

# Utilizing Effective Way of Sketches for Content-based Image Retrieval System

Dipalee J. Fendarkar  
Research Student  
Government Engineering College Aurangabad  
Maharashtra, India

A. K. Gulve  
Asst. Professor CSE Dept.  
Government Engineering College Aurangabad  
Maharashtra, India

## ABSTRACT

This paper reviews the development of content based image retrieval system (CBIR) and sketch based image retrieval system (SBIR). In this paper, we present the problems and challenges concerned with the design and the creation of CBIR systems, which is based on a free hand sketch (i.e. SBIR). The use of the existing methods, describe a possible results, how to design and implement a task specific descriptor, which can handle the informational gap between a sketch a colored image to make an opportunity for the efficient search. The CBIR system first computes the similarity between the query and the images stored in the database.

This paper outlines a description of primitive features of images like textures, colors and shapes. The review of the feature extraction, feature based matching and indexing which represent the base of retrieving images and it allows the comparison of database images with queries containing various levels of detail, thanks to a hierarchical representation of the database images. In this paper, we create the mosaic of the images and compared the methods of matching sketches descriptor.

## General Terms

Sketch based Image Retrieval (SBIR), Content Based Image Retrieval (CBIR)

## Keywords

Content Analysis and Indexing, Indexing methods, Digital Images, Image Descriptors, K-means algorithm, Image Database.

## 1. INTRODUCTION

In computer technology, Database and Internet are the most basic part of the searching information. These are mostly based upon the text search. Although there are many advanced machines to handle queries and storing the vast amount of information. The efficiency of search cannot be same. The methods of searching using text based content are not efficient because most of the visual Information cannot be expressed textually. This situation can be avoided by using CBIR technology [1].

The CBIR technology is necessary for the image; the image search is made efficient and dependable. CBIR is allows to extract information which itself represent the image. Content-based image retrieval (CBIR) is also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR). Application of computer vision to the image retrieval problem that is the problem to search for digital images in the large databases [2]. The term 'content-based' means that the search can be analyzing the actual content of the image. Content must be relay on metadata to examine the image content, searches without the facility like captions or keywords. This type of metadata must be

generated by a human and stored alongside each image in the database. The rise of interest in techniques for retrieving

Images on the basis of automatically-derived features such as color, texture and shape. This technology generally referred to as Content-Based Image Retrieval (CBIR) [2].

The aim of this paper is to review the current state of the art in content-based image retrieval. The word 'content' in CBIR refers to feature of images like color, shape, texture, edges etc. which can be extracted from that image itself. Thus a system which uses features of image will give better indexing and give accurate results. The work is done on inventing an easier and smarter system called Sketch-Based Image Retrieval (SBIR). It is a system that basically operates to show the user's query by line-based hand drawing. The query picture can be well expressed by SBIR through shape only while it needs more than that to be expressed by CBIR. The SBIR system adds new interesting variety of hand drawing style and even more flexibility to the query users.

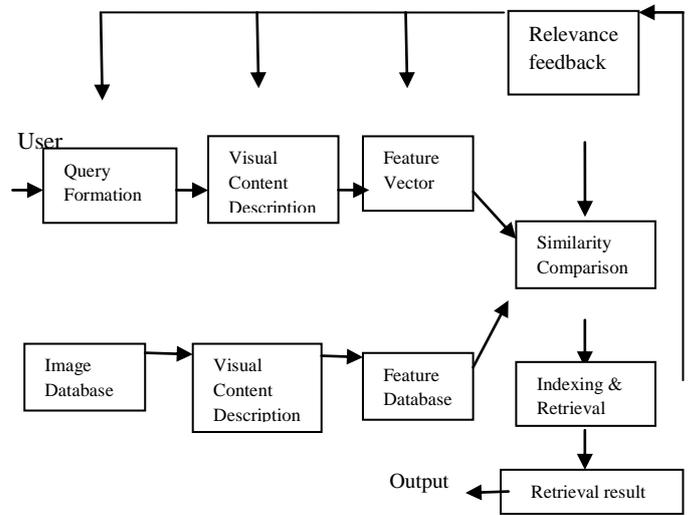


Fig. 1 Diagram for content-based image retrieval system

Above figure shows the flow diagram of content based image retrieval system [3].

