ISSN No. 0976-5697

Volume 5, No. 7, September-October 2014

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Study of Different Methods of Content Based Image Retrieval (Cbir)

Rupesh Lahori
PG, Scholar, CSE
RITS, BHOPAL

Anurag Jain
Head, CSE
RITS, BHOPAL

Himanshu Yadav
Assistant Professor, CSE
RITS, BHOPAL

Abstract: Content-based image retrieval (CBIR) is widely adopted process for searching and retrieving images from large image databases. Now the day's volume of production of digital images is increasing day by day, traditional retrieval techniques are not satisfactory and not time efficient. CBIR is a fast retrieval method. Researchers developed so many techniques for CBIR. In this paper, we present a detailed survey of these CBIR techniques with their pros and cons. In addition, we discussed issues regarding image retrieval and evaluation of retrieval performance. Finally, on the basis of present CBIR methods we suggest some future research scope which can help researchers in their work.

Keywords: CBIR, Color histograms, Database, Image processing.

I. INTRODUCTION

Content based image retrieval (CBIR) has been one amongst the foremost vital techniques for the image retrieval in computer vision. It's used for searching images in massive image databases [1]. Content based image retrieval (CBIR) or query by image content (QBIR) and content based visual information retrieval (CBVIR) is that the technique to the image retrieval drawback. The main drawback of searching for digital images within the huge databases [2] the amount of strategies already existed for image retrieval in computer vision. The primary image retrieval system developed by IBM's query by image content (QBIR). Then the foremost efficient use of an image retrieval system is content based image retrieval (CBIR). It's searching the image on the premise of content or image features (like color, texture and shape, etc.). The term content means that color, shape, texture or the other data that may be derived from the image. CBIR could be a helpful technique as a result of most the web based image search engines believe metadata and this produces plenty of garbage within the result. All image retrieval strategies needed to enter key words manually, that isn't efficient, expensive and not lined all the main points of the image. So a system that may filter images based on their content is needed to produce higher categorization and returns a lot of correct results.

The number of applications including military, industrial and civilian generates and will continue to generate more gigabyte of color images per day. Therefore, a large amount of information cannot be accessed unless it is organized. Here organization means that appropriate indexing is available in order to allow client browsing, searching and retrieving as in keywords searches of text databases. Query by example is the simplest way to search which means that the user has to present an image to the system and the letter searches for a leak by extracting features from the query image and comparing them to the ones stored in the database. The extraction of meaning features is critical in CBIR and therefore an open and active field of research area.

There are three categories of image retrieval [3] as follows: (a) text based (b) content based (c) semantic based.

The difficulties faced by text based retrieval became more severe. The well-organized management of the quickly expanding visual information became problematic, so content based image retrieval focuses mainly on the visual information management systems.

The rest of this paper is organized as follows. Section 2 presents the image processing. Section 3 presents the content based image retrieval. Section 4 presents the application of CBIRs. Section 5 describes the various previous methods. Section 6 presents current content based image retrieval methods. Section 7 presents the various image databases. Section 8 presents the literature survey. Section 9 presents the future research scope. Section 10 presents the conclusion. Finally section 11 presents the references.

II. IMAGE PROCESSING

Image processing is a method which converts image into digital form and performs some action on it. The image processing extracts some functional information from large image databases. Image processing is rapidly growing technologies today with its large number of applications for business and computer field. Image processing technique use mainly three steps for image extractions like:-

a. Fetch the image by digital photography or with optical scanners.
b. Manipulating and analyzing the image.
c. An altered image is output of the last stage in image processing.

Digital image processing is concerned with extracting useful information from the image. This is done by

computers with little or no human intervention. Digital image processing have number of steps: -The first step is image acquisition i.e. to acquire a digital image. After a digital image has been fetched, it will preprocess that image. The next step deals with image segmentation [5].Image segmentation partitions an input image into its constituent parts and objects. The term CBIR has been used to describe the process of retrieving image from large image databases on the basis of features.

There are two types of visual content descriptor the global and the local. A global descriptor employs the visual features of the entire image where as local descriptor employs the visual features of area objects to depict the image content. In local visual descriptor is an image divided into part first. The simplest way of dividing an image is to use partition which cuts the image into tiles of equal shape & size. A straightforward partition does not produce perceptually meaningful regions, but is a way of representing the global features of the image at a better resolution. An improved method is to split the image into homogenous regions according to same criterion using a region segmentation algorithm that have been already proposed in computer vision. A more complicated way of separating an image is to undertake a complete object segmentation to obtain semantically meaningful objects (like a ball, car, and hours) [4].

