Volume 5, Issue 4, 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Study of Different Techniques for Image Retrieval

[1]Rajdeep Kaur, [2]Kamaljit Kaur
[1]Student Masters of Technology, [2]Assistant Professor
[1, 2] Department of CSE, Sri Guru Granth Sahib World University
Fatehgarh Sahib, Punjab, India

Abstract*—From last few years, there is rapid increase in the size of digital image collections. Everyday, millions of images are being generated.In this paper,various techniques for image retrieval have been studied.Furthor,brief discussion has been done on content based image retrieval which is included with feature extraction methods based on color, texture and shape.At the end of this paper,comparative study between various content based image retrieval techniques has been done.

Keywords*—Image retrieval, Text based image retrieval (TBIR), Color-Based Image Retrieval (CBIR)

I. INTRODUCTION

Image Retrieval:
An image retrieval is defined as a method of browsing, searching and retrieving images from a large database of digital images. Image retrieval has been a very active research area since the 1970s[1]. Some research about image retrieval technology has begun that focuses only on the text-based ImageRetrieval (TBIR) which utilize some method of adding keywords, or descriptions to the images so that retrieval can be performed over the annotation words. In 1990s,content-based image retrieval (CBIR) has begun. To overcome from limitations of traditional methods, CBIR retrieves similar images based upon image features such as color, texture, and shape from large image database.

1.1 Image Retrieval Architecture: There are three databases in this system architecture. The image collection database contains the raw images for visual display purpose. During different stages of image retrieval, different image resolutions may be needed [1].
The visual feature database stores the visual features extracted from the images using techniques. This is the information needed to support content-based image retrieval.
The text annotation database contains the key words and free-text descriptions of the images.
The retrieval engine module includes a query interface sub-module and a query-processing sub-module. The interface collects the information need from the users and displays back the retrieval results to the users in a meaningful way. The query-processing sub-module manipulates the user query into the best processing procedures [1].
There are two major characteristics of this system architecture:

- Multidiscipline and Inter-discipline nature.
- Interactive nature between human and computer

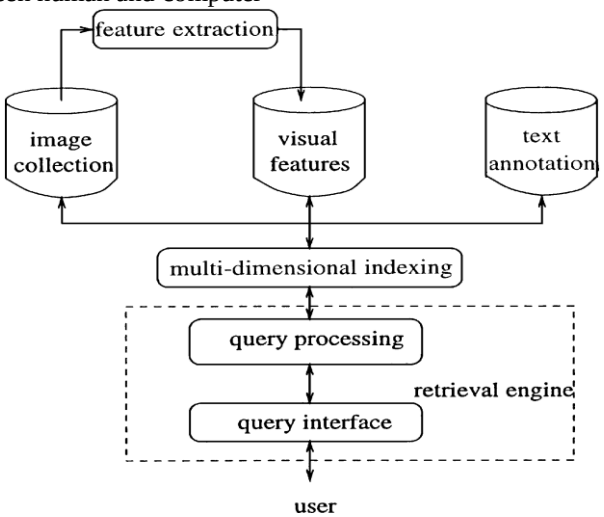

Fig. 1 Image Retrieval System Architecture[1]

The two methods, which are used for image retrieval, are:

- Text based image retrieval
- Content based image retrieval

Table I Comparision Between TBIR and CBIR :

TBIR	CBIR
In text based, user entered the query in the form of text to search an image from image database and the system will return images similar to the query entered by the user. It is also known as annotations based image retrieval (ABIR) **Advantages:** • Easy to implement • Fast retrieval • Web image search **Disadvantages:** • Annotation of each image requires domain experts • It is necessary to use unique keyword for each image,so this is a very complex task. • Annotation for each and every image in a large database is impossible • Sometimes, Text descriptions are incomplete	Content-based means that the search analyses the contents of the image not the metadata such as keywords, labels or tags associated with the images. It is also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) **Advantages:** • Features such as color, texture, shape and spatial are retrieved automatically • Similarities of the images are based on distance between the features • No need of domain experts • Description of image in text form doesn't required **Disadvantages:** • High semantic gap between low level features and high level features

II. CONTENT BASED IMAGE RETRIEVAL

Content-based image retrieval was introduced in 1990's to address the problems associated with text-based image retrieval as discussed above. CBIR is also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR). To search and retrieve digital images, CBIR uses content (color, texture, shape etc.) of the images. Content-based means that the search analyses the contents of the image not the metadata such as keywords, labels or tags associated with the images. In CBIR systems input is provided in terms of an image and based on image attribute matching the most similar images from database are retrieved [2].