## 2. LITERATURE SURVEY

### 2.1 Content based Image Retrieval (CBIR)

The visual content of image is used to match similar images in content based image retrieval system. The term 'visual content' refer to image's characteristics such as its color, shape, texture, object or any other information which can be derived from image itself. CBIR is popular because most of the web-based image search engines purely consistent on metadata and this produces a lot of garbage in the results [4].

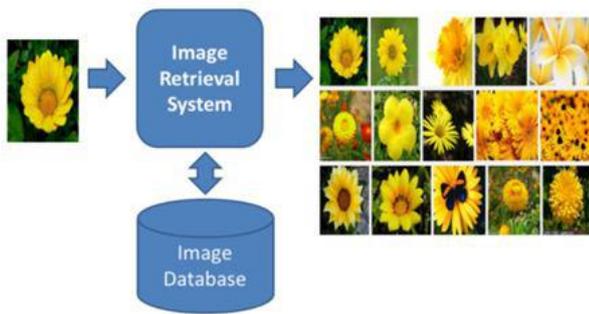


Fig.2 Content Based Image Retrieval System

Human are manually entering the keywords for images in large databases. It is efficient, expensive and may not capture every keyword that describes the images. Thus a system that can filter images based on their content would provide better indexing and return more accurate results. Above figure 2 shows working of CBIR [4]. A CBIR consist of two main parts as follows:

**1) Offline part:** The offline part deals with building the database for images. All the images are analyzed, their features are extracted and according to their features images are clustered and indexed.

**2) Online Part:** In online part, user gives the image as an input. The image features are extracted and those extracted features are matched with the features present in database. If the features get matched then the image corresponding to that feature is similar to the input image. In this way, the matching images are retrieved and they are displayed on output screen in descending order [4].

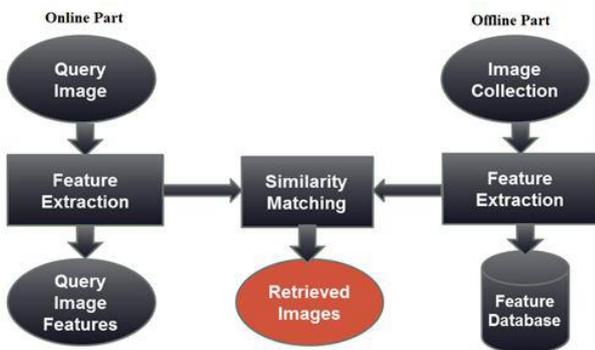


Fig.3 Architecture Diagram of CBIR

Figure 3 shows the architecture diagram of CBIR [4].

There are different techniques followed by CBIR as follows:

**a) CBIR Using Color Feature Extraction**

Color is very crucial aspect which is used by humans to differentiate things. Color feature extraction can be done in following ways:

- I) Color Histograms
- II) Color Correlograms [4]
- III) Color Layout Descriptor (CLD)

IV) MPEG-7 Descriptor

V) Color and Edge Directivity Descriptor (CEDD)

VI) Fuzzy Color and Text Histogram (FCTH)

Explanation of each color feature extraction techniques is as follows:

**I. Color Histograms:** Color is most important aspect for humans, because through color they can distinguish things properly. Color of an image is generally expressed through color model. Color model is defined in three dimensional color spaces as follows [4, 5]:

**a. RGB:** RGB colors are known as primary colors. They are additive. New colors can be prepared by combining them differently.

