iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** VI **Month of publication:** June 2017

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

A Review of Content Based Image Retrieval System

Manpreet Kaur[1], Suchitra[2]
[1] M. Tech, Student Dept. Of CSE, GGGI Dinarpur, Ambala, Haryana
[2]Asst. Prof., Dept. Of CSE, GGGI Dinarpur, Ambala, Haryana

Abstract- ***content based image retrieval (cbir), also called as query by image content (qbic). It has been an active research field since last decades. In contrast to traditional systems, where images are retrieved on the basis of keywords but in the cbir system images are retrieved on the basis of visual content. Content based image retrieval is an approach that enables a user to extract an image based on a query, from the database containing huge amount of images. Here, the term "content" refer to color, shape, texture, or any other information that can be derived from image itself. Cbirs is becoming a necessity for various applications such as medical imaging, geographic information systems (gis), space search and many others. In this paper, we have discussed various techniques that enable users to extract images from huge database, thus improving accuracy, efficiency and precision. The main issue in designing a content based image retrieval system is to select the image features that best represent the image contents in a database.***

Keywords: cbirs, image databases, color string comparison, feature extraction, query image, target image.

I. INTRODUCTION

Since last few decades, systems working with retrieving large amount of multimedia data have been growing rapidly. Systems like search engines, e-business systems, online tutoring system, GIS, and image archive are few to them. These systems involve retrieving multimedia data based on visual content. In image archive for example, a simple query such as searching for bird with green feathers requires system to be able to find all images in the database which contains a bird with green feathers. This is a challenging task since it requires system to browse every single image in database and compare it to query image. Manual browsing database to search for identical images would be impractical since it takes a lot of time and requires human efforts as well. A more practical way is to use Content based image retrieval (CBIR) technology. CBIR has provided an automated way to retrieve images based on content or features of images itself. The CBIR system simply extracts content of the query image matches them to contents of the search image.One of the most interesting aspects of this or finding alternative way of looking and overcoming the limitations forced by TBIR systems more natural and easy to understand content based image retrieval systems (CBIR) were developed. Content Based Image Retrieval or CBIR is defined as a process to search similar images from the database when a query image is given. Given an image of an apple, the system should be able to present all similar images of an apple in the database to user. This is done by extracting features of the images such as color, texture and shape. These image features are used to compare between the query image and images that are present in the database. A similarity algorithm like Cosine, Spearman, Jaccard index etc. is used to calculate the degree of similarity between those two images. Images in the database which has similar features to the query image is then ranked and presented to the user.

