

A Review of Content Based Image Retrieval System

Manpreet Kaur¹

Dept. Of CSE, GGGI Dinarpur, Ambala, Haryana

ABSTRACT- *An image retrieval system provides a way to access, browse and retrieve images efficiently, from the large databases. 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.*

Introduction

With the headway in internet and multimedia technologies, a immense amount of multimedia data in the form of audio, video and images has been utilized as a part of numerous fields like medical treatment, satellite data, video and still images repositories and surveillance system. This has made a progressing interest of frameworks that can store and retrieve mixed media information in a powerful way. Numerous multimedia information storage and retrieval systems have been developed till now to cater these requests. The most common retrieval systems are Text Based Image Retrieval (TBIR) systems, where search is based on automatic or manual explanation of images. A conventional TBIR searches

database for the similar text surrounding the image as given in the query string. The TBIR systems are fast as the string matching is computationally less time consuming process. In addition annotation of images is not always correct and takes a lot of time. Image Processing is a field that continues to grow, with new applications being developed at an ever increasing pace. It is fascinating and an exciting area to be involved in today with application area ranging from the entertainment industry to the space program. 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. The system retrieves comparable images when an query image or sketch is presented as input to the system. Querying in this way removes need for describing the visual content of images in words and is near to human perception of visual data. A portion of representative CBIR systems is Query by Image Content (QBIC).

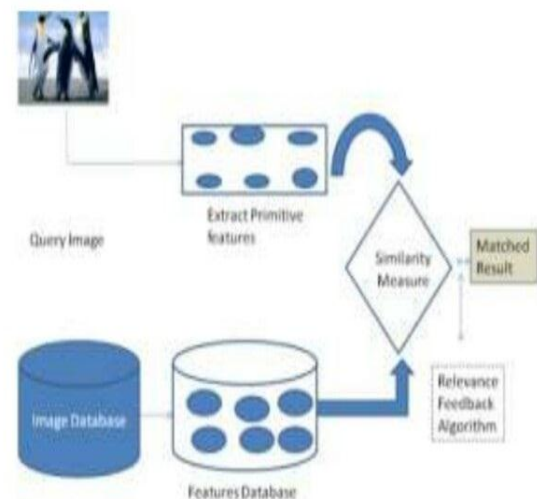


Fig. 1 Architecture of a typical CBIR system

The feature vectors of images in database form a feature database. The retrieval process is initiated when a user query the system using an image or sketch of the object. 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. At last, the retrieval is performed using an indexing scheme which facilitates an efficient searching of the image database. A CBIR system is composed of a query interface for the acquisition of the query image, databases for storing indexing data and metrics, similarity and retrieval system. Figure 1 illustrates these schema:

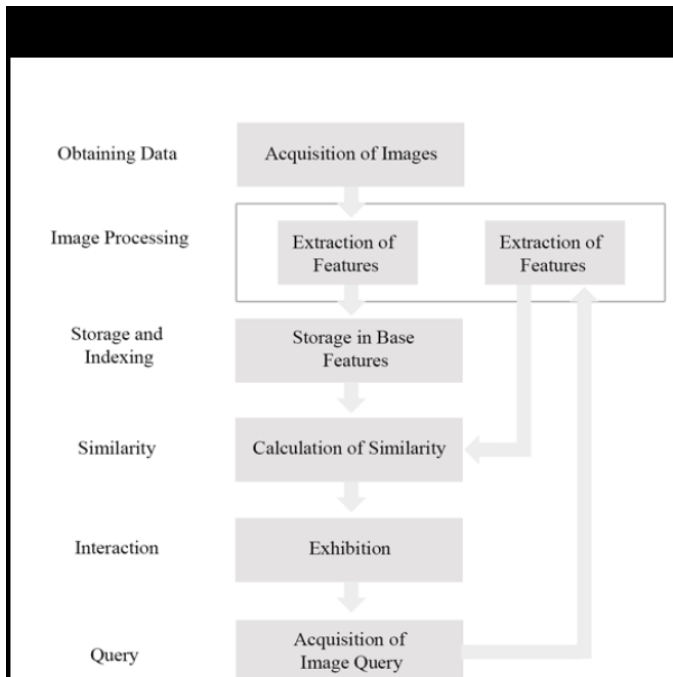


Fig. 2 Basic Steps of CBIR system

It may be noted that CBIR is a process that may require long processing time due to the amount of images to be analyzed in a database, so comparison between images is made using a set of features extracted from the images [6]. 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. An image retrieval system provides a way to access, browse and retrieve images efficiently, from the large databases. Since then, the image database techniques have become popular area for researchers. Retrieval of image from large databases are exponentially increasing and becoming a challenging task. For image retrieval, two methods are used namely Content Based Image Retrieval System (CBIR) and Text Based Image Retrieval System (TBIR). Out of the two, 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. But many problems links to this traditional text based retrieval systems such as manual explanation of keywords, discrepancy in perceptions. 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

Feature Extraction

It is a process to extract the features in terms of information of an image based on their visual contents for comparing images.

