

# Content Based Image Retrieval

Roshni Sahu<sup>1</sup>, Shital Zade<sup>2</sup>, Sonal Kamble<sup>3</sup>, Payal Bhagat<sup>4</sup>, Bhaghyashri Awachat<sup>5</sup>, Mayur Dakhane<sup>6</sup>  
<sup>1,2,3,4,5,6</sup>Computer Science and Engineering Department, Datta Meghe Institute of Engineering, Technology and Research,  
 Rashtrasant Tukadoji Maharaj Nagpur University, Maharashtra, India

**Abstract:-** To the explosive growth of digital image there is need of efficient and effective technique to image retrieval from database or from the internet. Now day no one can have that much of time that they will check or visit each image and check, is it required image or not. So Present new technique of image retrieval system which is based on the combination of both text and content based. The idea behind TBIR (text-based image retrieval) is that it uses the title, tag, character which is present inside the image i.e. metadata using this feature retrieve image. In CBIR (content-based image retrieval) image retrieving based on image feature i.e. colour, texture, shape, etc. Using these two techniques we can extract image from their feature. Main aim by using this type of system is to reduce the time when we retrieve one image from database.

**Keywords:-** Image Retrieval, TBIR (Text-Based Image Retrieval), CBIR (Content-Based Image Retrieval), Color, Shape, Texture, Metadata.

## I. INTRODUCTION

There is huge amount of image are creating in digital life and all are stored in database or gallery. In most of area this image are most important like academia, commerce, government sector, medical or health care, and internet.

When retrieve one particular image in this large database, and then they cannot access this image unless by using searching the image, scrolling, browsing and retrieving. One main problem is that it is difficult to locating or finding a desired image in this large collection of database.

Image retrieval is the way of searching the image, browsing the image and retrieving image from the large collection database of digital image. The main objective is that to develop a new technique which extract the image from this huge database by using the image feature which is based on colour, shape, texture, etc. [1]

To search image in database use two type of image retrieval technique i.e. text based image retrieval and content based image retrieval.

Text based image retrieval method which used the metadata to retrieve the image. In TBIR use text of image which is present in image. For performing this type of retrieving method for extract the text of image use OCR(optical character recognition) algorithm is use. After

performing OCR on image we can extract the text and perform the copy this text and also modify this text.

CBIR process is used to retrieve image from image .feature like colour, shape, texture.

In CBIR image retrieval by using colour feature use Histogram. for shape use kNN (k nearest neighbours) algorithm, and for texture also use kNN algorithm. In kNN algorithm use Euclidean formula. By using this formula we find image with their actual value and original value.[2]

## II. RELATED WORK

The previous system implement or design the combination of text based and content based image retrieval image feature. They use the Flickr sidebar to perform search and use image URL as a query input. When user perform text based image retrieval and enter any text to search or the input contain text based feature, then in the output inverted file is used for extract the text or the set of image that contain the input text. But when user does not enter any text for search then image is search in linear fashion over all the collection of image. [5]

Proposed method to extract feature from image. For retrieving image they use following techniques.

In this the texture classification they used statistical technique. Some of the statistical technique representation of texture feature use technique like, tamura feature, co-occurrence matrices, multi-resolution technique such as Gabor filter and wavelet transformation.

Shape based image retrieval are classified in two different categories i.e. object segmentation and shape representation. They also categories the shape representation in two other sub categories i.e. boundary based and region based. For this two categories of shape representation they use two method one is the Fourier descriptors and other is moment variant. [4]

Some of recent research paper or in review paper li and ma this two person work on the shape based image retrieval. They show the shape representation in two different method i.e. geometric moments method (region based) and Fourier descriptor (boundary based) this two method is related to the simple linear transformation. After li and ma second research is by babu etal. He compare performance of both the method (region based representation and boundary based representation) which is use in li and ma in shape representation. Babu etal combine

this representation. In this paper showed that the combined representation outperformed the simple representation. [6]

In content based image retrieval the earliest work is done by the two person (Ning-san chang and king-san Fu) they describe their paper by taking the example of Query-by-pictorial.

The Ning-san chang and king-san Fu, they use low level feature and same system used semantic for content based image retrieval. They also use the local feature in another perspective. The local feature is use to extract feature from the image such as handle the same image, matching the image, or image searching that contain the same image/same object or same scene in query image with different direction or different viewpoint, they have different scale, change in illumination, occlusions. [3]

### III. METHODOLOGY

#### A. Text based (Algorithm: OCR (optical character recognition))

OCR algorithm is use to scan the text of image and this text can be use to modified the data. OCR extract the text from the image. It capture the image text a via pattern either word by word or letter, line, etc. OCR algorithm is

used to capture the image content i.e text, line, word, character, etc. In text based image searching OCR gives better performance.