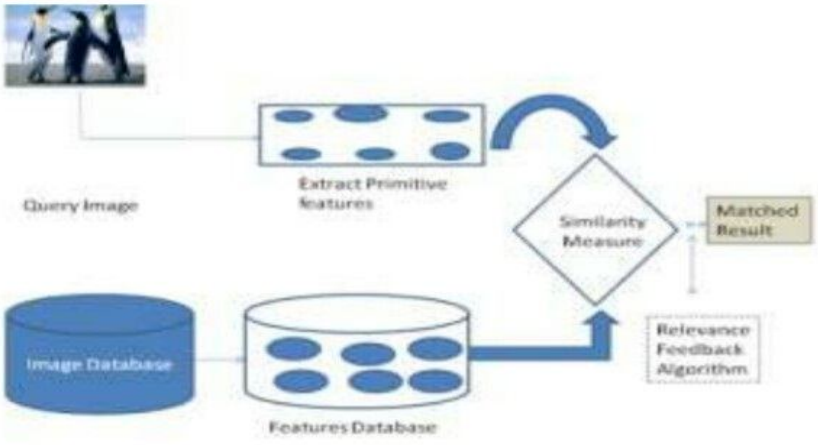
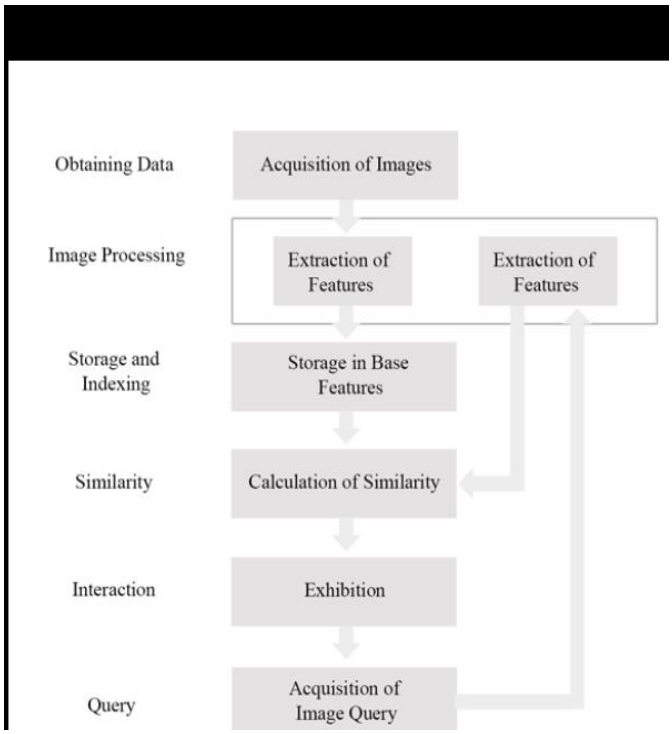


Fig. 1 Architecture of a typical CBIR system

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

The feature vectors of images in database form a feature database. The similarity measure is utilized to calculate the distance between the feature vectors of query image and those of the objective images in the feature database.



: Fig. 2 Basic Steps of CBIR system.

The implementation of such a system requires the extraction and storing of the image features to be compared with the features of the query image. With this flow, the implementation process is more dynamic, since all features have already been stored somewhere.

Since then, the image database techniques have become popular area for researchers. CBIR is most popular area of research; it has been introduced from 1990's. Today, in various fields such as medical, engineering designs, crime prevention, fashion, interior designs, and education etc. The content based image retrieval plays an important role for searching, browsing and retrieving of images. The earlier searching and retrieving processes of images were done by traditional text based retrieval system. To overcome these problems content based image retrieval is developed where image are retrieved based on their visual contents, they are as:

Low level features (human vision related).

Middle level features (object related).

High level features (semantic related).

Among them, the low level features which are most trendy for its simplicity when compared to features of other levels. A low level feature is classified as color, texture and shape. The Content based image retrieval involves two phases

A. Feature Extraction

It is the heart of the content based image retrieval. As we know that raw image data that cannot 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.

B. Matching

It performs comparison between extracted features of images to check whether they are similar or not and up to what extent they match.

C. colour

One of the most significant features of image that make possible the recognition of images by humans is color [9]. 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.

II. APPLICATIONS OF CBIRS

The CBIR technology has been used in several applications such as fingerprint identification, biodiversity information systems, digital libraries, crime prevention, medicine, historical research, among others. Some of these applications are presented in this section.

A. Biodiversity Information Systems

Biologists gather many kinds of data for biodiversity studies, including spatial data, and images of living beings. Ideally, Biodiversity Information Systems (BIS) should help researchers to enhance or complete their knowledge and understanding about species and their habitats by combining textual, image content-based, and geographical queries. A combination of this query with textual and spatial predicates would consist of "Show the drainages where the fish species with 'large eyes' coexists with fish whose fins are shaped like those of the fish in the photo". Examples of initiatives in this area include [55, 74].

