International Journal of Engineering & Technology, 7 (4.5) (2018) 471-476

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Survey on content based image retrieval

Ankitha Varma [1] *, Dr. Kamalpreet Kaur [2]

[1] School Of Electronics and Communication Engineering, LPU, Phagwara, India
[2] Division of Research and Development, LPU, Phagwara, India
**Corresponding author E-mail: ankithavarma52@gmail.com*

Abstract

Now-a-days, because of the advancement in the digital technology and the use of internet, a huge amount of digital data is available in the form of medical images, remote sensing, digital museums, geographical information, etc. This has lead to the need of accurate and efficient techniques for the search and retrieval of relevant images from such voluminous datasets. Content based image retrieval (CBIR) is one such approach which is increasingly being used to search and retrieve query image from the databases. CBIR combines features of color, texture as well as shape which ease out the process of extracting desired information from the retrieved images. This paper pre- sents a systematic and a detailed review of the CBIR method along with the different databases and evaluation parameters used for the analysis. An attempt has been made to include an exhaustive literature survey of the various CBIR approaches.

Keywords: *Image Retrieval; CBIR; Color; Texture; Shape; Database; Performance Evaluation.*

1. Introduction

In the present era, internet has become an integral part of the human life. Now-a-days, almost everyone is equipped with gadgets like laptops, tabs, multimedia mobile phones etc. These devices have good quality of inbuilt cameras, which increases tendency of capturing pictures and recording videos. It has become so easier to share the images/videos on social networking through these multimedia devices, which has lead to a huge amount of pictures of every kind i.e. desired and undesired being uploaded on internet. Different digital images are generated and uploaded on internet daily by persons who belong to various fields like medicine, education, engineering [1] [26]. The retrieving of these images from cyberspace is an interesting but a challengeable task. By using various search engines, whenever a user wishes to retrieve an image, large number of images from web are retrieved. Most of them are unrelated to user query. The user feels difficult to search the exact images, as useful images are distributed amongst the unrelated ones on different pages. On the other hand, efficient and effective manner results cannot be yielded, as users are forming random queries. This makes user unsatisfied most of the time. Lot of efforts has been done towards efficient retrieval of useful imag-
from large database has become crucial. This probably motivates for extensive research into a system of retrieval images [1] [19].

2. Background study

The earlier image retrieval systems are text based as images are required to be indexed and annotated accordingly. With the in- crease of image size and the size of image database, task of user- based annotation has becomes difficult, thereby incompletion as the text usually fails to convey structure of images. In the early 1990's, in order to overcome these difficulties, CBIR was pro- posed where retrieval is based on automatic matching feature of query image and also some evaluation of similarity images are also annotated. Technology is growing vast day-by-day and usage of digital images increases in different fields for extensive mecha- nism of image retrieval and has become to a great extent for pro- cessing of desired image where researchers has been working [2] [3] [15].
Feature extractions from web by various researchers. But till date an efficient solution to this problem has not been proposed which remains as an
Open challenge [2] [19] [20]. A lot of review work exists for image retrieval process, but still a lack of systematic image database and performance evaluation of existing techniques for CBIR is realized. From the past years, CBIR systems have enhanced a lot, but still some problems have not been responded satisfactorily which can explore research gaps and statistical information for future researches. In few years, size of multimedia files increases because of increase of memory storage and wide spread of World Wide Web. So the need for the efficient tool for image retrieval
Multi-dimensional indexing

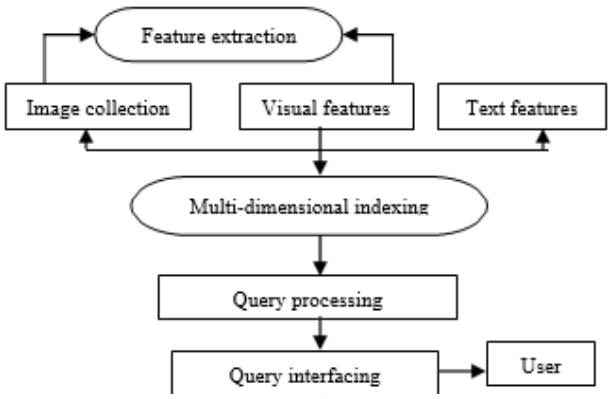


Fig. 1: Architecture of Image Retrieval.

