

Enhanced Semantic Image Retrieval using Feature Extraction and KNN Techniques

Nitesh Rastogi

M.Tech Student

Computer Science and Engineering Department
IET-Alwar, India

Deepak Chaudhary

Assistant professor

Computer Science and Engineering Department
IET, Alwar, India

ABSTRACT

In addition of that, the technique required some additional techniques to correct the retrieval process such as user feedback, these methods consumes additional time of search. Thus a new technique with hybrid concept is proposed for improving the content based image search. The proposed technique includes the technique to train the system using the image features and text for annotation of image. For identifying the images more accurately the text and image features are used. Finally to retrieve the data (image) using user query (image or text) a KNN algorithm is implemented with it. The implementation of the proposed model is performed using visual studio technology and their performance in terms of time and space complexity is estimated. In addition of that the performance in terms of accuracy and error rate is also provided for demonstrating the relevancy of image search.

General Terms

Your general terms must be any term which can be used for general classification of the submitted material such as Pattern Recognition, Security, Algorithms et. al.

Keywords

CBSIR, Image Retrieval, Tag based, feature, computation, system modeling, KNN, LBP.

1. INTRODUCTION

Content Based Semantic Image Retrieval (CBSIR) [1] may be a theme that searches pictures from an outsized info by means that of visual contents, as per the interest of user. Since Nineteen Nineties, it's been fast growing analysis space. Moreover, within the past years, the researchers have created notable results. In late Seventies, the sphere got its foundation in a very conference on info Technical for Pictorial Applications that was control in Florence [2]. This created the researchers to be attracted towards the sphere. At early stage, the technique was supported matter annotations of pictures i.e. pictures were 1st annotated with text then searched employing a text-based theme from typical info systems [3]. This theme was very little easy and will typically fail to deliver precise results. Moreover, it's challenging to mechanically generate annotations for every image, thus manual annotation was followed that may be a clumsy and sophisticated task and a lot of high-ticket if we've got bulk databases.

Further in early Nineteen Nineties, with ascent in web and digital image sensors, usage and production of pictures accrued that more created a desire of CBSIR systems. Since then, analysis on content-based image retrieval has developed apace [4].

This article lists out some essential works and contributions place within the direction of CBSIR and presents a quick survey concerning completely different techniques.

Additionally of that the work is extended and obtained a brand new image retrieval model. The careful discussion of the image retrieval model is given in more sections.

2. BACKGROUND

This section provides the fundamental understanding of the content based image retrieval and their functioning.

2.1 Image Retrieval

Image retrieval systems with their basic elements square measure simulated exploitation figure one. that may be known as looking, browsing, and retrieving pictures from huge databases consisting of digital pictures. Though ancient and standard techniques of retrieving pictures employs adding information i.e. captioning keywords thus on perform annotation of words. but image search are often delineate by dedicated technique of search that is routinely accustomed realize pictures. For that, user provides the question image and also the system returns the image like that of question image [5].

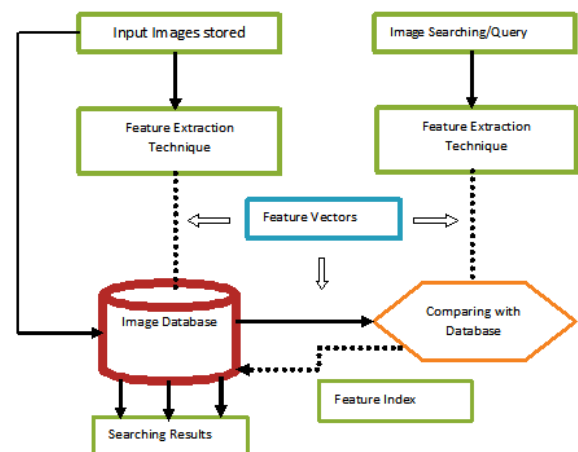


Figure 1 Semantic Image Retrieval System

The image Retrieval has been adopted by most of the main search engines, as well as Google, Yahoo!, Bing, etc. encompassing texts is taken into account by an oversized range of search engines and image names to index them, since there are unit solely 2 main places wherever text are often placed. 1st is title and therefore the alternative is caption or tags that area unit projected and enforced exploitation net two. concepts. Sometimes users build question within the text format for search contents over any computer program.

