ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

# Review on: Content Based Image Retrieval

[1] Manisha Sharma [2]Anand Singh Bisen [3]Abhinav Vidwans
[1]Head. Department of Computer Science, Department of Computer Science, VITM, Gwalior
[2][3]M.Tech Scholar, Department of Computer Science, VITM, Gwalior

***Abstract:*** **The paper presents a review of different techniques in content-based image retrieval. Content Based image retrieval is a system by which several images are retrieved from a large database collection. The paper starts with discussing the fundamental aspects of CBIR. Features for Image Retrieval like color, texture and shape are discussed next. We briefly discuss the similarity measures based on which matches are made and images are retrieved. Another essential problem in content-based image retrieval is effective indexing and fast searching of images based on visual features. Dimension reduction and indexing schemes are also discussed. For content-based image retrieval, user interaction with the retrieval system is crucial since flexible formation and modification of queries can only be obtained by involving the user in the retrieval procedure. Finally Relevance feedback is discussed which helps in improving the performance of a CBIR system.**

***Index Terms*— Image retrieval, Content Based Image retrieval (CBIR)**

## I. INTRODUCTION

Due to exponential increase of size of so called multimedia files in recent years because of the substantial increase of affordable memory storage on one hand and the wide spread of World Wide Web (www) on the other hand, the essential for the efficient tool to retrieve the images from the large data base becomes crucial. This motivates the extensive research into image retrieval systems. From the historical perspective, the earlier image retrieval systems are rather text-based with the thrust from database management community since the images are required to be annotated and indexed accordingly [1]. However with the substantial increase of the size of images as well as size of image database, the task of user-based annotation becomes very cumbersome and at some extent subjective and thereby, incomplete as the text often fails to convey the rich structure of images. In the early 1990s, to overcome these difficulties this motivates the research into what is referred as content based image retrieval (CBIR) where retrieval is based on the automating matching of feature of query image with that of image database through some image-image similarity evaluation. Therefore images will be indexed according to their own visual content such as color, texture, shape.

## II. OVERVIEW: CONTENT-BASED IMAGE RETRIEVAL

Content Based Image Retrieval (CBIR) is any technology that in principle helps to organize digital image archives by their visual content. By this definition, anything ranging from an image similarity function to a robust image annotation engine falls under the purview of CBIR The most common form of CBIR is an image search based on visual [1] .The increasing amount of digitally produced images requires new methods to archive and access this data. Conventional databases allow for textual searches on Meta data only. Content Based Image Retrieval (CBIR) is a technique which uses visual contents, normally called as features, to search images from large scale image databases according to users' requests in the form of a query image. Apart from the usual features like color and texture, a new feature extraction algorithm called edge histogram is introduced. Edges convey essential information to a picture and therefore can be applied to image retrieval. The edge histogram descriptor captures the spatial distribution of edges. This model expects the input as Query by Example (QBE) and any combination of features can be selected for retrieval. The focus is to build a universal CBIR system using low level features. These are mean, median, and standard deviation of Red, Green, and Blue channels of color histograms. Then the texture features such as contrast, energy, correlation, and homogeneity are retrieved. Finally the edge features that include five categories vertical, horizontal, 45 degree diagonal, 135 degree diagonal, and isotropic are added [2]. Human being gets images, sound and any other information by seeing, hearing and perception and analysis. Human judge similarity of images and sounds according to their semantic contents, for instance the searching for a star's picture is based on his facial characters

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

or other contents. So the retrieval methods based on text or keywords for the digital multimedia apparently can't meet the demand that human being get multimedia information exactly. With more and more multimedia information appear on the Internet and other digital multimedia as well as human beings' thirst for exact and fast retrieval, based on contents multimedia information retrieval becoming the focus of the academe research as well as images retrieval of contents is one of the important study aspects of multimedia information retrieval [3]. Existing color-based general-purpose image retrieval systems roughly fall into three categories depending on the signature extraction approach used: histogram, color layout, and region-based search. And, histogram-based search methods are investigated in two different color spaces. A color space is defined as a model for representing color in terms of intensity values. Typically, a color space defines a one- to four- dimensional space. A color component, or a color channel, is one of the dimensions. Color spaces are related to each other by mathematical formulas. The two three-dimensional color spaces, RGB and HSV, are investigated. CBIR involves the following four parts in system realization: data collection, build up feature database, search in the database, arrange the order and deal with the results of the retrieval.