B. Digital Libraries

There are several digital libraries that support services based on image content. One example is the digital museum of butterflies aimed at building a digital collection of Taiwanese butterflies. This digital library includes a module responsible for content-based image retrieval based on color, texture, and patterns. The system manages air photos which can be retrieved through texture descriptors. Place names associated with retrieved images can be displayed by crossreferencing with a Geographical Name Information System (GNIS) gazetter. In this same domain, Bergman *et al.* describe an architecture for storage and retrieval of satellite images and video data from a collection of heterogeneous archives. For example, while research presented in [77,78] concentrates on new searching strategies for improving the effectiveness of CBIR systems, another popular focus is on proposing image descriptors [79].

III. LITERATURE SURVEY

There are various methods have been proposed to extract the images from very large database. Here, we have some of the papers that use different techniques to retrieve the images:

A. Ammar Huneiti et al. [1]

proposed a CBIR method by extracting both color and texture feature vectors using the Discrete Wavelet Transform (DWT) and the Self Organizing Map (SOM) artificial neural networks. At query time texture vectors are compared using a similarity measure which is the Euclidean distance and the most similar image is retrieved. In addition, other relevant images are also retrieved using the neighbourhood of the most similar image from the clustered data set via SOM. Results showed that the proposed method is able to retrieve images with higher average precision values than other methods proposed in literature by just comparing the texture similarity and without any need to compare color similarities.

B. Anuja khodaskar et al. [2]

proposed an advanced content based image retrieval system using topical rule based classification strategy which improve retrieval performance significantly. The proposed classification strategy used three training rules, low level, high level and expert rules which improve classification accuracy and effectiveness, ultimately encroachment in quality of classification. Experimental result shows performance evolution in precision, accuracy and retrieval time of image retrieval.

C. Bulo et al.,[3]

proposed a novel approach to content-based image retrieval with relevance feedback, which is based on the random walker algorithm introduced in the context of interactive image segmentation. The idea is to treat the relevant and non-relevant images labelled by the user at every feedback round as "seed" nodes for the random walker problem. The ranking score for each unlabeled image is computed as the probability that a random walker starting from that image will reach a relevant seed before encountering a non-relevant one.

D. Garvita and Priyanka Camboj [4]

Content Based Image Retrieval (CBIR) is also called as Query By Image Content (QBIC). It has been an active research field since last decades. In contrast with traditional systems, where images are retrieved on the basis of keywords but in CBIR system images are retrieved on the basis of visual content. Content Based Image Retrieval is a methodology that allows a user to extract an image based on query from the database. Different types of classification we can use Neural network, Support Vector Machine (SVM), KNN, Bayesian etc.

We are using K Nearest Neighbour (KNN) classifier to find out the relevant images and after that we use Spearman's Rank Correlation Function to reduce the time complexity and improve F-measure From above discussion we conclude that the proposed approach perform better than existing approach with reduced time complexity and improves F-measure value. As here we have own Dataset which contain limited number of images. In future, we can use this concept for the huge database as well and use some other classifiers to enhance the results.

E. Gaurav Jaswal, Amit Kaul [5]

concluded that content based image retrieval is not a replacement of to text based image retrieval. But integration of the two can result in satisfactory retrieval performance. Both author reviewed the main components of a content based image retrieval system. Authors also analyzed image feature representation, and indexing . Also highlighted the past and current technical achievement.

F. Ivan Lee, et.al. [6]

have present the analysis of the CBIR system with the human controlled and the machine controlled relevance feedback. They proposed an analytical study over different network topologies including centralized and distributed content search. The experimental results for the interactive relevance feedback using RBF. They observed a higher retrieval precision by introducing the semi-supervision to the non-linear Gaussian shaped RBF relevance feedback.

G. Jain and Singh [7]

provided an overview of the functionality of content based image retrieval systems by combining advantages of HC and divide and conquer K-Means strategy. He proposed HDK method to use both advantages of HC and Divide and conquer K-Means by introducing equivalency and compatible relation concepts.

H. Kalyan Roy and Joydeep Mukherjee [8]

Image Retrieval means searching, browsing, and retrieving the images from an image database. Two different methods are used for image retrieval, namely text based image retrieval and content based image retrieval techniques.