Matching

It performs comparison between extracted features of images to check whether they are similar or not and up to what extent they match. It uses distance functions like Euclidean distance, Mean Square Error for measurement of similarity.

Color

One of the important features that make possible the recognition of images by humans is color. Color is a property that depends on the reflection of light to the eye and the processing of that information by the brain. We use color to tell the difference between objects, places, and the time of day.

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.

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].

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:

Manish K. Shriwas et al. [1] proposed a new technique for CBIR called as Local tetra patterns (LTrPs). In this technique images are retrieving from database based on the direction of pixel that is calculated by vertical and horizontal derivatives. By combining LTrPs with Gabor transform (GT) the effectiveness of proposed method will also be analyzed. By considering the diagonal pixel for calculation the derivation of vertical and horizontal direction the result of proposed method will be improved. Accuracy and efficiency of proposed approach will reaches a level of 93% and it can also be suitable for various applications such as biomedical research, educational, crime prevention, web images etc. Sushant Shrikant Hiwale et al. [2] proposed a CBrn system which extracts the features of digital image to retrieve similar images from huge databases. We have used Color Histogram, Color Auto-Correlogram, Color Moment, Gabor Wavelet and Discrete Wavelet transform to extract image features. The images are classified using Support Vector Machine (SVM) classifier which effectively distinguishes between relevant and irrelevant images. The results depict that proposed method has better precision and recall rate compared to other methods.

A. A. Khodaskar et al. [3] proposed innovative framework for effective, intelligent and efficient content based image retrieval is based on three soft computing techniques such as Artificial Neural Network, Fuzzy Logic and support vector machine. In this framework, the SVM based relevance feedback is introduced, that will provide feedback to the system that whether the retrieved image is relevant or not. The

main purpose is to avoid non relevant images and to reduce the un-necessary overhead of the system. Content based image retrieval based on soft computing techniques like ANN, fuzzy logic. Devyani Soni et al. [4] The use of Image data is growing tremendously in every field such as medical, engineering designs, fashion, interior designs, and education etc. In this paper the researcher presents a method for image retrieval based on text, color space approaches with color histogram and color correlogram. Color correlogram is a second order statistical approach which calculate spatial relationship. In this approach Euclidean distance is used for comparing query image with database images. During experimentation, both HSV color model as well as RGB color model is used for the same process of retrieval. The author have developed this mechanism for image retrieval based on color features of image with the help of MATLAB tool. Mrs. K. Jayanthi [5] An investigation comparative of many descriptors of various images in content-based image retrieval system (CBIR) is described in the paper. This paper describes more number of various features in CBIR system and compare the four different Color and texture based existing low level Feature Extraction The Proposed Techniques such as Fuzzy color and texture histogram (FCTH) and Color and edge directivity descriptor (CEDD) which retrieve the relevant images matching with edge, texture and color value from the Corel library. The proposed approach gives good result in Average Image Retrieval precision (IRP) and Recall value. An experimental feature extraction techniques comparison shows that that the performance of FCTH and CEDD techniques is good in image retrieval. Hanen Karamti [6] Visual information retrieval has become a major research area due to increasing rate at which images are generated in many application. This paper addresses an important problems related to the content-based images retrieval. It concerns the vector representation of images and its proper use in image retrieval. Indeed, we propose a new model of content-based image retrieval allowing to integrate theories of neural network on a vector space model, where each low level query can be transformed into a score vector. Preliminary results obtained show that our proposed model is effective in a comparative study on two dataset Corel and Caltech-UCSD.T. Dharani [7] An image is