2.2 Image Retrieval Techniques

With the increasing use of net and convenience of various image capturing devices like camera, bulk quantity of pictures area unit being created each day in numerous areas as well as

remote sensing, fashion, crime bar, publishing, medicine, design, etc. Therefore, development of economical and effective methodologies to manage massive image databases for retrieval is desperately required. 3 strategies of image retrieval area unit used i.e. text-based methodology, content-based methodology and hybrid methodology. This section explains in details every methodology. Image retrieval system are often classified in 2 key streams as given in figure 2:

- Text based Image retrieval system
- Content based Image retrieval system

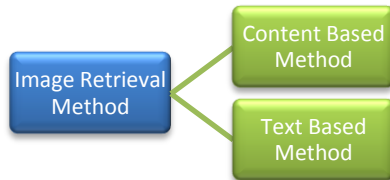


Figure 2 Image Retrieval Approaches

2.3 Text-based image retrieval

Text based Image Retrieval (TBIR) is presently employed in nearly all general net image retrieval systems these days. This approach uses the text allied with a picture to seek out out what the image contains. This text are often text encompassing the image, the image's name, a link resulting in the image, AN annotation to the image, or the other piece of text which will be related to the image. samples of systems victimization this approach are Google, Yahoo Image Search engines. These search engines having indexed over one billion pictures. though when being quick and strong, these search engines typically fail to retrieve relevant images.

2.4 Content - based semantic image retrieval

Content based semantic Image Retrieval may be a set of techniques for retrieving semantically-relevant pictures from a picture information supported automatically-derived image options [6]. CBSIR seeks to avoid the employment of matter question inputs. Rather, it retrieves pictures based on their visual similarity to a user-supplied question image or user-specified image options.

The core objective of CBSIR is potency whereas retrieval and categorization of image, thereby reducing the necessity for human intervention within the categorization method [7]. the pc should be capable of retrieving images from a information with none human assumption on specific domain (such as texture vs. non texture). one among the key tasks for CBSIR systems is similarity comparison, extracting feature of each image supported its pixel values and process rules for comparison images. These options become the image illustration for the activity of similarity with alternative pictures within the information. pictures area unit compared by scheming the distinction of its feature elements to alternative image descriptors. These image descriptors area unit internationally obtained on the premise of color histograms for color features; global texture data on coarseness, contrast, and direction; and form options concerning the curvature, moment's invariants, disk shape, and eccentricity

Majority of existing CBSIR systems takes every image as a complete; though, one image might comprise multiple regions/objects with altogether totally different semantic that means. a normal user can typically have an interest in barely one specific region of the question image not the entire. So, rather than viewing a picture as whole, viewing it as a group of regions is a lot of cheap. Most Image Retrieval systems embody color, texture, form and spatial layout. Such feature lose their effectiveness for CIBR if they get retrieved from a full image, as a result of they suffer from distinct backgrounds, overlaps, occlusion and cluttering in numerous pictures and that they don't possess satisfactory capability to capture necessary properties of objects, as a consequence the first focus of most well liked approaches is being shifted from the worldwide content description of pictures into the native content description by regions or maybe the objects in pictures. RBIR may be a showing to be a possible extension of the classical CBSIR: instead of exploitation global options over the complete content, RBIR systems partition a picture into variety of undiversified regions and extract native options for every region then options of regions area unit wont to represent and index pictures in RBIR. For RBIR, The user puts in a very question object by choosing a neighborhood of image so the resultant similarity live is computed between options of region within the query and a group of features of divided regions in features information and therefore the system returns a hierarchal list of pictures that contain identical object. The content-based approaches are often summarized as follows:

1. Computer vision and image process techniques are utilized to extract content features from the image.
2. Images are unit depicted as collections of their distinguished options. For a given feature, AN applicable illustration of the feature and a notion of similarity area unit determined.
3. Image retrieval is performed supported computing similarity or difference within the feature space, and results hierarchal supported the similarity measure.

3. RELATED WORK

This area provides the afresh fabricated contributions and efforts in the administration of Content based Image retrieval technique.