**b. HSV (Hue Saturation Value):** The HSV space is resulting from RGB space cube with main diagonal of RGBI model with vertical axis in HSV. As capacity varies from 0.0 to 1.0, the colors vary from unsaturated (gray) to saturate (no white component). With the dissimilarity beginning with red, going away through yellow, green, cyan, blue and magenta back to red in Hue ranges from 0 to 360. Information about color is expressed through color histograms. In Color Histogram, every image is analyzed and its histogram is produced. In that histogram, color proportion for each pixel is shown. Color Histogram is the most universal method for color based image retrieval. The color histogram is obtained by quantizing image colors into discrete levels and then counting the number of times each discrete color occurs in the image. Color Histograms are further divided into two types:

**a. Global Color Histograms:** In GCH, a single image will have a single color histogram [4, 5].

**Advantage:** Less computation is required.

**Disadvantage:** In GCH, less information about an image is extracted. GCH is considered as conventional method for color based image retrieval. Sometimes we can get conflicting results.

**b. Local Color Histograms:** In LCH, an image is divided into segments, and further color histograms of those segments are calculated.

**Advantage:** In LCH, much more information about an image is extracted.

**Disadvantage:** More calculation is required because; every segment has to be computed. Perfect results are obtained [4, 5 and 6].

**II. Color Correlogram:** It deals with color distribution of pixels and spatial correlation of pairs of colors. Let I be an image that comprises of pixels f(i, j). Each pixel has certain color or gray level. Let [G] be a set of G levels g1, g2, .. .g. G that can occur in the image. For a pixel f let I(f) denote its level g, and let Ig correspond to a pixel f, for which I(f) = g. Histogram for level gx is defined as [4, 5]:

$$hg_x(I) \equiv Pf \in I | f \in I_{g_x}$$

Second order statistical measures are Correlogram and autocorrelogram. Let [D] denote a set of D fixed distances d1, d2, . . . , dD. Then the correlogram of the image I is defined for level pair (gx, gy) at a distance d.

$$g_x, g_y(I) \equiv Pr$$

$$f_1 \in I_{g_x}, f_2 \in I_{g_y} [f_2 \in I_{g_x} || f_1 - f_2 = d]$$

Which gives the probability that given any pixel  $f_1$  of level  $g_x$ , a pixel  $f_2$  at a distance  $d$  in certain direction from the given pixel  $f_1$  is of level  $g_x$ . Auto correlogram captures the spatial correlation of identical levels [4, 5, and 6].

**III. Color Layout Descriptor (CLD):** The CLD is a very compact and resolution invariant representation of color for high speed image retrieval [7]. It captures the spatial layout of the representative colors on a grid superimposed on a region or image based on the Discrete Cosine Transform (DCT). It is expressed in the YCbCr color space. The size of the array is fixed to 8x8 elements to ensure scale invariance. It is then transformed using DCT followed by zig-zag re-ordering. This is shown in the figure given below. It can be used for fast searching of databases as well as filtering in broadcasting applications. Another application is description of video clips, where CLD is combined with the Time Series structure [7, 8].

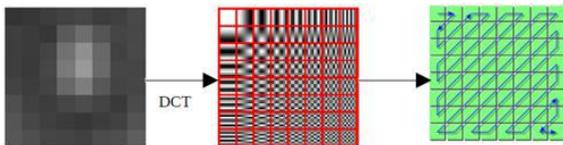


Fig.4 The CLD Extraction Process [7, 8]

**IV. MPEG-7 Descriptor:** MPEG-7 has been developed by the Moving Pictures Expert Group (MPEG), a working group of ISO/IEC [ISO00]. The goal of the MPEG-7 standard is to provide a rich set of standardized tools to describe multimedia (i.e. audio-visual) content. Unlike the preceding MPEG standards (MPEG-1, MPEG- 2, MPEG-4) which have mainly addressed coded representation of audio-visual content [9]. At the different levels, MPEG-7 focuses on representing information about the content. The structural level is supported in the same way as the (visual) feature level or the semantic level. The content itself is out of the scope of the standard and MPEG-7 states explicitly that the description tools are applicable for all kinds of multimedia content independent of its design and code. The method as well as technology generate and use the descriptions. These are not the part of the standard and the tools [9]. These are not restricted to a specific set or class of applications. This is the powerful concepts to reach this goal but MPEG-7 restricts itself. **MPEG-7** is also called as Multimedia Content Description Interface. It is a new standard for describing the content of multimedia data [ISO00, MPE99, MPE00, and MPE01]. MPEG-7 is a means that the attach metadata to multimedia content. It specifies a standard set of depiction tools, which can be used to describe various types of multimedia information. These tools shall be linked with the content itself to allow efficient and effective searching for multimedia material of users interests. MPEG-7 is a basic standard with broader application area.