an array, or a matrix, of square pixels organized in columns and rows. A picture is worth a thousand words. Image metrics are very useful to analyzing the raw images for further quality image processing. In this paper we are mainly focusing image metrics for better performance of CBIR system. Because Content Based Image Retrieval system mainly working with diverse images for retrieval purpose. Finally Image Quality Assessment (IQA) has been carried out using various evaluation measures are found better in the retrieval accuracy. In this paper, an image metrics are proposed. The proposed research works are very useful to found better performance and fast retrieval purpose in CBIR system. Nicoleta Angelescu[8] This paper presents a SQL query optimization method for Content Based Image Retrieval (CBIR) systems. The proposed approach is tested on infantile hemangioma images. Details about the textural features and implementation are provided and their effectiveness is proved. A SQL query optimization method for CBIR systems is presented. Modern web technologies for the server and client parts are used. The proposed approach is fast and has several advantages when applied on infantile hemangioma images. The main advantage is the improvement of running time for the users' queries. Nhat Quang Doan[9] Over the past decade, Content-Based Image Retrieval (CBIR) has been an active research approach to retrieve images based on their visual contents such as color, texture and shape. Image indexing is an important task to ensure the efficiency of CBIR systems. In this paper, we present a new stochastic clustering algorithm which allows automatic image indexing and quick neighbor search of similar images. In this paper, we introduce a hybrid algorithm based on dynamic tree structure for both image clustering and indexing. The proposed structure combines hierarchical and topological aspects to associate image feature vectors. It supports queries based on self-assembly similarity functions. Xiangyuan Zhao[10] In this paper, we propose an algorithm for a Content Based Image Retrieval (CBIR) system based on Wavelet Transformation and Deep Auto encoder (DAE). For the proposed algorithm, the image is first processed by wavelet transform and decomposed into wavelet coefficients. In this paper, a content based image retrieval system is presented using both wavelet transformation and multiple

input and multiple task deep auto encoder. Experiment shows that compared with a deep auto encoder that has only single input and single output and uses unprocessed image as the input, our new system can get better performance with a smaller model and fewer trainable parameters.. Khairul Munadi[11] This paper presents a content-based image retrieval mechanism for visually protected image databases. Visual protection is provided as a means to prevent illegal use of a database by those who could break access control, such as an untrustworthy server provider or intruders. The proposed scheme directly extracts a DC image from the lowest frequency DCT coefficients of a visually protected image and measures its similarity with a thumbnail of arbitrary size that is queried by a user. In this way, unscrambling of the image database is not necessary. The proposed method is compatible with the JPEG compression standard and the MPEG-7 Color Layout Descriptor (CLD).

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] Manish K. Shriwas, Prof. V. R. Raut, Content Based Image Retrieval: A Past, Present and New Feature Descriptor, IEEE International Conference on Circuit, Power and Computing Technologies [ICCPCT] 2015.
- [2] Sushant Shrikant Hiwale, Dhanraj Dhotre, Dr. G.R. Bamnote, Quick Interactive Image Search in Huge Databases Using Content-Based Image Retrieval, IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems ICIIECS'15.
- [3] A. A. Khodaskar and S. A. Ladhake, A Novel Approach for Content Based Image Retrieval in Context of Combination S C Techniques, International Conference on Computer Communication and Informatics (ICCCI -2015), Jan. 08 – 10,

2015, Coimbatore, INDIA

[4] Devyani Soni , K. J. Mathai, An Efficient Content Based Image Retrieval System based on Color Space Approach Using Color Histogram and Color Correlogram, Fifth International Conference on Communication Systems and Network Technologies 2015

[5] K. Jayanthi, M .Karthikeyan, “ Color and Edge Directive Descriptor Feature Extraction Technique for Content Based Image Retrieval System” Middle-East Journal of Scientific Research 23 (8):1590- 1597, 2015.

[6] Hanen Karamti and Mohamed Tmar and Anis BenAmmar, A New Relevance Feedback Approach for Multimedia Retrieval. IKE, 219–225 (2012).

[7] T. Dharani, I. Laurence Aroquiaraj, “CONTENT BASED IMAGE RETRIEVAL SYSTEM with MODIFIED KNN ALGORITHM”, International Journal of Innovations in Scientific and Engineering Research(IJISER), Vol. 1 Issue 1 Jan 2014.

[8] N. Angelescu, D. Popescu, and M.G. Ionita, "CBIR system based on texture characterization," 4th International Symposium on Electrical and Electronics Engineering (ISEEE 2013), pp. 1-6,2013.

[9] N.-Q. Doan, H. Azzag, and M. Lebbah, “Growing self-organizing trees for knowledge discovery from data.” in IJCNN. IEEE, 2012, pp. 1–8.

[10] Zhou, Xiang Sean, Sonja Zillner, Manuel Moeller, Michael Sintek, Yiqiang Zhan, Arun Krishnan, and Alok Gupta. "Semantics and CBIR: a medical imaging perspective." In Proceedings of the 2008 international conference on Content-based image and video retrieval, pp. 571-580. ACM, 2008.

[11] K. Munadi, M. Syaryadhi, F. Arnia, M. Fujiyoshi, and H. Kiya, “Secure online image trading scheme using dct coefficients and moment invariants feature,” in Consumer Electronics (ISCE), 2013 IEEE 17th International Symposium on, June 2013, pp. 291–292.