In this day and age, Content based Image retrieval (CBSIR) has become a bastion of Image retrieval systems. In adjustment to get added yields, a affection of appliance acknowledgment was added to CBSIR so that added absolute after-effects can be acquired by demography user's acknowledgment into considering. Though, accepted feedbacks are getting offered by absolute systems to access aesthetic results, decidedly in systems with ample calibration Image database. Unfortunately this is absurd in absolute applications. In 2011, Ja-Hwung proposed a atypical approach, Navigation-Pattern-based Appliance Acknowledgment (NPRF), to cope up with ample calibration Image database thereby convalescent ability and capability of CBSIR. With use of aeronautics patterns extracted from user concern log, iterations are reduced. As consequences, ability is increased. This proposed algorithm utilizes the extracted aeronautics patterns and three varieties of concern clarification strategies called Concern Point Movement (QPM), Concern Reweighting (QR), and Concern Expansion (QEX), to besiege the seek amplitude appear the blueprint of user. By appliance NPRF method, top superior of Image retrieval on RF can be able in a slight amount of feedbacks

[8]. In addition, arrangement actualization such as the anarchy based on the gray akin co-occurrence cast and the bend histograms of an Image are as well taken into consideration. Moreover, the IGA is active to arch the gap amid retrieved after-effects and user's apprehension and advice him to get the Image a lot of acceptable user's necessity. Experimental after-effects acutely authenticate achievability of the adjustment [9].

CBSIR systems are accepting their easily in so abounding altered fields and domains. A CBSIR based Image retrieval arrangement was developed for medical applications by Ramesh in 2012. In this appliance the CBSIR is acclimated for analytic medical Image collections in ample scale. The arrangement should be able of retrieving images from the aforementioned ache chic as the accommodating is suffering. The abstraction presented by Ramesh examines the abstraction of histogram from a medical image. That Image is resample and classified by appliance Sequential basal enhancement address for assorted allotment of dataset produced consecutive to the abstraction [10].

A amount of variants of appliance acknowledgment are accessible these days. In adjustment to acquisition out a lot of optimum methodology, an agreement was presented by Ghanshyam in 2012 in which assorted new applications of genetic algorithm to advice retrieval, a lot of of them accompanying to appliance feedback. Efficient accumulator and retrieval of Image abstracts is consistently binding in adjustment to accomplish assigned tasks and authoritative decisions. The arrangement presented a new framework for Image retrieval with two types of appliance acknowledgment called absolute acknowledgment and absolute feedback. It utilizes Interactive genetic Algorithm to ascertain a aggregate of descriptors that bigger characterizes the charge of user in agreement of Image affinity [11].

P. Jayaprabha has presented a abstraction in agnate domain, in 2013. They presented a top akin semantic retrieval process, in which the seek engine is activated for retrieving a ample amount of images appliance a accustomed argument based query. In a low akin Image retrieval process, agnate Image seek action is provided for user to ample in the ascribe concern for Image affinity characterizations. The internet anarchy and agenda technologies accept answerable a charge to accept a arrangement to adapt abundantly accessible agenda Image for effortless analysis and retrieval. The technique involves ample areas, i.e. Image segmentation, Image affection extraction, representation, mapping of actualization to semantics, accumulator and indexing, Image affinity ambit altitude and retrieval authoritative CBSIR arrangement development a acute assignment [12].

Local Binary Patterns are abundantly acclimated for arrangement classification, accept projected a adjustment for facial announcement acceptance appliance Local Binary Patterns (LBP) as actualization and Artificial Neural Network [13]. Six accepted expressions, i.e. anger, disgust, fear, happy, sad, and abruptness forth with seventh one neutral, are accustomed by the Generalized Feed-forward Neural Network. Levenberg - Marquart (LM) nonlinear enhancement algorithm is acclimated to alternation and analysis the Neural Network. They accept accomplished 93.3 % allocation amount with testing achievement 0.0573.

A. R. Ardakany, in 2012 provided allocation appliance a simple affection abstraction which performs geometric and actualization appearance at the aforementioned time. This affection abstraction is agitated out by accretion the acquired

in all pixels of face images and again amalgam a histogram based on edges magnitudes and directions. The abstracts are bright affirmation that presented adjustment is absolutely aggressive with 95.67% accurateness on FERET database [14]. A typical access which is assiduity the classical archetypal with all-encompassing knowledge-based, WordNet was proposed by Yohan in 2005. This access attempts to trim extraneous keywords by the acceptance of Word Net. To assort extraneous keywords, analysis on assorted semantic similarities of keywords and to agglutinate the outcomes of all these measures calm to accomplish a final accommodation with the advice of Dempster-Shafer affirmation combination. Assorted models accept implemented by them to accessory beheld tokens with keywords based on knowledge-based, Word Net and evaluated achievement appliance precision, and anamnesis appliance criterion dataset. The after-effects appearance that by assiduity knowledge-based with classical archetypal they can advance attention of comment by demography out extraneous keywords

4. PROPOSED WORK

The proposed system design is given using figure 3 which is given in two main phases. First for train the system or storing the data into the database and second is used for accepting the user query and producing the search results. Therefore the entire system is described in three major modules first feature extraction, Query interfacing and finally the results listing.