III. CONTENT BASED IMAGE RETRIEVAL

Content based image retrieval (CBIR) another name is queried by image content (QBIR) and content based visual information retrieval (CBVIR) is the application of computer vision technique to the image retrieval difficulty, i.e. the problem of exploring for digital images in large image databases. Therefore, two methods are used in image retrieval system: - Text-Based Image Retrieval (TBIR) and Content-Based Image Retrieval (CBIR). Text-Based Image Retrieval means retrieving the image one of the base text, e.g. human types the text on the Google, yahoo etc. sear engine not accurate result found because not any text or keywords or tag captured exactly. So new method found this is content based image retrieval in the method we are retrieving image on the bases of content from large image databases. The content means color, shape, texture; these are primary image features [2]. The Content Based Image Retrieval is the technique which retrieving desired images from a large database on the basis of image features (that is color, texture and shape etc.).

The main aim of content based image retrieval technique is transform query image or sketch with query resultant images obtained from the image databases, as shown in figure 1.

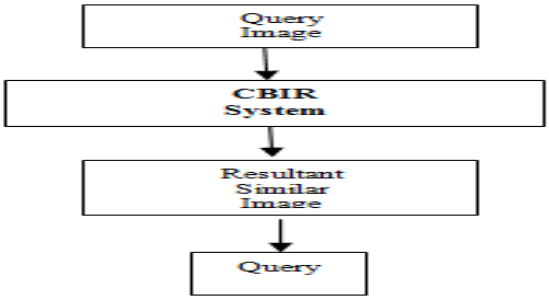

Figure 1:- Block diagram of a CBIRs

The system consist two types of databases such as image database and feature database. The image database is a collection of image features on which image search. While adding a new image in the image database, image features are extracted and stored in feature database. The features are represented in a convenient format before storing them for faster search. The features of the query image are extracted and compared features of the images, available in feature database. The similarity comparison and search are carried out with the image database for finding similar featured image .The corresponding images are retrieved and displayed as a result on the basis of similarity measure, as shown in figure 2.

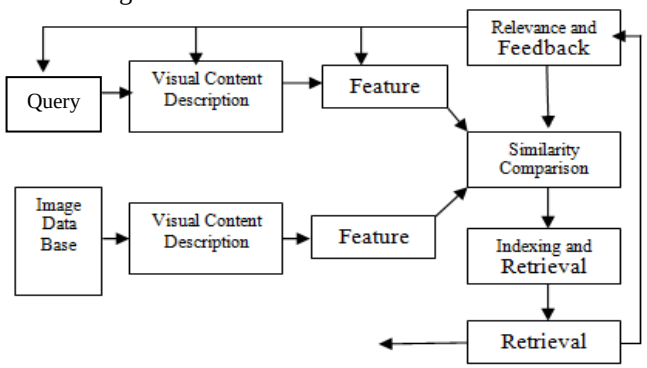

Figure 2:- Describe block diagram of a CBIRs

A. *Image Features:-*

The number of image features is used for content description of images as follows:-

a. Colors, Color Layouts and color distributions
b. Shapes
c. Textures
d. Boundaries and regions
e. Edges
f. Contours
g. Histograms (Local, Global, Cumulative).

The performance of any image retrieval methods in CBIR mainly dependent only two factors:-precision and recall. Precision is defined as the probability that a retrieved image is relevant. Recall is defined as the probability that a relevant image is retrieved in a search [5].

Precision = (Number of relevant image retrieved)/ (Total number of image retrieved)

Recall= (Number of relevant image retrieved)/ (Total number of relevant image in database) [5].

IV. APPLICATION OF CBIR

a. Consumer Digital Photo Albums
a) Digital Cameras
b) Flicker
b. Medical image, Crime Prevention
c. Trademarks search
d. The Military
e. Fashion and interior design
f. GIS and remote sensing
g. Education and training
h. Home entertainment
i. Web searches
j. Architectural and engineering drawing
k. Fingerprint and face recognition
l. Weather forecasting.

V. PREVIOUS METHODS

a) Edge Histogram Descriptor (EHD): - The edge histogram descriptor (EHD) captures the special distribution of local edges and is an efficient descriptor for images with a significant textural presence even once the underlying texture isn't same. The computation of this descriptor is straightforward and generates a histogram of the most edge field contains truth object boundaries and isn't saturated by the background texture. Within the EHD is used not solely to represent the special distribution of local edges in natural images, however also to shape features; it's very sensitive to object or edge distortions. In this technique the retrieval performance could also be unacceptable [10].

b) Color Auto Correlograms (CAC):-This method proposed by Huang et al. This technique planned by Huang et al. the utilization of color correlograms for image classification and retrieval employing a method of the probability. A subset of color correlograms, namely, color auto correlograms, is used in image retrieval because of the high-dimensional vector of color correlograms. Consistent with Huang's color auto correlograms are supported color division within the RGB color space and cause a 256-dimensional vector with four distance sets. The use of color auto correlograms captures the spatial correlation of colors in an image; therefore overcoming the foremost disadvantages of the classical histogram technique, and this rule out performs both ancient histogram technique and the color coherent vector method. Though, it ought to finest hassled that color auto correlograms accomplish superb performance in uniform color space, like the HSV, L*a*b* or LUV color spaces, as a result of color auto correlograms return from four distance sets, they'll end in very high dimensional vectors. Color provides powerful info and is merely one image attribute, and edge orientation is additionally a really vital factor in image representation that's neglected once using color auto correlograms [10].

c) Multi Texton Histogram (MTH): - The Multi Texton histogram is additionally novel image feature representation technique for image retrieval. The multi Texton histogram presents, descriptor with none image segmentation or learning method. The MTH is improved method of TCM. It specially designed for image retrieval and may succeed higher retrieval exactness than that of EOAC and TCM. It's conjointly a bonus of the co-occurrence matrix. The MTH plays a vital role for color and edge orientation while not human perceptually uniform color difference. Thus MTH technique is 82 dimensional vectors as similar to edge histogram descriptor (EHD) and color auto correlograms (CAC) strategies that are 240 and 256 respectively. It's efficiently used technique for image retrieval during a massive image database [3].

d) Color Difference Histogram (CDH):-This methodology is analogous to the human sensory system that that's very sensitive to color and edge orientation. The CDH presents novel image feature for image retrieval. It works as a visible attribute descriptor. This descriptor combines the utilization of orientation, color and color, abstraction layout while not using any image segmentation or learning processes. The authors planned method that initial describe the L*a*b* color space, then second it describe the edge orientation detection using L*a*b* color spaces. Third, it describes the color division in L*a*b* color space, and then finally we describe perceptually uniform color distinction [10].

VI. CURRENT CONTENT BASED IMAGE RETRIEVAL SYSTEMS(CBIRs)

The number of existing software used for retrieving image based on the properties of the image. Two categories of software have been found: open source software products and close source software products. Simplicity (also developed as ALIPR) is one of the online tool developed by Li and Wang at Penn State. CIRES (content based image retrieval system) are another online tool developed by Qasin Iqbal at the University of Texas [6].

The other various CBIR systems as follows:-

a. QBIC: - Query By Image Content System developed by IBM.

b. Ultimedia Manager Product: - Developed by IBM based on QBIC.

c. VisualSEEk: - Developed at Columbia University.

d. Photobook:-Developed at Media Laboratory, Massachusetts Institute of Technology (MIT).

e. FIRE: - Flexible Image Retrieval Engine.

VII. VARIOUS IMAGE DATABASES

The image database is a collection of the various content kind of image. A variety of image databases are used in image processing. In this proposed method they define mainly two kinds of image subset: First subset is the Corel-5K which contains 50 categories covering 5000 images. Each category contains 100 images of size 192*128 or 128*192 in JPEG format. The second subsets are the Corel-10K which contains 100 categories covering 10,000 images. Each category contains 100 images of size 192*128 or 128*192 in JPEG format [10].