CBIR involves with following steps:

- Data collection - Collect the images.
- Extract the features of collected images as well as query image.
- Search the Database - Visual features from the database images are extracted and stored in the feature matrix. Then userenters the query.
- The features of the query image are then matched with the feature matrix using some similarity matching methods.
- After searching, then the resulting images are retrieved.

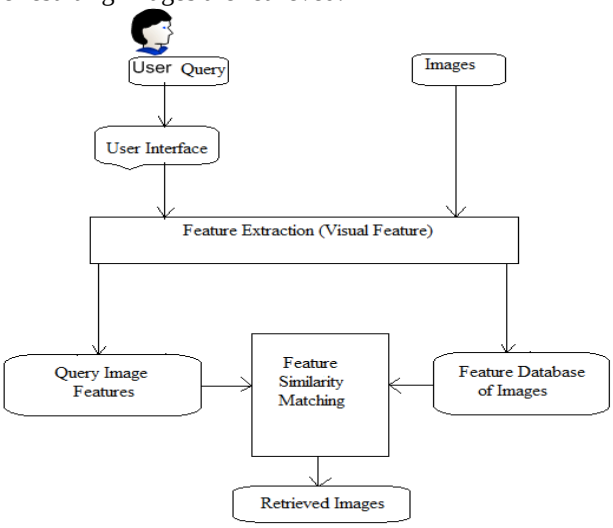

Fig. 2 Flow Diagram of a CBIR system [3]

2.1 Existing Techniques of CBIR:

2.1.1 Feature Extraction: A feature is defined as capturing a certain visual property of an image. It is the main task in the CBIR systems to retrieve the similar images from database similar to query image. In feature extraction, features such as color, texture or shape from image are extracted and creates a feature vector for each image. Several important features that can be used in image retrieval will be discussed in the next subsections.

Color: Color is the most important features that are easily recognized by humans in various images. Color features are the most widely used in CBIR systems. To extract the color features from an image, a color space and color feature extraction method are required [2] Colors in digital images may be represented in a variety of color models including RGB, HSV, YCbCr and CIELab etc [19]. The simplest way to represent colors in an image is to populate color histograms in which a count of the number of pixels of various colors is accumulated. Color quantization is generally employed to reduce the number of colors in the image into a few representative colors. Several color descriptors [19] have also been introduced which represents the importance of a dominant color in the image.

Texture: Texture is the natural property of all surfaces, which describes visual patterns. Like colors in the image, the textural characteristics are also effective ways of describing visual content. Texture features have also been widely used in CBIR applications. This is a feature that describes the distinctive physical composition of a surface [2]. Six texture features including directionality, coarseness, contrast, line-likeliness, regularity and roughness were decided to be more important than the rest after experiments. Gabor features have also been used for texture analysis tasks. The extensive adaptation of various methods for a variety of tasks involving texture analysis is a proof of their strength [19].

Shape: Shape description is an important task in content-based image retrieval. It is important in CBIR because it corresponds to region of interests in images [2]. Extensive research is being carried out in the field of shape based image retrieval. This section focuses on the most commonly used shape descriptors derived from the shape contour or shape interior. A shape descriptor needs to be accurate in retrieving similar shapes from the database. Region based Shape Descriptors are derived from the entire set of pixels that make up an object. Shape descriptors can be classified as global or structural [19] Boundary based Shape Retrieval Descriptors are derived by considering only the boundary of the shape. Shape recognition using shape contexts [19] is an enhancement to the classic Hausdorff distance based methods.

2.1.2 Segmentation:

Segmentation is very important to image retrieval. Segmentation extracts the boundaries from a large number of images without occupying human time and effort. The user defines where the object of interest is, and than the algorithm groups regions into meaningful objects [1]. Reliable segmentation is especially critical for characterizing shapes within images, without this shape estimates are meaningless. The normalized cut segmentation method in [8] is also extended to textured image segmentation by using cues of contour and texture differences.