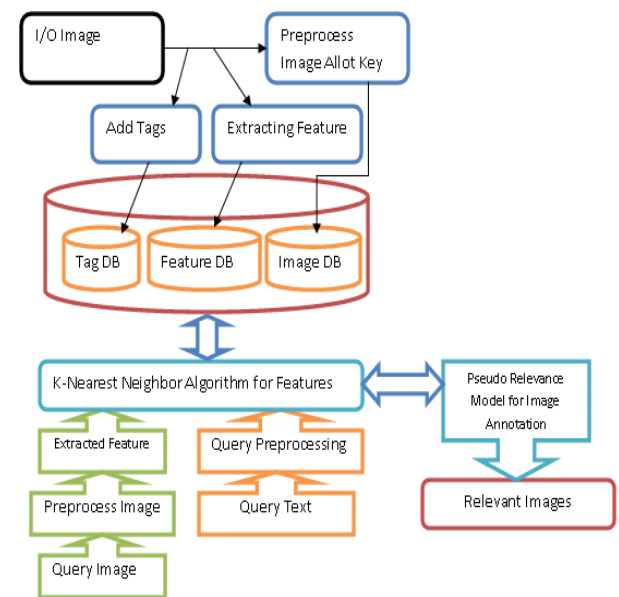


Figure 3 proposed system

4.1 Feature extraction

The content based images are retrieved by their image properties such as image objects edges, colour distributions and the image textures. Therefore all the tree image features are computed and normalized first which is stored in a database table for image feature representation. At the same time the image are also tagged with some kind of text which indicates the objects available in the input image during training phase. These tags are preserved separately in a table. But in order to recognize the image a key is assigned which is also preserved with the databases.

4.2 Query interface

As database is filled with the image contents the training session of the presented model is completed. Now for accepting the user query the system can accept the text query and image query also through the individual user interface.

4.3 Search outcomes

The user produced query is supplied to the KNN algorithm where the KNN having to inputs first the user query tokens and second the database of images, image features and the tag associated with images. Thus by finding the distance between the user query input and the data base scenarios the nearest distance images and their objects are recognized.

The basic features of the proposed work model are explained in this section. In addition of that their modular distribution for implementation of the CBSIR model is also explained. The next section discussed the conclusion and the future extension of the presented work.

4.4 Implementation

4.4.1 Grid Colour Moment

Colour feature is one of the most widely used features in low level feature. Compared with shape feature and texture feature, colour feature shows better stability and is more insensitive to the rotation and zoom of image. Colour not only adds beauty to objects but also more information, which is used as powerful tool in content-based image retrieval. In colour indexing, given a query image, the goal is to retrieve all the images whose color and texture compositions are similar to those of query image. In color image retrieval there are various methods, but here we will discuss some prominent methods.

The feature vector we will use is called "Grid-based Colour Moment". Here is how to compute this feature vector for a given image:

- Convert the image from RGB for HSV colour space
- Uniformly divide the image into 3x3 blocks
- For each of these nine blocks
- Compute its mean colour (H/S/V)

$$x' = \frac{1}{N} \sum_{i=1}^N x_i$$

Where N is the number of pixels within each block, x_i is the pixel intensity in H/S/V channels.

- Compute its variance (H/S/V)

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - x')^2$$

- Compute its skewness (H/S/V)

$$\gamma = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - x')^3}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - x')^2 \right)^{3/2}}$$

Each block will have 3+3+3=9 features, and thus the entire image will have 9x9=81 features. Before we use SVM to train the classifier, we first need to normalize the 81 features to be within the same range, in order to achieve good numerical behaviour. To do the normalization, for each of the 81 features:

Compute the mean and standard deviation from the training dataset

$$\mu = \frac{1}{M} \sum_{i=1}^M f_i$$

$$\sigma = \sqrt{\frac{1}{M} \sum_{i=1}^M (f_i - \mu)^2}$$

Where M is the number of images in the training dataset, and f_i is the feature of the i-th training sample.