1) Data collection Using the Internet spider program that can collect webs automatically to interview Internet and do the collection of the images on the web site, then it will go over all the other webs through the URL, repeating this process and collecting all the images it has reviewed into the server.

2) Build up feature database using index system program do analysis for the collected images and extract the feature information. Currently, the features that use widely involve low level features such as color, texture and so on, the middle level features such as shape etc.

3) Search the Database The system extracts the feature of image that waits for search when user input the image sample that need search, then the search engine will search the suited feature from the database and calculate the similar distance, then find several related webs and images with the minimum similar distance.

4) Process and index the results after researching Index the image obtained from searching due to the similarity of features, then return the retrieval images to the user and let the user select. If the user is not satisfied with the searching result, he can re-retrieval the image again, and searches database again. The retrieval of content based image involves the following systems

***A. Color-based retrieval***

Color feature is the most intuitive and obvious feature of the image, and generally adopt histograms to describe it. Color histograms method has the advantages of speediness, low demand of memory space and not sensitive with the images' changes of the size and rotation, it wins extensive attention consequently.

***B. The retrieval based on texture feature***

When it refers to the description of the image's texture, we usually adopt texture's statistic feature and structure feature as well as the features that based on special domain is changed into frequency domain.

***C. The retrieval based on shape feature***

There is three problems need to be solved during the image retrieval that based on shape feature. Firstly, shape usually related to the specifically object in the image, so shape's semantic feature is stronger than texture [4,9 ].

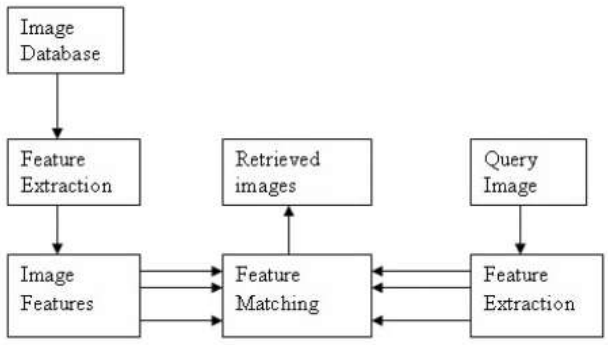


***Fig. 1: A typical Content Based Image Retrieval System***

## III. TYPES OF CBIR BASED IMAGE RETRIEVAL

***1) Region-based:*** The Netra and Blobworld are two earlier region based image retrieval systems [4]. During retrieval, a user is provided with segmented regions of the query image, and is required to assign several properties, such as the regions to be matched, the features of the regions, and even the weights of different features.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

***2) Object-based:*** Object-based image retrieval systems retrieve images from a database based on the appearance of physical objects in those images. These objects can be elephants, stop signs, helicopters, buildings, faces, or any other object that the user wishes to find. One common way to search for objects in images is to first segment the image in the database and then compares each segmented region against a region in some query image presented by the user. Such image retrieval systems are generally successful for objects that can be easily separated from the background and that have distinctive colors or textures [5].

***3) Example-based:*** Users give a sample image, or portion of an image, that the system uses as a base for the search. The system then finds images that are similar to the base image.

***4) Feedback-based:*** System shows user a sample of pictures and asks for rating from the user. Using these ratings, system re-queries and repeats until the right image is found.

## IV. CBIR TECHNIQUES

Different implementations of CBIR make use of different types of user queries.

***CBIR system using QBE***: Query-by-example or pictorial-query approaches make the system return similar images to the example image given by a user. The underlying search algorithms may vary depending on the application, but result images should all share common elements with the provided example.
Options for providing example images to the system include:

- A pre-existing image may be supplied by the user or chosen from a random set.
- The user draws a rough approximation of the image they are looking for, for example with blobs of color or general shapes.