In Content Based Image Retrieval many visual feature like color, shape, and texture are extracted, next when we query an image its feature are compared with the stored feature and we get most similar kind of image. In our proposed method we firstly extract low level image feature like- color histogram, color coherence vector.

Then we add edge detection technique sobel edge detection method to get better output. Finally we use Manhattan distance to find the similar images from our database. Content Based Image Retrieval, means extracting a range of images which is relevant with the given image from a large database of images.

I. Kommineni Jenni et al. [9]

presented a Content Based Image Retrieval approach based on the database classification using Support Vector Machine (SVM) and color string coding feature selection. In SVM method, the feature extraction was done based on the basis of color string coding and string comparison. Here, they succeed in transferring the images retrieval problem to strings comparison. Thus the computational complexity is decreases obviously and increased the accuracy in obtaining results for image retrieval. Using database classification we can improve the performance of the content based image retrieval than compared with normal CBIR that is without database classification.

J. Lionel Gueguen and Mihai Datcu [10]

addressed the problem of extracting relevant information from Satellite image time series (SITS) based on the information-bottleneck principle. The method depends on suitable model selection, coupled with a rate–distortion analysis for determining the

optimal number of clusters. They presented how to use this method with the Gauss–Markov random fields and the auto binomial random fields model families in order to characterize the spatio-temporal structures contained in SITS. In this approach spectral or geometrical information was not taken into account.

IV. CONCLUSION AND FUTURE SCOPE

Nowadays, CBIR has been an active research field and is used in various fields like Medical, Geographic Information System (GIS), and Space Research etc. In this paper, we have discussed various techniques that enable users to extract the relevant image from huge database improving the accuracy, precision and reducing retrieval time.

In future work, we will work on feature extraction and classification to get the better results for image retrieval.

REFERENCES

[1] Ammar Huneiti, Maisa Daoud, "Content-Based Image Retrieval Using SOM and DWT", Journal of Software Engineering and Applications, 2015, 8, 51-61 Published Online February 2015

[2] Anuja khodaskar, "Advanced Image Retrieval with Topical Classification Strategy", International Conference on Intelligent Computing, Communication & Convergence (ICCC-2014) Conference Organized by Inter science Institute of Management and Technology, Bhubaneswar, Odisha, India

[3] Bulo, S.R., M. Rabbi and M. Pelillo, "Content based image retrieval with relevance feedback using random walks", Patt. Recogn., 44: 2109- 2122, 2011.

[4] Garvita and Priyanka Kamboj, "Content Based Image Retrieval System: Review", International Journal of Computer Trends and Technology (IJCTT) – Volume 34 Number 3 - April 2016

[5] Gaurav Jaswal Amit Kaul, " Content Based Image Retrieval ", National Conference on Computing, Communication and Control , An Literature Review , National Institute of Technology, Hamirpur- 177001, Himachal Pradesh(India).