Perform the "whitening" transform for all the data (including both the training data and the testing data), and get the normalized feature value:

$$f'_i = \frac{f_i - \mu}{\sigma}$$

The computed color features of images are preserved in a feature database. And in next step the edge detection technique is applied on image for recovering the edges of the image therefore the canny edge detection technique is applied on image, the canny edge detection technique can be described as:

4.4.2 Canny Edge Detection

The purpose of edge detection in general is to significantly reduce the amount of data in an image, while preserving the structural properties to be used for further image processing. Several algorithms exists, and this worksheet focuses on a particular one developed by John F. Canny (JFC) in 1986.

The algorithm runs in 5 separate steps:

Smoothing: Blurring of the image to remove noise.

- Finding gradients: The edges should be marked where the gradients of the image has large magnitudes.
- Non-maximum suppression: Only local maxima should be marked as edges.
- Double thresholding: Potential edges are determined by thresholding.
- Edge tracking by hysteresis: Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

4.4.3 Local binary pattern

Given a pixel in the image, an LBP [14] code is computed by comparing it with its neighbours:

$$LBP_{p,R} = \sum_{p=0}^{P-1} s(g_p - g_e) 2^p$$

$$s(x) = \begin{cases} 0 & x \geq 0 \\ 1 & x < 0 \end{cases}$$

Where g_e the gray value of the central pixel is, g_p is the value of its neighbors, P is the total number of involved neighbors and R is the radius of the neighborhood. Suppose the coordinate of is (0, 0), then the coordinates of g_p are

$$\left(R \cos\left(\frac{2\pi p}{P}\right), R \sin\left(\frac{2\pi p}{P}\right) \right)$$

The gray values of neighbours that are not in the image grids can be estimated by interpolation. Suppose the image is of size $I \times J$. After the LBP pattern of each pixel is identified, a histogram is built to represent the texture image:

$$H(k) = \sum_{i=1}^I \sum_{j=1}^J f(\text{LBP}_{p,r}(i, j), k), k \in [0, k]$$

$$f(x, y) = \begin{cases} 1 & x = y \\ 0 & \text{otherwise} \end{cases}$$

where K is the maximal LBP pattern value. The U value of an LBP pattern is defined as the number of spatial transitions (bitwise 0/1 changes) in that pattern

$$U(\text{LBP}_{p,r}) = |s(g_{p-1} - g_e) - s(g_0 - g_e)| + \sum_{p=1}^{p-1} |s(g_p - g_e) - s(g_{p-1} - g_e)|$$

The uniform LBP patterns refer to the patterns which have limited transition or discontinuities ($U \leq 2$) in the circular binary presentation. In practice, the mapping from $\text{LBP}_{p,r}$ to $\text{LBP}_{p,r}^{u2}$ which has $P \times (P-1) + 3$ distinct output values, is implemented with a lookup table of 2^P elements. To achieve rotation invariance, a locally rotation invariant pattern could be defined as:

$$\text{LBP}_{p,r}^{\text{riu2}} = \begin{cases} \sum_{p=0}^{p-1} s(g_p - g_e) \text{ if } U(\text{LBP}_{p,r}) \leq 2 \\ P + 1 \text{ otherwise} \end{cases}$$

The mapping from $\text{LBP}_{p,r}$ to $\text{LBP}_{p,r}^{u2}$ which has $P+2$ distinct output values.

Finally the tag on the image is also included and all three features and their tags are saved on the database. This preserved features and tags of the image objects are used for search the data according to the user supplied query.

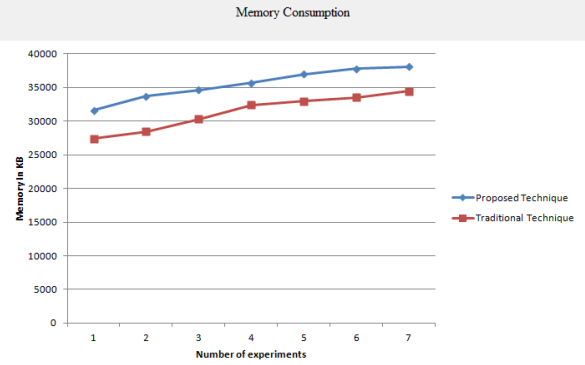
4.5 Results and Graph

The proposed model is compared in this chapter over different performance parameters. Thus the similar parameters are used as stated in the subjected in the proposed system.