Fig.5 Three color images and their MPEG-7 histogram color distribution [9]

**V. Color and Edge Directivity Descriptor (CEDD):**

The unit associated with the extraction of color information is called Color Unit. Similarly, the Texture Unit is the unit associated with the extraction of texture information. In Histogram, descriptor specifies the information of color, Edge and Texture by using Color and Edge Directivity Descriptor and incorporates color and texture information. The Color and Edge Directivity Descriptor (CEDD) size is limited to 54 bytes per image, representation of this descriptor suitable for use in large image databases. The CEDD is the one of the most important attribute is the low computational power needed for its extraction and in comparison with the needs of the most MPEG-7 descriptors [10].

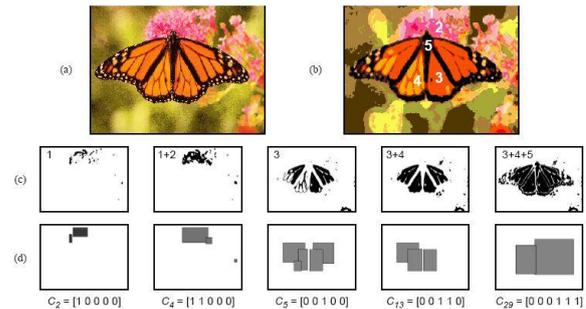


FIGURE 2. (a) Butterfly color image, (b) processed color image with 30 colors, (c) pixels from image (b) belonging to color set  $C_i$ , (d) minimum bounding rectangles (MBRs) for extracted regions used to index the image collection.

Fig. 6 Image Retrieval Using CEDD [10]

**VI. Fuzzy Color and Text Histogram (FCTH0):** This descriptor is considers degree of color similarity between color and texture. Fig. 6 shows image retrieval using FCTH. Fuzzy color histogram-based shot-boundary detection algorithm specialized for content based copy detection Applications and the color histogram generated with the fuzzy linking method on  $L^*a^*b^*$  color space, the system extracts a mask for still regions and the window of picture-in-picture transformation for each detected shot. This will be useful in a content-based copy detection system. The results show that our method effectively detects shot boundaries and reduces false alarms as compared to the state-of-the-art shot-boundary detection [10].



Fig. 7 Image Retrieval using FCTH [10]

**2.1.1 Common Feature (Color, Shape and Texture)**

Descriptors which represented color and used as invariant color, features color space, histogram, and moments color coherence dominant color, these features give the brightness and geometry view as a result of an image retrieval process [11]. The idea t of the moment, which represent the main factor in color distribution by using an image retrieval system as a method for choosing a query image by calculating color histogram and comparing it with the stored image database. The main two shape characteristics are bound-based and region-based , which are computed for every object specified in the database, Yasmin at 2013 proposed the geometric

details of image which can provide us the local shape features, which have a large area in CBIR systems applications. Also, a part of color features of the shape can be calculated from the image segmentation. The main shape feature of any image is one of the invariant features (translation, rotation, and scale) along the axis. Many different available methods improved to extract and represent textures from image property. Basically space based frequency-based models, and texture signature model which are defined in the term of granularity, directionality and repetitiveness features. These are the characterized of this feature [11].