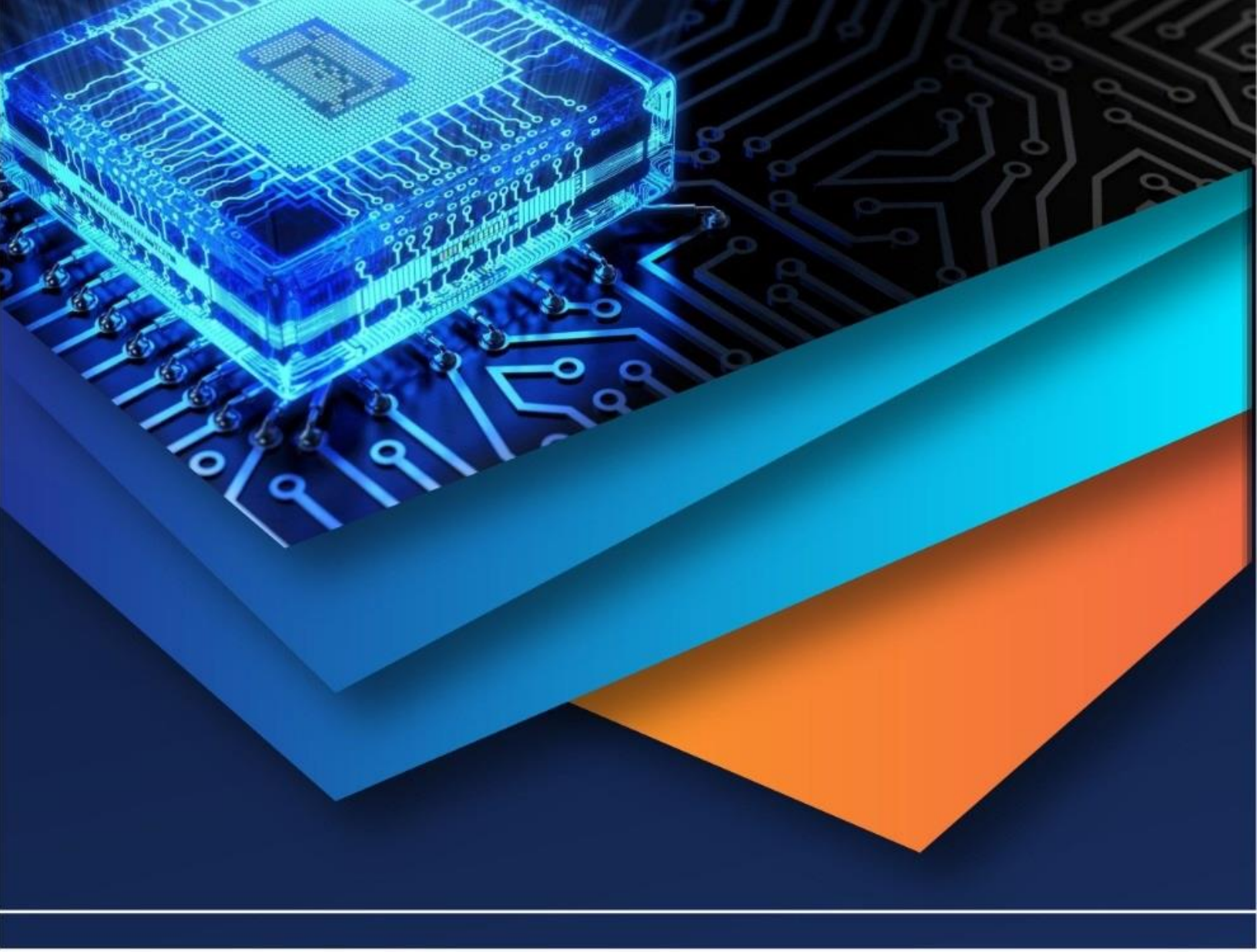
[6] Ivan Lee, Paisarn Muneesawang, Ling Guan, "Automatic Relevance Feedback for Distributed Content-Based Image Retrieval",ICGST, ieee.org FLEXChip Signal Processor (MC68175/D), Motorola, 1996.

[7] Jain, M. and S.K. Singh, "A survey on: Content based image retrieval systems using clustering techniques for large data sets", Int. J. Manage. Inform. Technol., 3: 23-39, 2011.

[8] Kalyan Roy and Joydeep Mukherjee, "Image Similarity Measure using Color Histogram, Color Coherence Vector, and Sobel Method", International Journal of Science and Research (IJSR), Volume 2 Issue 1, January 2013.

[9] Kommineni Jenni, Satria Mandala, Mohd Shahrizal Sunar, Content Based Image Retrieval Using Colour Strings Comparison, 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

[10] Lionel Gueguen and Mihai Datcu, " Image Time-Series Data Mining Based on the Information-Bottleneck Principle", IEEE Transactions On Geoscience and Remote Sensing, Vol. 45, no. 4,827, April 2007.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY