International Journal of Advanced Trends in Computer Applications
www.ijatca.com

REVIEW OF CONTENT BASED IMAGE RETRIEVAL

[1]Shilpa, [2]Navdeep Singh
[1]Post Graduation Student
Department of Computer Engineering, Punjabi University,Patiala.
[2]Assistant Professor
Department of Computer Engineering, Punjabi University, Patiala.

Abstract: *Content-based visual information (CBVIR) or content-based image retrieval(CBIR) is one other important research areas in the field of computer vision. Many tools and programming have been developed to execute the queries based on the audio or visual content and that will help us to browse large multimedia repositories. A content-based image retrieval (CBIR) system is required to effectively and efficiently use information from these image repositories. This system helps the users to retrieve the relevant images from the database based on the content. In this paper we discuss various techniques for retrieving the images based on the contents of the image like color, shape and texture. Application areas where we used CBIR are numerous and diverse. We studied merits and demerits of various techniques of content based image retrieval.*

Keywords: *Content based image retrieval, Medical image retrieval, Biomedical research.*

I. INTRODUCTION

Content Based means that the search analyses the contents of the image rather than the metadata such as tags, keywords. The term "content based image retrieval" is originated in 1992 by T.Kato[1] which describes the automatic retrieval of images from the databases. Content Image Based retrieval is the process of searching and retrieving images from a huge dataset. As there are many images in dataset so it is difficult to search for the images which are similar to our query image. In medical and satellite image processing a large amount of images are there. To get an image from these databases require a technique which is called Content Based Image Retrieval(CBIR).Content Based Image Retrieval uses content like shape ,color and texture to retrieve query image from various image database. Characteristics of query image depend on the interest of the user like which type of image, scene, object, and texture he wants. On the basis of processing these specific queries leads to a different query types. First level tells that shape, texture and color are directly from images without using any background knowledge. Second level tells about queries that consist of logical features that require logical inference from low level information of the image. Third level contains queries consists of attributes that require prior knowledge and high level reasoning about the objects.

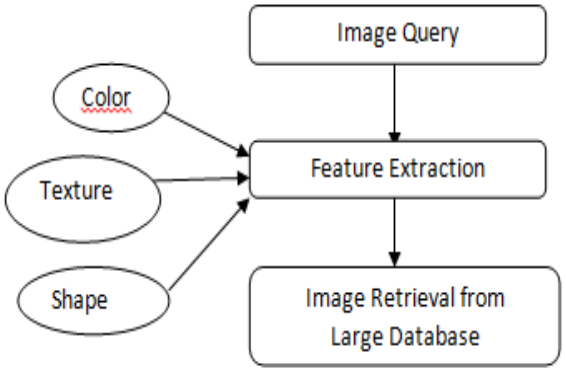

Figure 1: Image query Processing System.

Feature Extraction from image plays an important role in content based image retrieval. Feature extraction consists of both visual and metadata based features. Visual features like color, shape and texture plays an important role in retrieving of images. Feature Vectors describes the characteristics of an image. This is basically used for indexing and retrieval purposes of an image. Visual features like color, shape and texture are commonly used in content based image retrieval.

Color Features

Color descriptors include color histogram, colormoments, color coherence vector and colorcorrelogram.

Color descriptors they will extract the proportion of colors in the images, and then the images present in the database which is having less or more the same proportion of the colors can be retrieved using the similarity measure. Color histogram is used for retrieving the images if color distribution is unique in the image. Color moments include mean and variance and skewness .These three values are used to represent the distribution of color in an image. Color coherence vector combines color histogram with spatial information. This vector divides the histogram bin into two different types of bins that is coherent and incoherent by measuring part of colored similar large region. Colorcorrelogram combines the color distributions and spatial correlation of pairs of colors both. Following steps are used in retrieving an image using color features:

- Feature Extraction
- Histogram Calculation
- Similarity Matrix Calculation
- Dissimilarity Calculation

Texture Features

Texture descriptors include coarseness, contrast, directionality, line likeness and regularity.

Coarseness is used to analyse the size and number of texture patterns. It divides the texture patterns into coarse texture and fine texture. Coarse texture contains large patterns in small range and fine texture contains a small texture patterns in large range. Contrast represents the intensity difference between the adjacent pixels. If the intensity pixels is large then texture primitives has high contrast while the difference is low then the texture primitives has low contrast. Directionality refers to the global property about the region. It provides shape and placement of the texture primitives. Line likeness to the shape of the texture patterns .If the texture patterns are in the form of straight or wave, it is represented by line likeness feature. Regularity divides text patterns into regular and irregular texture based on arrangement of variations in texture patterns. Periodicity and scale are the two qualities of texture .Accessing the image using texture include arithmetic and structural analysis.