Copyright © 2018 Ankitha Varma, Dr. Kamalpreet Kaur. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IMAGE RETRIEVAL: Image retrieval is a study that concerns with searching, browsing and retrieval images from digital image database on the basis of texture, color, structure or content. It plays a major role in the field of digital image processing, digital libraries applications [3] [12] [25].With the extensive increase in usage of internet, there are lots of requirements for efficient and effective retrieving from large database images, which is a very challengeable task. Availability of efficient image recording de- vices such as image scanners, cameras, size of image collection increases. Image retrieval can be classified into following methods [10] [19].

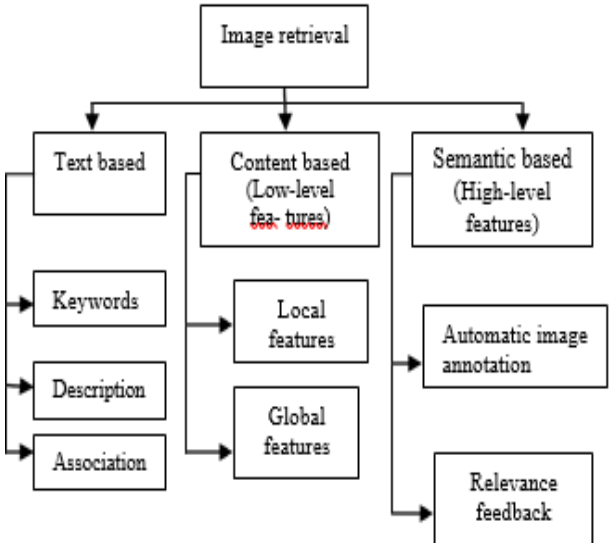


Fig. 2: Image Retrieval Classification.

a) Text based image retrieval (TBIR): Traditionally TBIR is also known as concept based image retrieval where searching is based on annotation of images. TBIR searches image database in terms of the similar text, tag, labels, keywords and system return images similar to the query string [10] .When images are well versed an- notated, text based methods are fast and robust which gives relia- ble results. But sometimes, text based approach fails to retrieve relevant information as manual annotations are not appropriate or else surrounding text may be improper to depicts the image. Hence these challenges of indexing, matching and retrieving mul- timedia data has been lead to development of different kind of techniques for retrieving images [6] [12] [27].

b) Content based image retrieval (CBIR): Another way of brows- ing and covering up the limitations that are imposed by TBIR is CBIR. Researchers have proposed term CBIR in 1992 and alter- nate name for CBIR given as "search engines as images". Some limitations which over comes by CBIR over TBIR as follows [2] [5] [26]:

- Image features that contain color, text, and structure data are extracted automatically.
- In addition to labels, tags, keywords that allocated with images, there is no need of field specialists.
- No language barrier or any ambiguity of languages as depiction of text images form is not necessary.
- Effect of manual error is minimized because of an au- tomated approach, which does not require any human in- tervention.
- Provides results that are more accurate.

3. Introduction of CBIR

CBIR plays a vital role in digital image processing techniques with the extensive increase of growing demands for multimedia fields in areas such as fashion and graphic design, crime preven- tion and traffic control which has forced application developers to manage search engines more efficiently. CBIR also termed as by query by image content (QBIC) and content based visual infor- mation retrieval (CBVIR) that acts as an application of computer vision [2] [4].It develops the search which images analyse actual

contents of image. It is process of getting a desired image from vast collection, on basis of features that constraints color, texture and shape which extracts automatically. It also exploits a more perceptual representation of an image that indulges more features properties such as color, text, shape and spatial relationships [24]. The following figure represents the pictorial image databases [11] [19] [22].

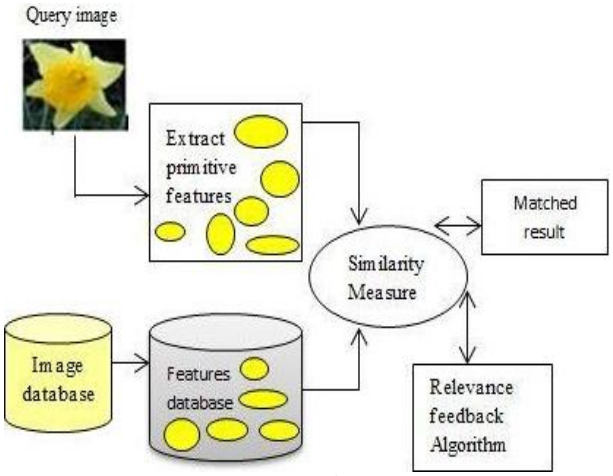


Fig. 3: Block Diagram for CBIR System.