#### B. Content based image retrieve

##### ➤ color based: Histogram

Histogram is graphical representation of a image. In histogram plot graph according to the image pixel value. It text the image pixel value and draw the graph and this graph is compare to all other image in database. Create histogram for query image find all 256 value of pixel instead of to store image color. By using distance function calculate all distant image. Retrieval image 256 scale in each RCB channel.

##### ➤ shape based: KNN (k nearest neighbors)

KNN is a type of classification algorithm. The neighbor are classified by nearest object. Distance of all image is calculated from different formula like manhattan, eculidean distance. Number of output image is depend on 'K' value. If suppose value of K is equal to 2 then their is two image in output.

##### ➤ texture based: Gabor-filter

### IV. ARCHITECTURE

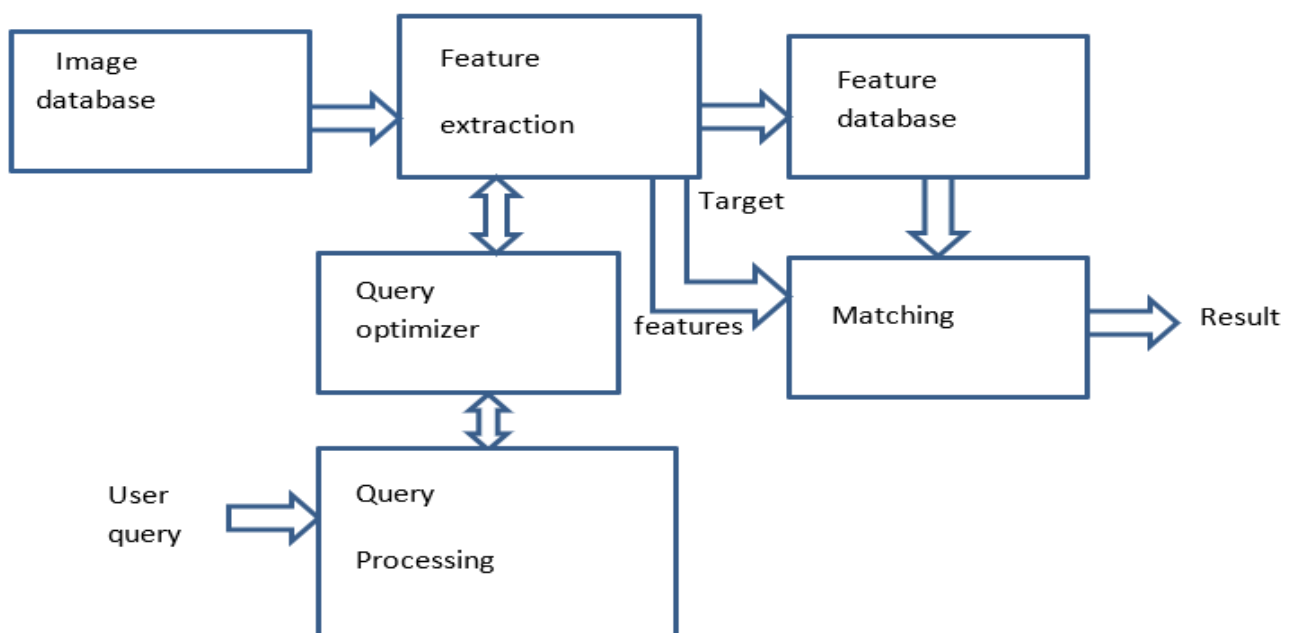


Fig 1:- Architecture of content-based image retrieval

- User query: There is user requirement viz, size, type, date, time of capture, texture, colour.
- Query processing: The image details can be either manually supplied or generated automatically using suitable image processing technique. The input query is then processed.
- Feature extraction: It extract the feature from input image such as text, colour, shape.

- Image database: It include image detail such as size, type, date and time.
- Feature database: It contain the information of the extracted feature of input image.
- Matching: These processed quires are then matched with the stored data based and the query images are retrieved. The resulting images is based on the queries entered by the user.

### V. WORKING

User can get input in the form of user query, when user have to search image or query goes input to the query processing. In this block the user query is processed using suitable image processing technique. After image processing, query optimization is use to optimize the dataset. Then the feature is extract in input image i.e. text, colour, shape, texture, etc and this feature is check or compare to the image database and feature database. After performing all this operation match the image database feature and feature database and at last after performing this image retrieval technique we have only similar data, shoe this data as result.