Shape Features

Shape representation can be divided into two categories Region based and Boundary based. The external edge of the shape is used in Boundary based representation. The outer description of the region like pixels is measured. Region based uses shape region in terms of 2D perspective but Boundary based uses one dimension approach. Specific characteristics of shapes are measured by number of corners, area and curvature. Shape descriptor is obtained by calculating coefficients using Fourier transform.

After feature extraction, various algorithms are used for similarity matching. K-means clustering algorithm is used to perform clustering. Similarity matching is mainly performed using Euclidean distance, Manhattan distance and histogram matching.

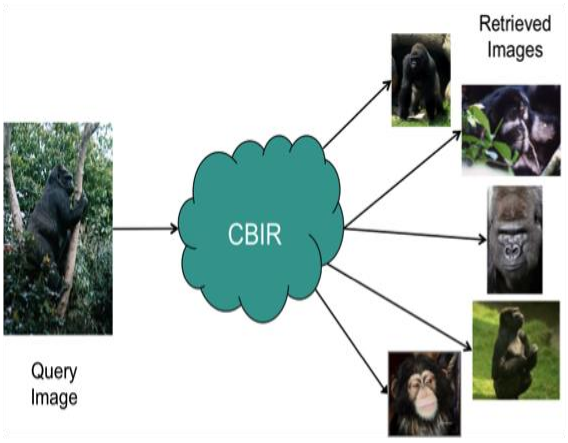

Figure 2: Example of content based image retrieval

II. Literature Survey

RoshiChoudharyet.al[2] proposed that contents of images are richer to get information as compare to text based image retrieval. By using single feature the results we get never be correct, so feature extraction is more beneficial to use for image retrieval. To extract color feature color moment and to extract texture feature local binary pattern (LBP) is used. The similarity matching which will compare query image with the images in the database used is Euclidian distance. Database used is Wang Database.LBP is always performed on gray scale images. LBP operator is used on natural images to extract texture from the images.LBP operator used in mainly for face recognition.

Manish K. Shriwaset.al[3] proposed a new technique for texture classification which will help to improve efficiency and accuracy of Content Based Image Retrieval system. Local Tetra Pattern (LTP) used for retrieving the images from database .LTP is based on the direction of pixel that are calculated by vertical and horizontal derivatives. By applying LTP with Gabor transform (GT) by considering diagonal pixel for calculation of derivatives will give the more accurate results. This is suitable mainly for biomedical research, crime prevention and for web images.

Mr SomenGhoshet.al[4] proposed an algorithm which is used in clinical support database system. An efficient technique is used for retrieval of malaria positive images from clinical database. The proposed method will identify the malaria presence by identifying the presence of chromatin dots in Red Blood Cells(RBCs).Algorithm use $-connected labelled region which will maps to

analyse and modify the image. Precision and Recall measures are used for efficiency. Precision is defined as the fraction of the images we retrieved relevant to the query image. Recall states the fraction of the relevant images that are retrieved from the database. X-rays, MRI images use this algorithm. This algorithm automatically helps to get different stains that are used in laboratories.

Jan-Ming Hoet.al[5] proposed a new technique for Content Based Image Retrieval system. To overcome the drawbacks of text based approach content based approach is used. In this technique we used segmentation and grid module, feature extraction module, K-means clustering and the neighbourhood color analysis model. neighbourhood concept module will recognise side of every grids of image. Results shows that this technique will not only works well for content based image retrieval but also improves its precision measure.

N.kavitha et.al[6] proposed a technique for the Content Based Image Retrieval for extracting the features of the query image .In this paper we will explore the technique by extracting the features of an image using color histogram(HSV),Polar Raster edge sampling technique,Fast discrete curvlet transformation for shape,color and texture of the image. Using this technique the feature vectors of the query image for color, shape and texture are extracted and then it is fused using genetic coding and the we use Euclidean distance for retrieving the similar images from the database which will match the query image. Curvlet is used to represent the texture information of an image at different pixel position and is given significant details for getting the better results. Similarity measures we use here is the Euclidean distance between the vectors of input image and the images in the database. This results prove that we get better retrieval based on multifeatures compared to single content features.