A query image is delivered to system and then system changes it into its internal representation of features vectors. A similarity is matched in between feature vector of user's query image and among database images. CBIR method has three phases:

a) extraction of features
b) matching of features,
c) System de- signing retrieval.

SEARCHING OPTIONS FOR CBIR SYSTEM: From 1990's, CBIR has become an active research area both in commercial and research system for retrieving images. Most of them support following options [14] [23].

- Randomly browsing
- Search by example
- Navigation with customized image categories
- Search by sketch, text (also includes tags, labels, speech, keywords)

There is supply of rich set of search options, but in practical application fields involves, actual users still requires systematic ways among different search options mentioned.

4. Steps for CBIR

The proper organization of large amount of image generated database for CBIR system is important. Following consists of basic steps for CBIR system [16] [11] [21].

- Image acquisition: It is used to acquire digital images.
- Image database: It collects "n" number of images that depends on users range and choice.
- Image pre-processing: Images which describes its contents, processing is done which involves filtering, normalization, seg- mentation and object identification in order to extract features.
- Feature Extraction: It involves features such as shape, text and color and is classified into a low level and high level feature which describes content of images.
- Similarity Images: Images are to be searched in database to predict whether similar images are exists and also involves dis- tance method such as Euclidean distance, Canberra distance and many other.

- Retrieved images: Involves searching of matched images according to the previously maintained information to determine whether result is similar or has any closest features as that of que- ry images.

User interface: Involves display of results, ranking etc.

5. CBIR benchmark database

Until date various databases having different features has been used, to cover up wide range of applications in different domains, where CBIR is used. Hence approximately 17 databases are discussed/presented in this review work [9]. The databases includes colors, gray scale, texture, shape, medical, segmented images including different kinds of natural scenes, faces, buildings, retinal and many others. These images motivate to concentrate in different kinds of dimensions for improving retrieval process [17] [18].

Table 1: Different Types of Database

Types Of Database	Summary
CXRO Database	• Optical database
DRIVE Database	• Digital retinal images for vessel extrac- tion • Image size 768 x 584 in ".jpg" format • Segmentation of blood vessels in retinal image
VISTEX Database	• Vision texture database • Image size 786 x 512 • Consists of 165 compressed and 240 uncompressed image
Outer Database	• Provide high quality texture images • Texture image database • Image pixel size 1712 x 1368
Bro darts Database	• Mainly emphasized on texture classifi- • cation, supervised and unsupervised texture classification • Two components MBT(multiband bro darts texture database of 154 color im- ages) and CBT (color bro darts texture • database of 112 gray scale images)
BSDS 500 Database	• Berkeley segmentation database • Images are of different size in ". Seg" format • Emphasizes on image segmentation and boundary detection application
INRIA Database	• Created by Navneet Dalai (1800 images of persons) and by Hervey Jingo (1441 images of holidays) • Developed at AGENCE NATION- ALLE DE LA • CHE,INRIA ,France
NUSWIDE Database	• Image size 64 x 128 ,binary format • Based on real world web image data- base • Image are of different size in ".jpg" format • KNN algorithm implemented
CALTECH Database	• It is a face database developed at Cali- fornia institute of technology, USA • Images are of fixed size 304 x 312 in .tiff format
Olive Database	• Created by Audi Olive, USA • Consists of natural and urban scene col- or images in ".jpg" format
ETHZ Shape classes Database	• Created at ETH Zurich Switzerland • Images are of different size in ".png" format • Used for testing object class detection algorithms
Zu-BuD Data- base	• Zurich Buildings database for image based recognition • Provides information on good combina- tion of color and shape descriptors • Images are of different sizes in ".jpg" and ".png " format
MIRFLICKR	• Media lab Image retrieval Committee • Images in ".jpg" format
IRMA-10000 Database	• Consists of 10000 fully annotated radi- ographs • Evaluate only gray value images rather • than color histograms nor the MPEG7 scalable color descriptors
UW Database	• Created at University of Washington • Consists of vacation pictures of various • locations such as spring flowers, Barce- lona etc.
WANG Database	• Used for feature extraction

6. Evaluation methods

Parameters like effectiveness and efficiency, on which CBIR systems are evaluated. Effectiveness relates with the high accuracy and efficiency with the speed of retrieval. CBIR systems use sev- eral evaluation methods to extract and classify features. Following are some of the several evaluation measurements for CBIR system for CBIR systems [7] [16] [21].

Table 2: Different Types of Evaluation Method

Parameters	Description
Measurement Evaluation	
Euclidean Distance (Ed)	• Termed As Displacement Of Pixel From Nearest Back- • Ground
Chi Square Distance	• Ed Between Components Of Profiles, On Which Weighting • Is Defined
Weighted Euclidean Distance (W.E.D)	• Multiply Squared Differences By Corresponding Weights
Bhattacharya Dis- Tance	• Measures Similarity Between Two Discrete Or Continuous Probability Distributions • Popular Distance Of Similarity Between Two Gaussian Distri- Butions
Mahalanobis Distance	• Based On Correlations Be- Tween Variables • Used To Analyze Various Pat- Terns • Useful In Determining Simi- Larity Between Unknown Sample Set And Known One
Performance Measurement	
Precision	• In Binary Classification, It Is Analogous To Positive Predic- Tive Value
F-measure	• Recall and precision are evenly weighted • Weighted harmonic mean of precision and recall

7. Literature survey

The focus of this paper is to build the review model. The following table pinpoints the research contributions and gaps of study and key areas of research [14] [26].

Table 3: Review of Literature on Techniques of CBIR

Author	Title	Contributions	Analyzing Of Pro- Posed Contributions	Gaps
Faiq Bji , Mihas Mocan U (2017)	Connected Compo- Nents Objects Features For Cbir [28]	A New Algo- Rithm Which Is Region Of In- ter- Est (Roi) For Image Retrieval And Texture Features Using Glcm Is Used	System Has Highest Average Precision Which Is Up To 78% Accuracy	Shape Features Are Not Consid- ered
Raj- Kumar Jai, Pumit Kumar Jo- hari (2016)	An Im- Proved Approach Of Cbir Using Color Based Hsv Quantiza- Tion And Shape Based Edge De- Tection Algorithm [29]	Used Color And Shape Features To Analyze In- tegrating Hsv Color His- togram And Prewitt Edge Detection Tech- Nique	Similarity Measure- Ment Has Been Car- Ried Out, Which Is Manhattan Distance	Texture Features Are Not Focused
Nidhi Tripa- Thi, Pankaj Shar- Ma (2016)	A New Technique For Cbir With Con- Trast En- Hance- Ment Us- Ing Multi- Feature And Multi- Class Svm Classifica- Tion [30]	Uses Multi- Feature And Multi-Kernel Sup- port Vector Machine For Cbir	Gabor Fil- Ter, Dwt & In- verse Difference Moment (Idm) For Textures, Skewness And Kurto- Sis' For Shape And Variance, Root Mean Square For Color Fea- Tures Are Extracted	Focuses Only On Image Files
Sandh ya R Shinde et.al (2015)	Experi- ments based on content based	Feature vectors are formed using color features and also to classify	Experiment outcomes is a maximum accuracy of 81.25%	
Prince Shakta watt and V K Go- vindan (2015)	Novel scheme for image retrieval using combina- tion of color- texture features [32]	CBIR technol- ogy based on color and tex- ture features are considered by applying Gabor filter and discrete cosine trans- form coeffi- cients of image	Manhattan distance for measure- ment of similarity is used	Shape and texture features we're not
K. Hard- est al. (2014)	Well or- ganized con- tent based image re- trieval system in RGB Col- or histo- gram, Tamura Texture and Gabor fea- ture [33]	A novel algo- rithm for con- tent based im- age retrieval based on RGB color Histo- gram, Tamura features and Gabor filter are considered	using Ga- bor features is 81.7% accuracy in CBIR Parameters like recall, and preci- sion value and accura- cy are also con- sidered	
Sadat al.(201 3)	Visual feature extrac- tion for con- tent based image retrieval [34]	color histo- gram method for image re- trieval is con- sidered	Includes design of data- base and also specific image fea- tures re- trieval	