### 2.1.2 Semantic Gap

The relationship between low level features (color, shape, Texture, object detection) and high level user features (abstract, objects, event) categories defined by Wang. CBIR indexes the images by using the low-level features, then displays an interpretation opposition between image description and high-level semantics, this process call semantic gap. The researchers tried to bridge the semantic gap by proposing many techniques [11].

### 2.1.3 Performance Measure

The image retrieval system is the main technique for precision and recall graph which represents the main way to evaluate and measure the performance [11].

**Precision= no. Of relevant images retrieved /Total no. Of image retrieved.**

**Recall= no. Of relevant images retrieved / no. Of relevant images in the database.**

## 2.2 Sketch Based Image Retrieval (SBIR)

Sketch-based image retrieval (SBIR) is a relevant means of querying large image databases. All of researches focus on how to solve the gap between sketch and image matching problem [11]. A lot of ways are discussed or discovered about this gap. Recently, we have reviewed bellow a method that deal with the main three points of sketch based image retrieval. The global structure includes the blocks such as Preprocessing subsystem, feature vector generating subsystem, Retrieval subsystem, and Database management subsystem & displaying subsystem. Fig. 1 shows the global structure of the system [11].

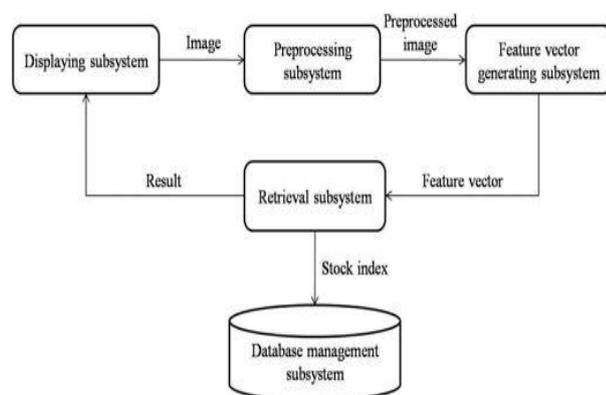


Fig. 8 The global structure of the system [12]

### 2.2.1 Preprocessing subsystem

The system is for databases containing simple images. This eliminates the problems due to multiplicity of images. To avoid the occurrence of noisier images, a multi-step preprocessing generates descriptors. In this subsystem input is one image, and the processed set of result is the output [ 12].

### 2.2.2 The Feature Vector Generation Subsystem

Feature vector is point at which particular value is stored or interesting point of image. This subsystem consists of descriptor vectors, which signify the content of the image. There are three descriptor vectors that is Edge Histogram descriptor (EHD), Histogram of Oriented Gradients (HOG), Scale invariant feature transform (SIFT) [12].

The function of descriptor is Preprocessing of free hand sketches, Compression of free hand sketch with gallery of images, Retrieval of matched images from the database[1].

In Sketch Based Image Retrieval (SBIR) system, we get the required images using other image. We discover images via some features of image. This Histogram characteristic is representing the feature composition of an image.

#### 1. EHD (Edge Histogram Descript or) concept

In EHD method, find the edges in image and these edges are used to store the feature value. To restrict the edge distribution to a certain area of the image, this image space is divided into 4 x 4 sub – images as shown in Fig. 5. An edge histogram generated for each sub-image to characterize sub-image edge distribution. [12]

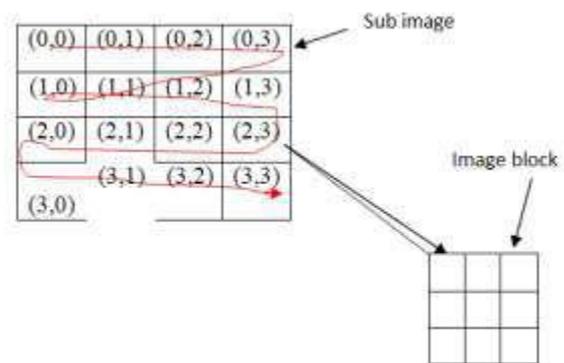


Fig. 9 Sub-image & image-block

The sub-image is more divided into small square blocks called image-blocks. For classify different edge types [12].