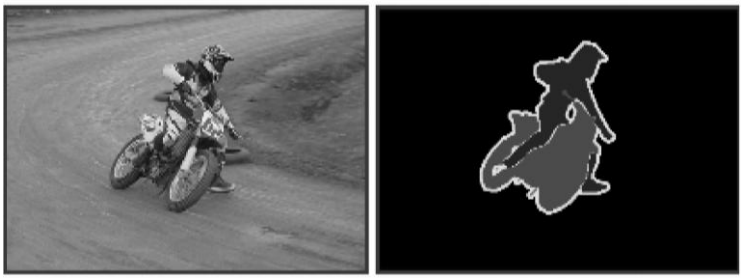

Fig. 3 Segmented image

2.1.3 Image Clustering: Clustering will be more advantage for reducing the searching time of images in the database. Various clustering techniques [8] like Side-information, kernel mapping, k-means, hierarchical, metric learning are used in image retrieval. Performance of K- mean algorithm is better than Hierarchical Clustering Algorithm. In another discussion, Fuzzy k-means is better than k-means by many factors like :It gives better results when compared with k-means algorithm by increasing the fuzzy factor. It takes lesser time to cluster the images than K-means. K-means is considered to be a hard clustering and in hard clustering, after some iteration most of the centers are converged to their final positions where as Fuzzy K-means is known as soft clustering in which the data points, which are present in the fuzzy K-means, can belong to more than one cluster with having certain probability [21]. Density based methods typically consider exclusive clusters only, and do not consider fuzzy clusters. Fuzzy C-means (FCM) is one of the clustering methods, which allow one piece of data to belong to two or more clusters. FCM groups data in specific number of clusters [14].

III. EXISITING WORK

Pattanaik, Bhalke (2012)[18] has worked to prove that Content Based Image Retrieval has overcome all the limitation of Text Based Image Retrieval by considering the contents or feature of image. From the experimental result it is seen that combined features can give better performance than the single feature.

Chin-Chin Lai et.al. (2011)[6] Proposed an interactive genetic algorithm (IGA) to reduce the gap between the retrieval results and the users expectation. They have used color attributes like the mean value, standard deviation, and image bitmap. They have also used texture features like the entropy based on the gray level co-occurrence matrix and the edge histogram. They compared this methods with others approaches and achieved better results.

Meenakshi Madugunki et.al. (2011)[10]Describe detailed classification of CBIR Systems. They have used the Global color histogram, Local Color histogram, HSV method for extracting the color feature and matched the result by using Euclidean distance, Canbera distance and city block

J´erˆome Da Rugna, Gael Chareyron and Hubert Konik (2011)[16] discuss the segmentation step in the feature extraction chain. This step frequently used in CBIR systems. The goal in this paper is to propose an objective evaluation of the stability of classical image segmentation process.

Meenakshi shruti pal and Dr.Sushil kumar garg (2013)[17] paper provide a comprehensive review and characterize the various problems of image retrieval techniques. They present a survey of the most popular image retrieval techniques with their pros and cons.

A.Kannan et.al. (2010)[11] Proposed Clustering and Image Mining Technique for fast retrieval of images. The main objective of the image mining is to remove the data loss and extracting the meaningful information to the human expected needs. The images are clustered based on RGB Components, Texture values and Fuzzy C mean algorithm.

IV. COMPARISION OF DIFFERENT CBIR TECHNIQUES:

Techniques Used	Description	Advantages	Limitations
Color Histogram Wavelets and Gabor Filter	Two methods were implemented on image database. Color and Texture taken as features and chi-square and euclidean distances as matching criteria. For the performance measure, precision ratio have taken	• Effective multi-scale image analysis • Lower computational cost • Feature vector is less incombined approach of wavelet and color moments	Poor retrieval efficiency
HSV Color Moment Hough Transform Ranklet Texture Moment	Combination of color, texture and shape feature is used to compare and retrieve image is more accurately than using one of them only. Euclidean Distance is used to measure the similarity between two images with N-dimensional feature vector.	• Higher precision than the combination of two features • Dimensions of features vector are low • Lower computational complexity	High computational time
Color Moment Gabor Filter GVF(Gradient Vector Flow Fields)	In this method, an image is partitioned into non-overlapping tiles. It captures the local colour and texture descriptors in a segmentation framework of grids and shape describable in terms of invariant moments	• Create robust feature set • High retrieval efficiency	High semantic gap Time consuming
Color Moment Gabor Wavelet Co-occurrence Matrix	In this combination is done in two levels. One is the combination of color and texture features and the other is the combination of two textures extracted by two different methods.	• Minimize the semantic gap using RF with SVM	Not suitable for large and heterogeneous Image database
CBIR Using Genetic K-means Algorithm (GKA)	In this GA hybridize with K-means algorithm, which define K-means operator, one of the step of K-means algorithm, and use it in GKA as a search operator instead of crossover.	• Searches faster • Retrieval performance is good	Consumes long time for computation
CBIR Using C-mean Clustering	It computes the distance between the centroid of the cluster and seed point. It also uses prior identified number K		

	(number of clusters to be formed) therefore the results depends on the cluster number K and initial choices of seed points.	• Minimizes intra cluster variance	High computational time
Color moment Block Truncation Coding (BTC) Algorithm	Block Truncation Coding (BTC) used to extract features of images and K-Means clustering algorithm is used to group the image dataset into various clusters	• Performance is superior to that of color moments	

V. CONCLUSION

After comparing various techniques of content-based image retrieval, it is conclude that there is no such technique that can achieve the accuracy of image retrieval system. So, in future including some advance techniques, which will achieve the accuracy of image retrieval, can extend this comparative study.