VIII. LITERATURE SURVEY

a. Sang Jeong, "Histogram-Based Color Image Retrieval", illustrated that images retrieved by exploitation the universal color histogram might not be semantically associated although they share similar color distribution in some results. An image retrieval presentation system was considered to form it easy to check the retrieval performance and to expedite any formula examination. And six histograms-based image retrieval approaches in two color areas were thoroughly compared by providing accuracy vs. recall graphs for every image category and for all check images. In general, histogram-based retrievals in HSV color area showed higher performance than in RGB color area. The histogram Intersection-based image retrieval in HSV color area was found to be most fascinating among six retrieval strategies.

b. D. Koubaroulis, et.al, "Color-based Image Retrieval from Video Sequences", demonstrated the potential of the Multimodal Neighborhood Signature (MNS) technique for image and video retrieval. Typical

region-based queries were created from a range of frames from a sports video sequence of the Olympic Games and retrieval results were reportable. The algorithmic rule performed well and relevant images were with success retrieved notwithstanding background muddle, partial occlusion and/or non-rigid deformation. Specially, terribly tiny regions were with success matched just like the tiny Irish flag on the swimmer's cap. MNS signatures were computed in period (0.1. Sec) on a Sun radical Enterprise 450 with quad CPUs at 400 megahertz and search speed was 600 image matches per second. Additionally, signature size was typically little (average 900 bytes) that, combined with quick signature computation and retrieval, looks promising for exigent web-based retrieval applications. Though the MNS technique supports looking out with illumination invariant options and use of abstraction data for retrieval (e.g. For query localization), these options weren't tested during this work. Future enhancements to the algorithmic rule embrace introducing a training/learning stage to with efficiency exploit discriminate color characteristics inherent to the database at hand and a multi scale approach to make amends for scale changes [7].

c. Hiremath. P. S, "Content Based Image Retrieval based on Color, Texture and Shape features using Image and its complement", illustrated a unique structure for merge all the three i.e. color, texture and shape data, and accomplish higher retrieval efficiency using image and its complement. The image and its complement are divided into non-overlapping tiles of equal size. The features drawn from conditional co-occurrence histograms between the image tiles and corresponding complement tiles, in RGB color space, function local descriptors of color and texture. This locality data is captured for two resolutions and two grid layouts that give totally different details of an equivalent image. An integrated matching scheme supported most similar highest priority (MSHP) principle and also the adjacency matrix of a bipartite graph shaped using the tiles of query and target image, is provided for matching the images. Shape data is captured in terms of edge pictures computed using Gradient Vector Flow fields. Invariant moments are then wont to record the shape features. The mixture of the color and texture features between the image and its complement in conjunction with the shape features give a strong feature set for image retrieval. The experimental results demonstrate the efficacy of the strategy.

d. Kerminen. P, "IMAGE RETRIEVAL BASED ON COLOR MATCHING", states properties of an image retrieval system supported color matching. The enforced procedure is given as a simplified pseudo code presentation of the system. When histogram has been calculated and saved on the disk, it are often used for examine [8].

e. K. Konstantinidis, "Image retrieval based on fuzzy color histogram processing", proposed a new fuzzy linking technique of color histogram creation is planned supported the L*a*b* color space and provides a histogram that contains solely 10 bins. The histogram creation methodology in hand was assessed supported the performances achieved in retrieving similar images from a wide numerous image assortment. The experimental results prove that the planned technique is less sensitive to numerous changes within the images (such as lighting variations, occlusions and noise) than alternative strategies of histogram creation [9].

IX. RESEARCH SCOPE

The content based image retrieval is the technique for efficiently retrieve the image from large image database. Color is one of the important features of image so here we will propose on the color based image retrieval. The algorithm which would be proposed in the subsequent work will be based on the text and content based, which also incorporated the histogram of the image, since histogram is also one of the keys in the imaging field, so it will search the image not only on the basis of either content or text, but it will search the images in the database by three factors those will be text, color, and histogram of the queried image. We propose a new descriptor for image retrieval. This descriptor combines the use of orientation, color difference and considers the spatially out without the use of any image segmentation or learning processes.