There are different 5 types of edges.

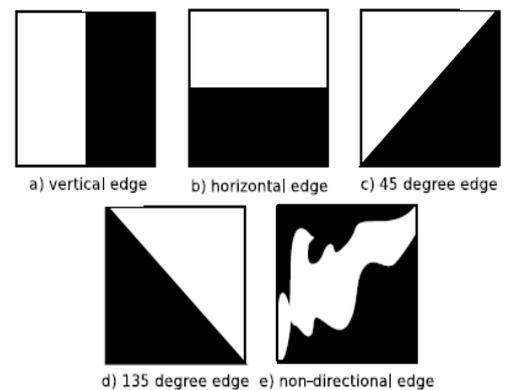


Fig. 10 Types of Edges

In each sub-image for each edge type count the total number of edges, after the edge extraction from image-blocks. There are five types of edges, for each sub-image classify five histogram bins. While there are 4 x 4 = 16 sub-images, total 16 x 5 = 80 bins for the edge histogram. For the histograms bin values can be obtained from the local histogram [12].

### Advantages of EHD

- i. Semi global and global histogram generated from the local histogram bins help to improve the retrieval performance.
- ii. Extra histogram information can be obtained directly from the local histogram bins without feature extraction process.

### Disadvantages of EHD

- i. The expansion is so difficult and robust descriptor is emphasized.
- ii. Not provide invariant opposite rotation, Scaling and translation [12].

### 2. HOG (Histogram of Oriented Gradient Descriptor) Concept

Gradient is on which intensity of image change as per direction. Here consider gradient at low to high level, in HOG occurrences of gradient count. In these descriptor, first it converts the image into small connected regions, called cells. Pixels within the cell declare gradient directions & edge orientations [2]. This descriptor results in better invariance to change in illumination or shadowing [12].

### Advantages of HOG

- i. HOG based retrieval is much better than EHD based retrieval.
- ii. It is used for large database.
- iii. For information poor sketch problem can be solved by HOG.

### Disadvantages of HOG

- i. Humans are highly expressed that including a parts based model with a greater degree of local spatial invariance [12]

### 3. SIFT (Scale invariant feature transform) concept

The SIFT feature detection which is used for object recognition. The invariant features extracted from images can be used to perform reliable matching between different views of an object or scene [12].

### Advantages of SIFT

- i. The image gradient magnitudes orientations are sampled around the key point location.
- ii. Image rotation and scale and robust across a substantial range of affine alteration, addition of noise, and change in illumination.

### Disadvantages of SIFT

- i. Perform the feature matching even the faces are small.
- ii. It is difficult to define and it is hard to detect but large number of key points can be extracted from typical images [12].

#### 4.1.1 K-mean Algorithm

The K-mean algorithm is used for better retrieval of image. In this k-mean clustering grouped object belong to same class. Find the mean between two classes. It is used to check the similarity. K-means method to represent the cluster centroid is used and sensitive to outliers [13].

**Input:** k is the number of clusters and a database containing n objects.

**Output:** k is the set of cluster that minimizes the squared error criterion.

The input is a query, and then the system extracts the image feature and measure the distance with images in the database and generates output which is the list of retrieval images. User choose the related image & compare with the query and perform the loop as long as it get same image belong to query image.

### Disadvantages of K-mean algorithm

Used for small database.

Unable to handle noisy data and outliers.

Does not work well with non globular clusters [13].

All the above algorithms are used for feature extraction or to get the feature vector of the SBIR system [13].