REFERENCES

[1] Yong Rui and Thomas S. Huang, Shih-Fu Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues", Journal of Visual Communication and Image Representation, 10, 39–62,1999

[2] Anuradha Shitole1 and Uma Godase, "Survey on Content Based Image Retrieval", International Journal of Computer-Aided Technologies, Volume 1,April 2014

[3] Suman Lata, Parul Preet Singh,"A Review on Content Based Image Retrieval System ",International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 5,May 2014

[4] M. Venkat Dass, Mohammed Rahmath Ali, Mohammed Mahmood Ali,"Image Retrieval Using Interactive Genetic Algorithm", International Conference on Computational Science and Computational Intelligence, 2014

[5] Xiang-Yang Wang, Hong-Ying Yang, Dong-Ming Li "A new content-based image retrieval technique using color and texture information", Computers & Electrical Engineering, Volume 39,April 2013

[6] Chih-Chin Lai, and Ying-Chuan Chen, "A User-Oriented Image Retrieval System Based on interactive Genetic Algorithm," IEEE transactions on instrumentation and measurement, Volume 60,October 2011

[7] P. S. Hiremath, Jagadeesh Pujari, "Content Based Image Retrieval using Color, Texture and Shape features", International Conference on Advanced Computing and Communications, 2007

[8] Datta R, Joshi D, Li J, Wang J Z,"Image retrieval: ideas, influences, and trends of the new age", ACM Computing Surveys, Volume 40, No. 2,april 2008

[9] Pranali Prakash Lokhande, P. A. Tijare, "Feature Extraction Approach for Content Based Image Retrieval", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 2**,** February 2012

[10] Dr. D.S.Bormane, Meenakshi Madugunki, Sonali Bhadoria, Dr. C. G. Dethe,"Comparison of Different CBIR Techniques", IEEE Conference, 2011

[11] A.Kannan, Dr.V.Mohan, Dr.N.Anbazhagan,"Image Clustering and Retrieval using Image Mining Techniques", IEEE Conference, 2010

[12] Priyanka P. Buch, Madhuri V. Vaghasia, Sahista M. Machchhar,"Comparative analysis of content based image retrieval using both color and texture", International Conference On Current Trends In Technology, December, 2011

[13] A.Kannan, Dr.V.Mohan and Dr.N.Anbazhagan,"An Effective Method of Image Retrieval using Image Mining Techniques", The International journal of Multimedia & Its Applications (IJMA), Volume 2,November 2010

[14] Source: Dr. Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng,"Fundamentals Of Content-Based Image Retrival"

[15] J. Assfalg, A. D. Bimbo, and P. Pala, "Using multiple examples for content-based retrieval", Proc.Int'l Conf. Multimedia and Expo, 2000

[16] J´erˆome Da Rugna, Gael Chareyron and Hubert Konik,"About Segmentation Step in Content-based Image Retrieval Systems", Proceedings of the World Congress on Engineering and Computer Science, Volume I, October 2011

[17] Meenakshi Shriuti Pal, Dr Sushil kumar Garg,"Image retrieval: A Literature Review", international journals of advanced research in computer engineering and technology (IJARCET), Volume 2,Issue 6, June 2013

[18] Swapnalini Pattanaik, Prof.D.G.Bhalke, "Beginners to Content Based Image Retrieval", International Journal of Scientific Research Engineering &Technology (IJSRET), Volume 1 Issue2, May 2012

[19] Jamil Ahmad, Muhammad Sajjad, Irfan Mehmood, Seungmin Rho, Sung Wook Baik,"Describing Colors, Textures and Shapes for Content Based Image Retrieval – A Survey", Journal Of Platform Technology Volume 2, December 2014

[20] A. Malcom Marshall, Dr. S. Gunasekaran,"Image Retrieval-A Review", International Journal of Engineering Research & Technology (IJERT),Volume 3, Issue 5,May 2014

[21] Md. Khalid Imam Rahmani, Naina Pal, Kamiya Arora, "Clustering of Image Data Using K-Means and Fuzzy K-Means", International Journal of Advanced Computer Science and Applications (IJACSA), Volume 5, 2014

