

# Fundamental of Content Based Image Retrieval, Challenges and Issues

Shaheen Fatima

Assistant Professor of Electronics, Government College, Kalaburagi, Karnataka India

**Abstract:** This paper is to review the present state of the art in content-based image retrieval (CBIR), a technique for retrieving images on the basis of automatically-derived features like color, texture and shape. The CBIR method is used to retrieve relevant images from the database based on the query image submitted by the user. The retrieval of images from a database relies purely on the image features such as color, shape and object identification using texture(s) in the query image. It can be useful to illustrate image queries into three levels of abstraction first is primitive features such as color or shape, second is logical features such as the identity of objects shown and last is abstract attributes such as the significance of the scenes depicted.

**Keywords:** CBIR, Feature Extraction, Distance Measures, Spatial Analysis.

## INTRODUCTION

A typical CBIR system automatically extract visual attributes like color, shape, texture and spatial information of each image in the database based on its pixel values and stores them in to a dissimilar database within the system called feature database [1,2]. The feature data for each of the visual attributes of each image is very much smaller in size compared to the image data. The feature database contains an abstraction of the images in the image database; each image is represented by compact illustration of its contents like color, texture, shape and spatial information. The users generally prepare query image and present to the system. CBIR system can be used in one of two ways. First, exact image matching, that is matching two images, one an example image and the other, image in image database. First, exact image matching, that is matching two images, one an example image and the other, image in image database. Furthermore is approximate image matching, which is finding most closely match images to a query image [3]. CBIR involves the subsequent four parts in system realization [2], data collection, build up feature database, search in the database, arrange the order and deal with the results of the retrieval [4].

- 1) Data gathering
- 2) Extract feature database
- 3) Searching in the Database
- 4) Process and index the results

## FEATURES EXTRACTION:

Feature extraction is the heart of the content based image retrieval. As we know that raw image data that can not used straightly in most computer vision tasks. Mainly two reason behind this first of all, the high dimensionality of the image

makes it hard to use the whole image. Further reason is a lot of the information embedded in the image is redundant. Therefore instead of using the whole image, only an expressive representation of the most significant information should extract. The process of finding the expressive representation is known as feature extraction and the resulting representation is called the feature vector [5].

## COLOR FEATURE:

One of the most significant features of image that make possible the recognition of images by humans is color[6]. Color is a property that depends on the reflection of light to the eye and the processing of that information in the brain. We use color everyday to tell the distinction between objects, places, and the time of day .Images characterized by color features have many advantages. Efficiency, Strength, Simplicity, Low Storage Requirements.

## TEXTURE FEATURE:

In the field of computer vision and image processing there is no exact definition of texture [7,8]. Texture is a main component of human visual perception. Like colour, this also makes it an essential feature to consider when querying image databases. Texture has qualities like periodicity and scale; it can be described in terms of coarseness, direction,contrast. It contains important information related to the structural arrangement of the surface, such as; clouds, leaves, bricks, fabric. It also describes the relationship of the surface to the surrounding environment [7].

## SHAPE FEATURE:

Another major image feature is the shape of the object contained in the image Shape feature of image may be defined as the characteristic surface configuration of an object; an outline or contour. It permits an object to be distinguished from its surroundings by its outline [9,10] Shape representations can be generally divided into two categories: Boundary-based, and Regionbased. Boundary-based shape representation only uses the outer boundary of the shape. This is done by describing the considered region using its external characteristics, like the pixels along the object boundary. But the Region-based shape representation is totally dissimilar from the prior method .It uses the entire shape region by describing the considered region using its internal characteristics; i.e., the pixels contained in that region .

## SPATIAL ANALYSIS:

Even though color and shape are the most important features of image, it is still possible that two images with totally different appearance might have the same color and shape information in the index file [5]. To distinguish two different

images having similar color and shape information, we can apply spatial analysis. The images are partitioned into a number of regions. Then we apply the color and shape analysis in to each small region separately. The color and shape information of all the small regions together form the index of this image. The selection of the number of partitions is critical to the performance of image retrieval.

#### **DISTANCE MEASURES:**

Instead of exact matching, content-based image retrieval calculates visual similarities between a query image and images in a database. Accordingly, the retrieval result is not a single image but a list of images ranked by their similarities with the query image. Many similarity measures has been developed for image retrieval based on empirical estimates of the distribution of features. Different similarity or distance measures will affect retrieval performances of an image retrieval system significantly.

#### **PERFORMANCE MEASUREMENT:**

In the content-based retrieval, because there is variety of algorithm for retrieval, so it needs to estimate the retrieval results mainly using the two indexes, precision and recall [4].

Recall=Number of relevant images retrieved / Number relevant images in collection  
Precision= number of relevant images retrieved /total of images retrieved.  
Recall reflects the system's ability of retrieval related images, while the precision reflects the ability of rejecting the unrelated images.

#### **CHALLENGES AND ISSUES:**

- 1)The semantic gap
- 2)User seeks semantic similarity
- 3)How to represent visual content of images
- 4)How to retrieve images efficiently.

#### **CONCLUSION:**

According to literature reviewed, One of the main challenges is to find significant features that adequately represent an image, especially for the general purpose CBIR applications and the way the user communicates with the content based image retrieval system, the size of the databases , the features used and the speed of the retrieval are the most important factors that judge the success of a CBIR system.

#### **REFERENCES**

- [1] Alaa H. Halawani, Alexandra Teynor, Lokesh Setia, Gerd Brunner,Hans Burkhardt, " Fundamentals and Applicationsof Image Retrieval": An Overview"Datenbank Spektrum 18/2006
- [2] Fahui Long,Hongjiang Zhang and David Dagan Feng," Fundamentals of content-based image retrieval",Microsoft corporation research articles, 2003.
- [3] Zijun Yang and Jay Kuo, "Survey on content-based analysis,indexing and retrieval techniques and status report of MPEG-7",Tamkang journal of science and engineering, vol.2, No.3 pp.101-118, 1999
- [4] Yu Xiaohong,Xu Jinhua " The Related Techniques of Content based Image Retrieval ".,2008 International Symposium on Computer Science and Computational Technology
- [5] J.R.Smith and S.-F Chang, "Integrated spatial and feature image query", Multimedia Syst. 7(2) ,129-140,1999
- [6] Shengjiu Wang, "A Robust CBIR Approach Using Local Color Histograms," Department of Computer Science,University of Alberta, Edmonton, Alberta, Canada, Tech.Rep. TR 01-13, October 2001
- [7] Sharmin Siddique, "A Wavelet Based Technique for Analysis and Classification of Texture Images," Carleton University,Ottawa, Canada, Proj. Rep. 70.593, April 2002.
- [8] Barbeau Jerome, Vignes-Lebbe Regine, and Stamon Georges,"A Signature based on Delaunay Graph and Co-occurrence Matrix," Laboratoire Informatique Systematique, Universiyt of Paris, Paris, France, July 2002.
- [9] Roger Weber and Michael Mlivoncic, "Efficient Region-Based Image Retrieval", ACM CIKM '03 November 3-8,USA,2003
- [10] M.Trimeche, "Shape Representations for Image Indexing and Retrieval", Master of Science Thesis, Tampere University of Technology, May 2000