### VI. RESULT

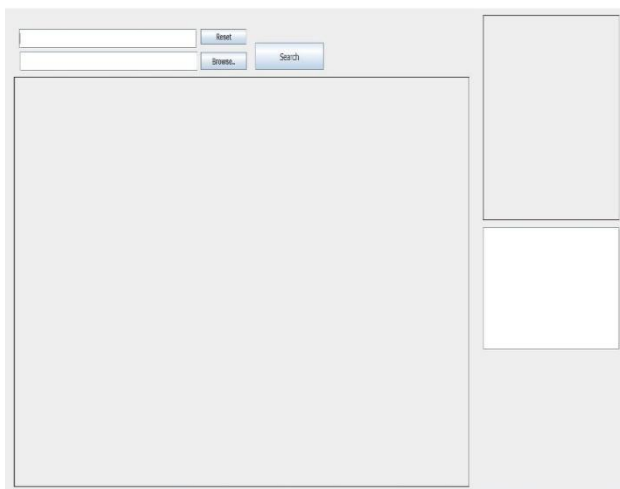


Fig 2:- Front View

As shown above fig. here it is interface of our project there are 3 buttons reset button is used to remove all text or image from text field and panel. Browse button is used to browse image and path of images is shown in text field. Search button is used to search the browse image.

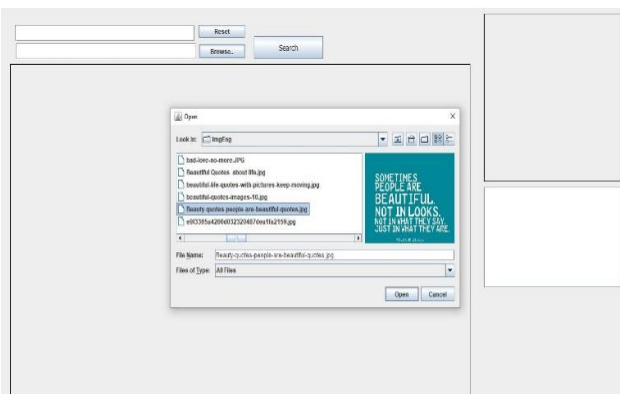


Fig 3:- Browse image for performing OCR

Here the image is browse when we click on browse button dialog box is open and here is path of imaged from that file we select one image. In right panel image preview is shown, which is use to help select required image.

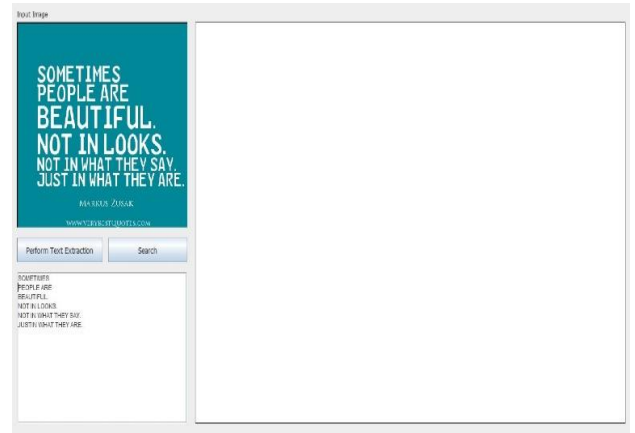


Fig 4:- OCR Perform for selected image

Here is window after clicking on search button. there is one panel that show the browse image and here is one button "perform text extraction" after clicking on this button image text is extracted and shown in text area below the button.

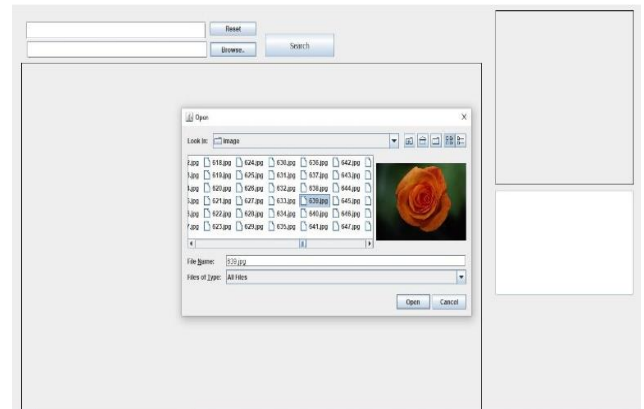


Fig 5:- browse image for retrieving color image

In this browse the image for performing color base image retrieving. Click on browse button after clicking dialog box is open in this path of image is given, where the image is store then choose the color image which color image you have to extract. Then click on open button. In right panel image preview is shown, which is use to help select required image.