Table 1. Result table in Database

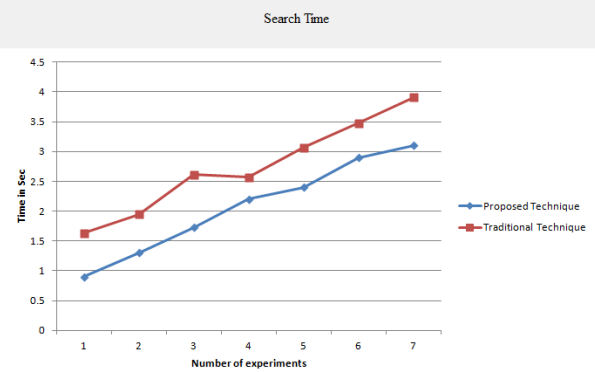
accuracy	errorrate	memoryused	timetaken
50	50	33087.6878	1404.0025
100	0	33087.6878	15.6001
100	0	43917.1122	47.0253
0	100	48057.06927	0
100	0	48057.06927	15.551
0	100	48057.06927	0

RESULT ANALYSIS



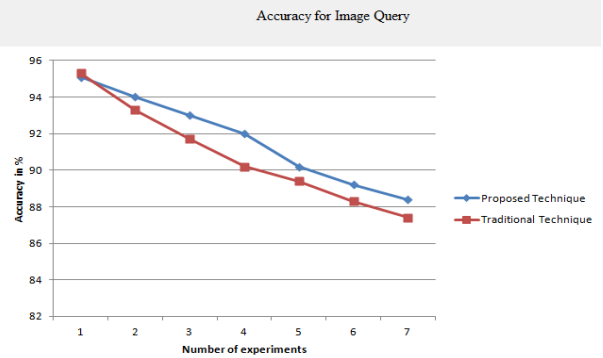
Graph 1: Memory Consumption

RESULT ANALYSIS



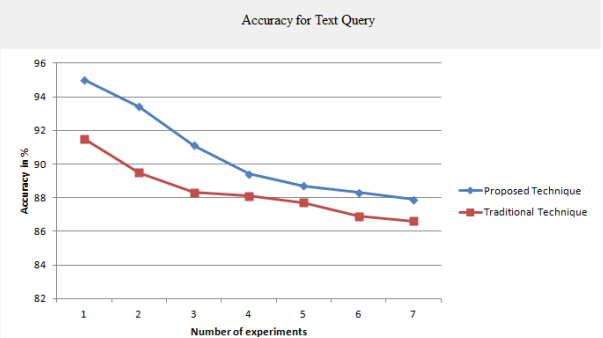
Graph 2: Search Time

RESULT ANALYSIS

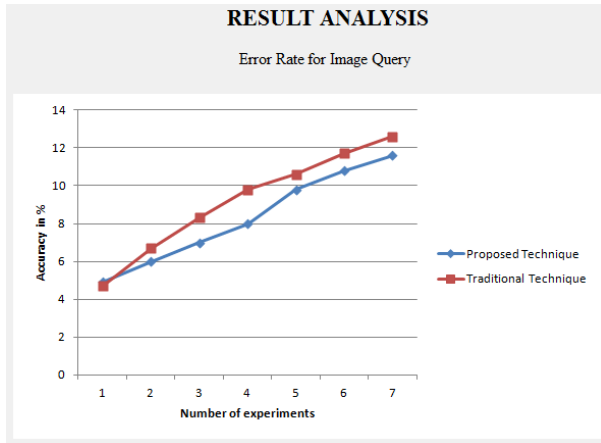


Graph 3: Accuracy for Image Query

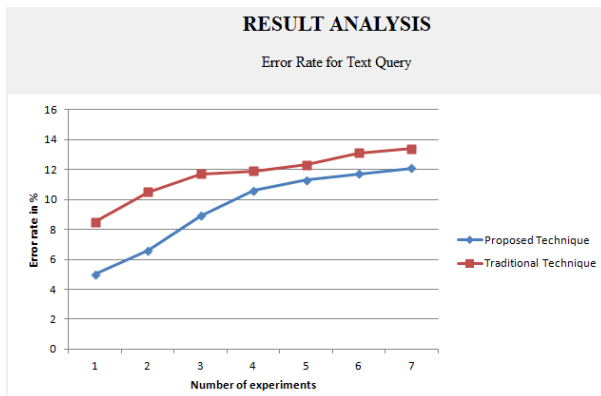
RESULT ANALYSIS



Graph 4: Accuracy for Text Query



Graph 5: Error rate for Image Query



Graph 6: Error rate for Text Query

5. CONCLUSIONS

In this age need of efficient and accurate systems are increases due to uneven increasing demand of image contents. Thus in this paper first a survey on Content based Semantic image retrieval (CBSIR) is performed in addition of that the different approaches developed in recent years are also discussed. During investigation that is found that there are a number of techniques that have been introduced so far but most of them are not much accurate or sometimes lake in performance. This paper has portrayed some of major contributions made in this field and discusses the fundamentals. Additionally a new model is also suggested for future implementation.

6. ACKNOWLEDGMENTS

My thanks to Mr. Deepak Chaudhary (Guide) who have contributed towards development of this research work.

7. REFERENCES

- [1] Datta, Ritendra, Jia Li, and James Z. Wang. "Content-based image retrieval: approaches and trends of the new age." Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval. ACM, 2005.
- [2] Khokher, Amandeep, and Rajneesh Talwar. "Content-based Image Retrieval: Feature Extraction Techniques and Applications." International Conference on Recent

Advances and Future Trends in Information Technology (iRAFIT2012). 2012.

- [3] Paton, Norman W., and Oscar Díaz. "Active database systems." ACM Computing Surveys (CSUR) 31.1 (1999): 63-103.