To begin a search, the user has an example image to submit as a query. The example images can be a photograph, user-painted example, or line- drawing sketch. The query serves as an approximation of the objective image being sought. The CBIR system accesses the image in the database, matches the query against the information in the database, and scores the images in terms of similarity. . In this method, images are retrieved by their contents: color, texture, shape, or objects. The matching is based on chromatic and textures features with equal weights. Thus, the degree of similarity between query images and images in databases can be measured by color distribution, texture distribution, shape similarity, or object presence between the two images. The top *k*-best images are returned as results. Upon receiving the result, user evaluates if the images in the result are relevant and selects another image from the result or database to refine the query. This query technique removes the difficulties that can arise when trying to describe images with words.

***Semantic retrieval:*** The ideal CBIR system from a user perspective would involve what is referred to as semantic retrieval, where the user makes a request like "find pictures of dogs" or even "find pictures of Abraham Lincoln". This type of open-ended task is very difficult for computers to perform - pictures of chihuahuas and Great Danes look very different, and Lincoln may not always be facing the camera or in the same pose. Current CBIR systems therefore generally make use of lower-level features like texture, color, and shape, although some systems take advantage of very common higher-level features like faces. Not every CBIR system is generic. Some systems are designed for a specific domain, e.g. shape matching can be used for finding parts inside a CAD-CAM database.

***Other query methods***: Other query methods include browsing for example images, navigating customized/hierarchical categories, querying by image region (rather than the entire image), querying by multiple example images, querying by visual sketch, querying by direct specification of image features, and multimodal queries (e.g. combining touch, voice, etc.) .
CBIR systems can also make use of relevance feedback, where the user progressively refines the search results by marking images in the results as "relevant", "not relevant", or "neutral" to the search query, then repeating the search with the new information [6, 7].

## IV. APPLICATIONS

Examples of CBIR applications are [8]:

- **Crime prevention:** Automatic face recognition systems, used by police forces.
- **Security Check:** Finger print or retina scanning for access privileges

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

- **Medical Diagnosis:** Using CBIR in a medical database of medical images to aid diagnosis by identifying similar past cases.
- **Intellectual Property:** Trademark image registration, where a new candidate mark is compared with existing marks to ensure no risk of confusing property ownership.
- **Architectural and engineering design**
- **Photograph archives**
- **Commerce** (fashion, catalogue)
- **Cultural** (art galleries, museums)
- **Military** (radar, aerial )
- **Entertainment** (personal album)

## VI CONCLUSION

Content Based Image Retrieval is an active and fast advancing research area since the 1990s. During the past decade, remarkable progress has been made in both theoretical research and system development. The impetus behind content-based image retrieval is given by the wide availability of digital sensors, the Internet, and the falling price of storage devices. Given the magnitude of these driving forces, it is to us that content-based retrieval will continue to grow in every direction: new audiences, new purposes, and new styles of use, new modes of interaction, larger data sets, and new methods to solve the problems. A wide researches of have been made on image retrieval. Each work has its own techniques, contribution and limitations. As a review paper, it might not include each and every aspects of individual work, however this paper attempts to deal with a detailed review of the most common, traditional and modern content based retrieval

## REFERENCES

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

[2] N. S. Chang and K.S.Fu, "image query by pictorial – example," IEEE Trans Software

[3] V.N.Gudivada and J.V. Raghvan, "Special issues on content based image retrieval system," IEEE Com. Magzine, 1985

[4] J. Assfalg, A. D. Bimbo, and P. Pala, "Using multiple examples

[5] for content-based retrieval," Proc. Int'l Conf. Multimedia and Expo, 2000.

[6] http://www.igi-global.com/dictionary/content-based-image-retrieval-cbir/5587

[7]http://www.ijsret.org/pdf/beginners_to_content_based_image_retrieval.pdf

[8] Meenakshi Shruti Pal and Dr. Sushil Kumar Garg, ―International Journal of Advanced Research in Computer Engineering and Technology (IJARCET) Volume 2, Issue 6, June 2013.

[9] H. Yamamoto, H. Iwasa, N. Yokoya, and H. Takemura, ―Content-Based Similarity Retrieval of Images Based on Spatial Color Distributions‖, ICIAP '99 Proceedings of the 10th International Conference on Image Analysis and Processing.