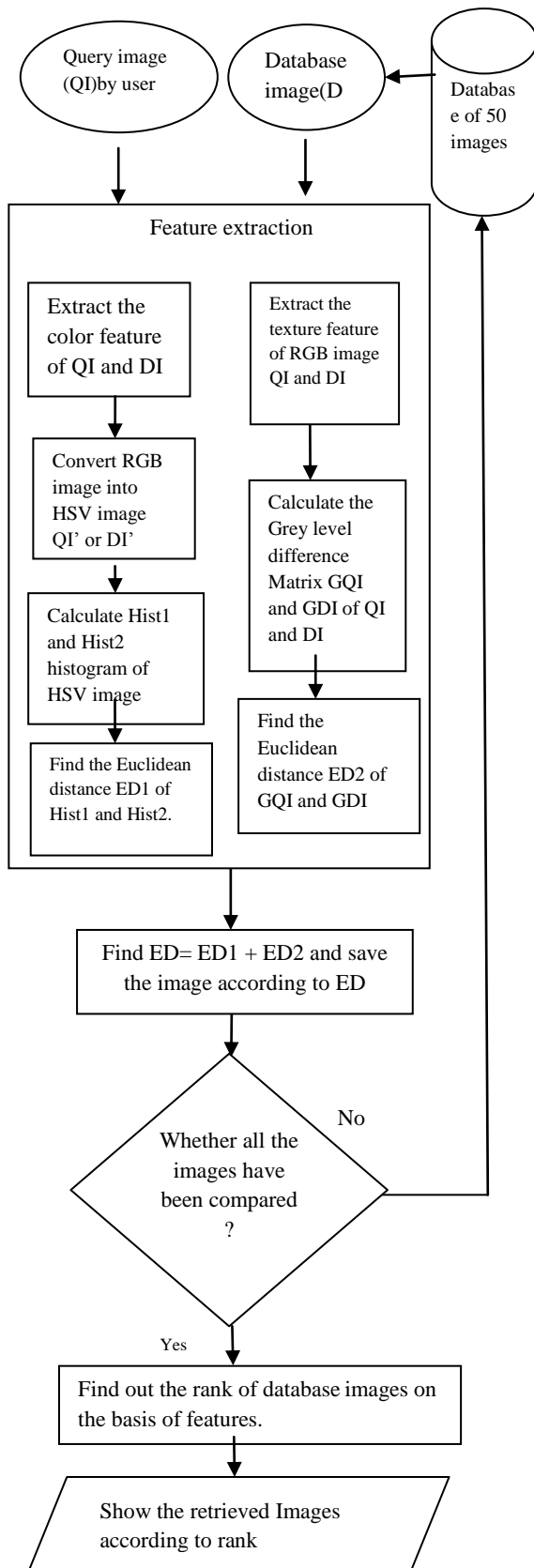





Fig 2: Flowchart of Proposed Algorithm

6. RESULTS

We have implemented the proposed algorithm using MATLAB. The results are shown in TABLE 1.

TABLE 1. Results of implemented algorithm

Input Image	No. of Relevant Images in 30 Retrieved Images	Total no. of Relevant Images Present in the Database
	16	20
	6	10
	9	10

7. CONCLUSIONS

Performance of implementation of the algorithm were tested with respect to the number of images in the database. The time taken for retrieval increases with the increase in the database as for each image in database the color histogram, Euclidean distance and GLD Matrix is calculated. So, the time complexity depends upon the size of database. This time complexity can be reduced using map-reduce framework in future. The result shows that image retrieval using the above methods works and results were satisfactory for fifty images and further results can be improved by introducing shape features or any other feature.

8. REFERENCES

- [1] Yang-Hoon Kim, Hyuk-Jun Kwon, Jong-Gu-Kang and Hangbae Chang, "The study on content based multimedia data retrieval system", Multimedia Tools and Application an International Journal © Springer Science Business Media, LLC 2011.10.1007/s11042-011-0758-5, 1 March 2011.
- [2] Stian Edvardsen, "Classification of Images using Color, CBIR Distance Measures and Genetic Programming", An evolutionary Experiment, Norwegian University of Science and Technology Department of Computer and Information Science, 2006.
- [3] Amandeep Khokher, Rajneesh Talwar, "Content based Image retrieval: Feature Extraction Techniques and Applications, International Conference on Recent

Advances and Future Trends in the Information Technology, IJCA,2012.

- [4] Rafael Gonzalez, Richard E. Woods, Steven L. Eddins,” Fundamentals of digital Image processing”, 2006.
- [5] Wikipedia, the free encyclopedia/texture/texels.
- [6] Krishna Kumar Pandey1& Nishchol Mishra, “design & development of color matching algorithm for image retrieval using histogram and segmentation techniques”, International Journal of Information Technology and Knowledge Management Volume 4, No. 2, July-December 2011.