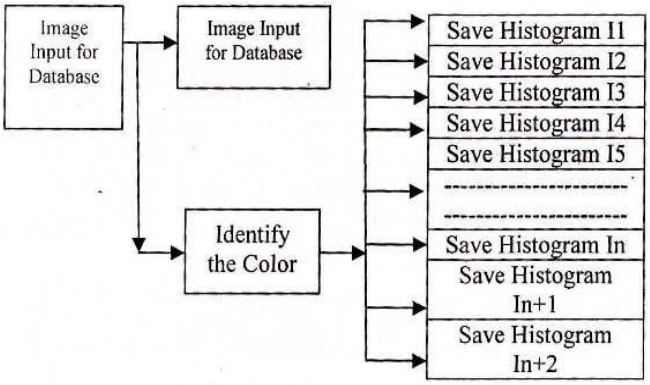

Figure: 3

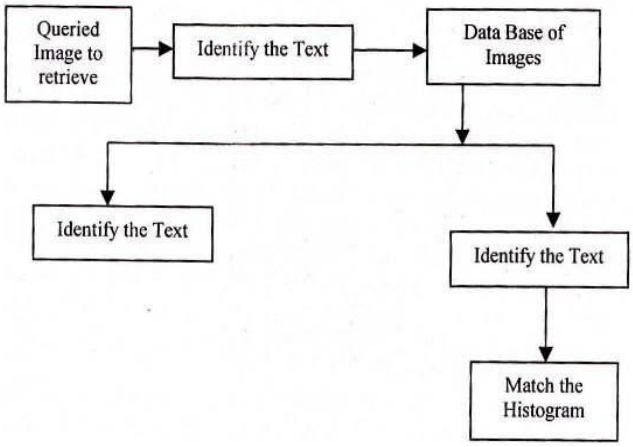

Figure:4

X. CONCLUSION

The content based image retrieval means retrieve the images from large image database on the basis of content or image features (i.e. Color, texture, shape, etc.). The color is the primary content for image retrieval. Then we will work

on the color based image retrieval. Image processing is a technique which is defined as a convert an image into digital form and we perform some operation on it. The number of existing methods used for image retrieval. But these are only calculating frequency or number of pixels. In the MPEG-7 edge histogram descriptor, color, auto correlograms and Multi Texton Histogram methods proposed, but these are not efficiently use so we will propose new descriptor for image retrieval. This descriptor combines the use of orientation, color and color difference as well as spatial layout without the use of any image segmentation or learning methods.

XI. REFERENCES

[1] Taileb, Mounira. "Content based image retrieval system using NOHIS-tree." In *Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia*, pp. 103-108. ACM, 2012.

[2] Kumar, R. Senthil, and M. Senthilmurugan. "Content-Based Image Retrieval System in Medical Applications." In *International Journal of Engineering Research and Technology*, vol. 2, no. 3 (March-2013). ESRSA Publications, 2013.

[3] Liu, Guang-Hai, Lei Zhang, Ying-Kun Hou, Zuo-Yong Li, and Jing-Yu Yang. "Image retrieval based on multi-text on histogram." *Pattern Recognition* 43, no. 7 (2010): 2380-2389.

[4] Khan, Sumaira Muhammad Hayat, Ayyaz Hussain, and Imad Fakhri Taha Alshaikhli. "Comparative Study on Content-Based Image Retrieval (CBIR)." In Advanced *Computer Science Applications and Technologies (ACSAT), 2012 International Conference on*, pp. 61-66. IEEE, 2012.

[5] Bankapur, Rahul, B. V. Bhoomaraddi, Meena S. Maralappanavar, and B. V. Bhoomaraddi. "Optimized trace transform based content based image retrieval algorithm." In *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*, pp. 690-696. ACM, 2012.

[6] Chakravarti, Rishav, and Xiannong Meng. "A Study of Color Histogram Based Image Retrieval." In *ITNG*, pp. 1323-1328. 2009.

[7] Koubaroulis, D., J. Matas, and J. Kittler. "Colour-based image retrieval from video sequences." In *Proceedings of the Czech Pattern Recognition Workshop*, pp. 1-12. 2000.

[8] Kerminen, Petteri, and Moncef Gabbouj. "Image retrieval based on color matching." *Proceedings of FINSIG* 99 (1999): 89-93.

[9] Konstantinidis, K., A. Gasteratos, and I. Andreadis. "Image retrieval based on fuzzy color histogram processing." *Optics Communications* 248, no. 4 (2005): 375-386.

[10] Liu, Guang-Hai, and Jing-Yu Yang. "Content-based image retrieval using color difference histogram." *Pattern Recognition* 46,no.1(2013):188-198.