Jia song et.al[7] proposed a technique which will measure the precision priority.Improving Precision Priority(IPP) algorithm is used for integrating vital features and method to increase the performance. Two phases of IPP algorithm, In first phase both the query image and the database images are divided into blocks respectively. Then we calculate the color histogram of each block. After that calculate the Euclidean distance to compare the similarities. Allocation of different blocks with different weights but central pixel always contain some useful information so it must have more weight and the corners have less weight. All of the distances then are accumulated together to be the distance of whole image. In second phase we retrieve the shape characteristics using Hu moment invariants.

III. CBIR system working

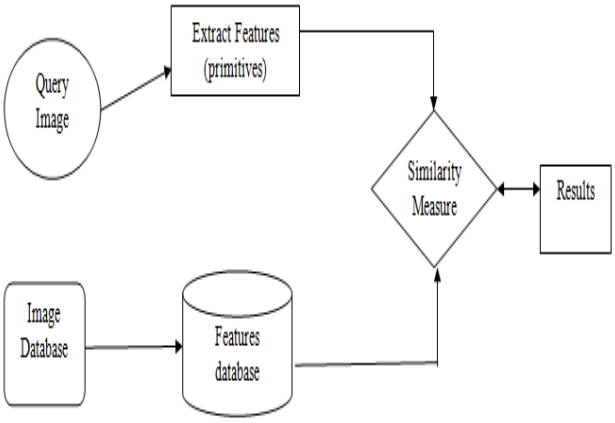

Figure 3: Flowchart of CBIR

First we will take the query image then we will extract the features from the query image by applying the extraction features using color, shape and texture features. Then we will select the dataset and extract the images from the dataset by applying the similarity measures like euclidean distance, City block distance, Cosine similarity. Then the images in the database which will match the features with the query images will be retrieved .There are various classifiers are used for the similarity measures. Figure 3 explains the processing of the query image and the images retrieved after applying the algorithms.

Applications: CBIR is used in crime prevention, Intellectual property, Face recognition, Industrial area, Medical Diagnosis, Education System,Earth Sciences.

IV. Techniques for content based image retrieval

Sr. NO	Domain	Researches	Proposed Method
1.	Color Images	RoshiChoudhary et.al[2]	LBP(local binary patterns).
2.	Natural color images	ManishK.Shriwas et.al[3]	Local Tertra Patterns.
3.	Malaria Positive images	Mr.Somenghosh et.al[4]	Presence of chromatin dots in RBC'S.
4.	Images	Jan-Ming ho et.al[5]	Segmentation and grid modules.
5.	Color images	N.Kavitha et.al[6]	Colorhistogram(HSV).
6.	Color and natural images	Jia song et.al[7]	IPP algorithm

Conclusion

Content-based image retrieval system is widely used in areas where we have large database containing images and we want to retrieve the images which are similar to query image. We have studied about various techniques like LTP(local tetra patterns) and LBP(local binary patterns) and color histogram to retrieve images from the database. These techniques are used in medical areas like face recognition. Similarity measures which are used here are distance between the query image and the images in the databases. In future we can use different techniques which will give better results in the different areas.IPP algorithm by calculating the color histogram is more beneficial for the retrieval of the images.

References

[1] Ms.Anchal" A Review of Query Image in Content Based Image Retrieval", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, April 2013.

[2] RoshiChoudhary, Nikita Raina, NesshuChaudhary, Rashmichauhan "An Integrated Approach to Content Based Image retrieval", 2014 IEEE.

[3] Manish K.Shriwas " Content Based Image Retrieval: A Past, Present and New Feature Descriptor" ,International Conference on Circuit, Power Computing Technologies[ICCPCT] ,IEEE 2015.

[4] Mr.SomanGhosh "Content Based Retrieval Of Malaria Positive Images From A Clinical Database" , 2013 IEEE Second International Conference on Image processing.

[5] Jan-Ming Ho, Shu-Yu Lin, Chi-Wen Fann, Yu-Chun Wang " A Novel Content Based Image Retrieval System using K-means with Feature Extraction", 2012 International Conference on systems and Informatics(ICSAI 2012).

[6] N.Kavitha ,P.Jeyanthi "Exemplary Content Based Image Retrieval" Using Visual Contents & Genetic Approach"peer-reviewed and accepted to be presented at the IEEE ICCSP 2015 conference.

[7] Jia song, Zunwen He "Content Based Image Retrieval by IPP algorithm" ,2015 IEEE.