- [4] Madugunki, Meenakshi, et al. "Comparison of different CBSIR techniques." Electronics Computer Technology (ICECT), 2011 3rd International Conference on. Vol. 4. IEEE, 2011.
- [5] Rui, Yong, et al. "Relevance feedback: a power tool for interactive content-based image retrieval." Circuits and Systems for Video Technology, IEEE Transactions on 8.5 (1998): 644-655.
- [6] Murthy, V. S. V. S., et al. "Content based image retrieval using Hierarchical and K-means clustering techniques." International Journal of Engineering Science and Technology 2.3 (2010): 209-212.
- [7] Siorpaes, Katharina, and Elena Simperl. "Human intelligence in the process of semantic content creation." World Wide Web 13.1-2 (2010): 33-59.
- [8] Su, Ja-Hwung, et al. "Efficient relevance feedback for content-based image retrieval by mining user navigation patterns." Knowledge and Data Engineering, IEEE Transactions on 23.3 (2011): 360-372.
- [9] Lai, Chih-Chin, and Ying-Chuan Chen. "A user-oriented image retrieval system based on interactive genetic algorithm." Instrumentation and Measurement, IEEE Transactions on 60.10 (2011): 3318-3325.
- [10] Ramesh BabuDurai, C., V. Balaji, and V. Duraisamy. "Improved content based image retrieval using SMO and SVM classification technique." European Journal of Scientific Research 69.4 (2012): 560-564.
- [11] Raghuwanshi, Ghanshyam, Nishchol Mishra, and Sanjeev Sharma. "Content based image retrieval using implicit and explicit feedback with interactive genetic algorithm." International Journal of Computer Applications 43.16 (2012): 8-14.
- [12] Jayaprabha, P., and RmSomasundaram. "Content Based Image Retrieval Methods Using Self Supporting Retrieval Map Algorithm." IJCSNS 13.1 (2013): 141.
- [13] Moore, S., and R. Bowden. "Local binary patterns for multi-view facial expression recognition." Computer Vision and Image Understanding 115.4 (2011): 541-558.
- [14] Ardakany, Abbas Roayaei, and A. M. Joula. "Gender recognition based on edge histogram." International Journal of Computer Theory and Engineering 4.2 (2012): 127.
- [15] Jin, Yohan, et al. "Image annotations by combining multiple evidence & wordnet." Proceedings of the 13th annual ACM international conference on Multimedia. ACM, 2005.