## 3. PROPOSED SYSTEM

In this paper we proposed to design, implement and test a sketch-based image retrieval system. Two main aspects were taken into an account. The retrieval process has to be unconventional and highly interactive. The robustness of the technique is important in some scale of noise, which might also be in case of simple images. The drawn image without modifications cannot be compare with color image, or its edge representation. On the other hand a distance transform step was introduced. The simple smooth and edge detection based method was improved, which had a similar importance as the previous step. User can choose the relevant image to further process the query, and this process can be iterated many times until the user find the desired images. Also we can compare the different methods of feature matching sketch which one is better and check the effectiveness of the methods. Also we can create the mosaics of images. The goal is to develop a content-based associative search engine, which databases are available for anyone looking back to by hand drawing. The user has a drawing area, where he can draw all shapes and moments, which are predictable to occur in the given location and with a given size. The most important task is to bridge the information gap between the drawing and the picture, which is helped by own preprocessing renovation process. In system the iteration of the utilization process is possible, by the current results look again, thus raising the precision.

## 4. REFERENCES

- [1] Surekha Karpe, Sneha Shree, Saurabh Kurkute, Abuzer Shah, Mrs. Archana Chaugule, " Sketch4Match: Content Based Image Recognition using Coloured Sketches", International Journal of Engineering Trends and Technology (IJETT)- Volume 9 Number 10- Mar 2014.
- [2] Junaid Khan, "Content Based Image Retrieval by Clustering", International journal of Advance Research in Computer Science and Management Studies (IJARCSMS) – Volume 2, Issue 1, January 2014 pg. 708-714.
- [3] D.T.V. Dharmaji Rao, BH. Anantha Laxmi, "A Content Based Retrieval System Using Graphs", International Journal of Advanced and Innovative Research (IJAIR) - ISSN: 2278-7844.
- [4] Prajakta Mahendra Ovah, Prof. B.D. Phulpagar, "A Survey On Various Feature Extraction Techniques

- Followed By Image Retrieval System”, *International Journal of Emerging Trends and Technology in Computer Science (IJETTCS)* , Volume 3, Issue 5, Sept-oct. 2014 – ISSN 2278-6856.
- [5] Ms. K. Arthi and Mr. J. Vijayaraghavan, “ContentBased Image Retrieval Algorithm Using ColourModels”, “*International Journal of Advanced Research in Computer and Communication Engineering*”, Vol.2, Issue 3, March 2013.
- [6] Swati V. Sakhare and Vrushali G. Nasre, “Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture”, “*International Journal of Computer Science & Informatics*”, Volume-I, Issue-II, 2011.
- [7] Sikora Thomas, “The MPEG-7 Visual Standard for Content Description—An Overview”, *IEEE Transaction On Circuits and System for Video Technology*, June 2001. Vol. 11, No. 6 pp. 696-702.
- [8] Roshini Johri, “Color Descriptors from Compressed Images”.pdf.
- [9] Ala’a Al-zoubi, Mohammad Al-zoubi, “Automatic Mapping of MPEG-7 Description to relational Database”, *The International Arab Journal of Information Technology*, volume 9 number 6, November 2012.
- [10] Priyanka B. Khairnar, Aarti A. Khairnar, Priyanka R. Khandre, Priyanka S. Mahajan, “Content Based Image Retrieval System”, 1st International Conference on Recent Trends in Engineering & Technology, Mar-2012 Special Issue of *International Journal of electronics, Communication & Soft Computing Science & Engineering*, ISSN: 2277-9477
- [11] Huda Abdulaali Abdul Baqi, Ghazali Sulong, Soukaena Hassan Hashem, “A Sketch Based Image Retrieval: A review of Literature”, *Journal of Theoretical and Applied Information Technology*, 10<sup>th</sup> may 2014, volume 63 no. 1, ISSN: 1992-8645.
- [12] Asmita A. Desai, “Review Paper on Sketch Based Image retrieval System”, *Proceedings of Sixth IRAJ International Conference*, 6th October 2013, Pune, India. ISBN: 978-93-82702-32-0.
- [13] B. Sz’ant’o, P. Pozsegovics, Z. V’amosy, Sz. Sergy’an, “Sketch4Match – Content-based Image Retrieval System Using Sketches”, *SAMI 2011 • 9th IEEE International Symposium on Applied Machine Intelligence and Informatics • January 27-29, 2011 • Smolenice, Slovakia*.