			Measure- ment of Euclidian distance has been considered	
Amanbir Sandhog (2012)	Content based image retrieval using texture, color and shape for image analysis [35]	Presents technique for CBIR by using three features such as shape ,texture and color	GLCM for texture ,color feature for color histogram and factors like area, eccentricity and filled area for shape are proposed and also its different combina- tion	
Yu-Chum Wang (2012)	A novel content based image retrieval system using K-means/KNN with feature extraction [36]	Uses K-means/KNN clustering along with extraction of features for novel image retrieval	To develop the CBIR system ,it Combines segmentation and feature extraction module, k-means clustering, grid module and neighborhood module	Optimizations are not proper for system architecture.
Rahul Mehta, Sanjeev Sharma (2011)	COLOUR-TEXTURE based image retrieval system [37]	Uses color and texture features of images by applying conventional color histograms (CCH) , Quadratic Distance Metric (QDM) and Pyramid structure wavelet transform model (PSWTM)	Similarity Measurement of Euclidean distance and quadratic distance metric is analyzed	Shape or edge features we're not taken into consideration
Zhi- chun Huang et al (2010)	Content based image retrieval using color moment and Gabor texture feature [38]	Uses texture and color features for ex- traction with the help of techniques such as color moment and Gabor de- scripters	Proposes combina- tions of color mo- ments of HSV color space and texture Gabor filter descriptors Euclidean distance is 63.6% accuracy for distance measurement	It has very low accuracy

In this survey paper, an exhaustive literature review of CBIR has been done from its origin to the new age. This paper presents different proposed contributions and techniques in CBIR used by researchers. This elucidates significance of content-based image retrieval systems [14] [26].

8. Conclusion

Because of large image collections in the fields like biomedical, military, education, remote sensing etc., the manual annotation approaches used in past years were very difficult to carry out. To overcome these challenges, CBIR approach was introduced. CBIR generally process the images to improve their pictorial representa- tion and extract useful and relevant information. In CBIR, instead of carrying out manual annotations using textual keywords, imag- es are indexed using their own visual contents such as color, tex- ture and shape. This paper presents detailed review of different Datasets and evaluation parameters in terms of both measurement and performance. The literature survey carried out discusses the various kinds of CBIR techniques, their feature, gaps and contributions.

References

[1] Nitish Barya, Himanshu Jaiswal (2015) "Survey on Content based Image Retrieval to Deal with Rapid Growth of Digital Imag- es", *International Journal of Computer Applications.*

[2] Prof.Vikram M Kakade1, Ishwar A. Keche (2017), "Content Based Image Retrieval (CBIR) Technique", *International Journal of En- gingering and Computer Science.*

[3] Y. Ruiz and T. Huang (*1999*) "Image retrieval: current techniques, promising directions and open issues," *J. Visual Commun. Image Representation.*

[4] Bansal et al. (*2014)* "Content Based Image Retrieval using SVM", International Journal of Advanced Research in Computer Science and Software Engineering.

[5] Retender Datta Jia James Z. Wang (2009) "Content-Based Image Retrieval - Approaches and Trends of the New Age ".

[6] Parul Preet, Kulvinder Singh Mann(2013),"An Approach of image retrieval using content based retrieval system", *International Jour- nal of Advanced Research in Computer Science and Software Engineering.*

[7] Performance Evaluation in Content-Based Image Retrieval: Over- view and Proposals MULLER, Henning, et al, *http://archive- ouverte.unige.ch.*

[8] Kenneth R. Castle man (*1996)*, "Digital Image Processing".

[9] A. Blazer (*1997)*, "Database Techniques for Pictorial Applications", *Lecture Notes in Computer Science, Vol.81, Springer Verlag GmbH.*

[10] Devbrat Aria, Jaimala Jha (Asst. Prof), Arya et al., (2016) "Review on Content Based Image Retrieval Using Feature Extraction", *In- ternational Journal of Advanced Research in Computer Science and Software Engineering.*

[11] Mohd. Danish, Ritika Rawat Ratika Sharma (2013), "A Survey: Content Based Image Retrieval Based On Color, Texture, Shape & Neuro Fuzzy" *Journal of Engineering Research and Applications.*

[12] Datta, Joshi Li, Z. Wang, Acm (*2008)*, "Image Retrieval: Ideas, Influences, and Trends of the New Age" Computing Surveys.