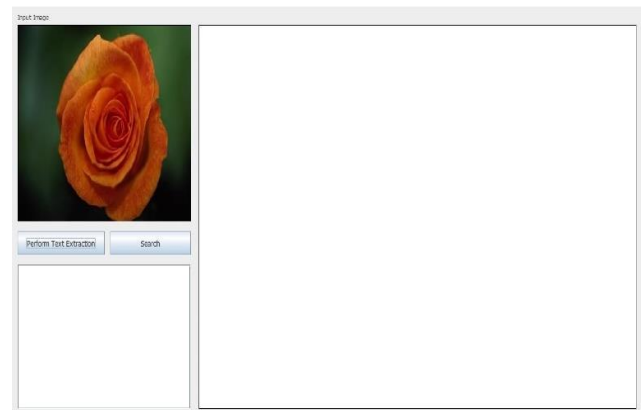


Fig 6:- Search query image

After browse the image for color-based image retrieval, this window is open. Here two button is available i.e. “perform text extraction” and “search”. By clicking on “perform text extraction” button text is extract if text is present in input image. And after clicking the “search” button the similar image (similar color image) is found.

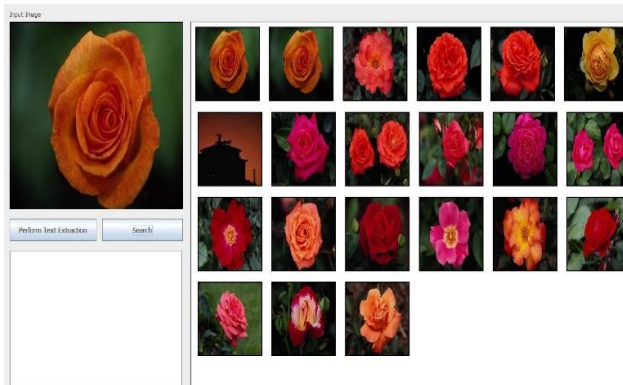


Fig 7:- Output of query image

After click on the search button histogram algorithm really work and apply histogram algorithm and calculate similar image by taking input as query image. List of similar image is shown in matlab because as compare to java matlab give better clarity of image. In final result first image is query image (input image) and other image is similar to that image.

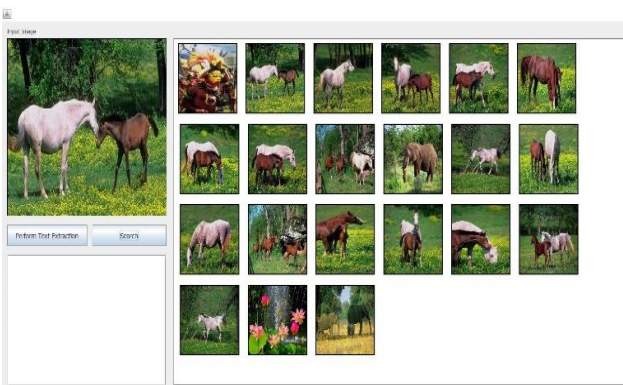


Fig 8:- Output of query image for performing shape-based image retrieval

After performing color-based retrieval next type is shape based image retrieval steps are same stated above in color based. After clicking on search button image is displayed and all this image have same shape as query image.

## VII. CONCLUSIONS

We present system that allow a combined search of text and content based feature. In text based we provide a keyword, description or image metadata in side bar. In CBIR we extract image feature i.e. shape, color, texture. Using both the technique we improved the performance speed and reduce searching time. Gives the better result using matlab and java both interface.

## REFERENCES

- [1]. Wengang Zhou, Houqiang Li, and Qi Tian Fellow, “Recent Advance in Content-based Image Retrieval: A Literature Survey” IEEE,2017
- [2]. Mohammed Alkhwilani, “Text-based, Content-based, and Semantic-based Image Retrievals: A Survey Dept.of Information Systems”, Faculty of Computers and Information, Mansoura University Mansoura, Egypt, 01 January 2015
- [3]. T.Karthikeyan,P.Manikandaprabhu, S.Nithya,Associate, “A Survey on Text and Content Based Image Retrieval System for Image Mining” , Department of Computer Science ,PSG College Of Arts and Science, Coimbatore, Tamilnadu, India.3 March 2014
- [4]. Amandeep Khokher,Rajneesh Talwar, “Content-based Image Retrieval: Feature Extraction Techniques and Applications” RIMT-Maharaja Aggrasen Engineering College Mandi Gobindgarh,2012
- [5]. Juan Manuel Barrios, Diego D’iaz-Espinoza, and Benjamin Bustos, “Text-Based and Content-Based Image Retrieval on Flickr: demo”, Department of Computer Science University of Chile Santiago, Chile,2009
- [6]. Debashish pal, Madhusudhan Mishra, “literature survey on content-based image retrieval (CBIR) technology”, North Eastern Regional Institute of Science and Technology Madhusudhan Mishra, April 2014