[13] NidhiSinghai, Prof. Shishir K. Shandilya (*2010*), "A Survey On: Content Based Image Retrieval Systems", *International Journal of Computer Applications.*

[14] Arun Singh Chouhan, Prabhleen Kaur, Saroj Bala, (*2016),*"Literature Survey on Latest trends in Content Based Image Retrieval (CBIR)",*International Journal of Computer Trends and Technology (IJCTT).*

[15] Shaila S. Tambe, Prof. B. S. Borkar (2014),"Image Retrieval System*", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET).*

[16] Miss. Mrunali D. Pawar , Miss. Shraddha N. Shinde, Mr. Ramdas P. Bag wade (*2016)*, "Content Based ImageRetrieval:A Study", *IJARCCE.*

[17] Desolaters, Daniel Keysers and Hermann Ney,Daniel.keysers@dfki.de, "Features for image retrieval, Data- base for CBIR", *computer science department, RwtH Aachen uni- versity, germany {deselears,ney)@cs.rwth-aachen.de.*

[18] Smeulders.A.W.M, Worring.M, Santini.S, Gupta.A, and R. Jain (*2000)*, "Content- based image retrieval at the end of the early years," *IEEE.*

[19] Christophe (2012), J. Next Generation Search Engine: Advanced Models for Information Retrieval. Hershey, PA: IGI Global. Retrieved, *http://www.igiglobal.com/book/next-generation-searchengines/59723.*

[20] R.Malini and C.Vasanthanayaki (2013) "An Enhanced Content Based Image Retrieval System".

[21] Manimala Singh (2012), "Content Based Image Retrieval using Colour and Texture" *Signal & Image Processing: An International Journal (SIPIJ).*

[22] M.Rehman, M.Iqbal, M.Sharif and M.Raza (2012), "Content Based Image Retrieval: Survey".

[23] Mussarat Yasmin *(2013)*, "Use of Low Level Features for Content Based Image Retrieval: Survey".

[24] Dr. Sanjay Silakari, Dr. Mahesh Motawani and Manish Maherswari (*2009*), "Color image clustering using block truncation coding algorithm",*IJCSI.*

[25] Kannan, Dr.V.Mohan, Dr.N.Anhazhagan (2010), "Image mining techniques",*International conference and computational intelligence and computing research, IEEE.*

[26] Jagsir kaurl, manoj kumar (2015),"Review paper on content based image retrieval for digital images" *International journal of research in computer applications and robotics.*

[27] Dong-Gyu Sim; Hae- Kwang Kim: Dae-II Oh (Kannan, Dr.V.Mohan, Dr.N.Anhazhagan (2010), "Image mining tech- niques", *IEEE ,International conference and computational intelli- gence and computing research.*

[28] Faiq Baji, Mihas Mocanu (*2017),*" connected components objects features for CBIR", *IEEE.*

[29] Rajkumar jai, Pumit Kumar Johari (*2016)*, "An improved approach of CBIR using color based HSV Quantization and shape based edge detection algorithm" ,*RTEICT.*

[30] Nidhi tripathi, Pankaj Sharma (*2016)*, "A new technique for CBIR with Contrast enhancement using Multi-feature and multi-class SVM classification", *IEEE.*

[31] Sandhya R Shinde et.al (*2015*)," Experiments based on content based image classification using color feature extraction", *IEEE.*

[32] Prince Shakta watt and V K Govindan (2015), "Novel scheme for image retrieval using combination of color-texture features", *(IJCTT).*

[33] K. Hardest al.(2014) "Well organised content based image retrieval system in RGB Colour histogram, Tamura Texture and Gabor fea- ture", *International journal of advanced research in computer and communication engineering.*

[34] Sadat al. (*2013)*, "Visual feature extraction for content based image retrieval", *IJASR.*

[35] Amanbir Sandhog, Aarti kochhar (*2012)*, "Content based image retrieval using texture, colour and shape for image analysis", *Inter- national journal of computers & technology.*

[36] Yu-Chum Wang, Ray-I Chang, Shu -Yu Lin. Chi- wen fan (2012), "A novel content based image retrieval system using K-means/KNN with feature extraction".

[37] Rahul Mehta, sanjeev Sharma (2011), "COLOUR-TEXTURE based image retrieval system", *International journal of computer applications.*

[38] Zhi-chun Huang et al. (2010) "Content based image retrieval using color moment and Gabor texture feature", *proceedings of the ninth international conference on machine learning and cybernetics, Qingdao.*