A Novel Method for Satellite Image Retrieval using Semantic Mining and Hashing

Linda Mary John
Research Scholar
Thakur College of Engineering and Technology
Computer Department

Kiran Ashok Bhandari
Associate Professor
Thakur College of Engineering and Technology
Computer Department

ABSTRACT
Satellite images play an important role for collection of geographical information. However, the use of such images is limited to a greater extent due to its retrieval complexities. Traditional methods using text have also failed to yield desired and time-saving results. Therefore, Content based image retrieval using high semantic features has been developed to overcome problems related to text retrieval and challenges associated with semantic gap. Keeping in consideration the importance of semantic features for CBIR, the proposed method introduces a novel method using Semantic mining and Hashing for fast retrieval and attainment of accurate results.

General Terms
Remote sensing, Satellites, Data sets, Content based image retrieval (CBIR) and Algorithms.

Keywords
Semantic Mining, Hashing, Hash codes, Semantic gap, precision, recall, feature extraction and similarity measures

1. INTRODUCTION
Due to enormous amount of satellite images obtained from various sources, retrieval of relevant information from these images has become a necessity. The retrieval of the information also depends on the requirements of the user. Therefore the retrieval system needs to be specific-content. An image retrieval system is a computer system used to browse, search and retrieve images from a huge collection of digital images. Image retrieval requires a specialized data search to find images. To search images user may give input query such as keyword, image file/link and retrieve results based on similar features.

Numerous retrieval systems have been developed with the recent development in various image processing technologies. The two techniques used for search and retrieval of images is Text-based image retrieval followed by Content-based image retrieval. Text Based techniques can be traced back to 1970s.In this approach keywords are used to search and retrieve images from the database. However, this approach is considered as invalid because of different keywords by different users for annotations [1].Therefore text based retrieval failed to give accurate results from complex images .Hence a new field known as Content-based image retrieval (CBIR) emerged in the year 1990 to overcome the disadvantages mentioned above. CBIR is the process to retrieve images from a database of digital images based on the visual content of the images. It retrieves images according to the properties of colors, textures or shapes [2].

Color is the most dominant and distinguishing visual feature used in CBIR and is invariant to image size and orientations. The color feature used in common is the color histogram followed by color correlogram and so on. Various applications also use texture features owing to its property to specify regular patterns. Some examples of texture based retrieval includes gray co-occurrences matrix, Gabor filtering,wavelet decomposition etc. Although shape feature being an important features, it has not gained much demand compared to color and texture features [3].

Although CBIR proved better results than TBIR, it failed to represent the semantic attributes of the images [4].This gave rise to Semantic based image retrieval. Semantic based retrieval attempts to capture the semantic meaning by extracting cognitive concept of a human for mapping the low-level features to high level features. Retrieval based on semantic features also allows users to input more easy and intuitive queries [5].

The further research is based on the topics as below. Section 2 deals with literature review followed by methodology used in Section 3.The analysis is given in Section 4 followed by experimental results in Section 5. The conclusion for the research work is presented in Section 6.

2. LITERATURE REVIEW
Begum Demir et.al have used a hashing based method on approximate nearest neighbor search. The method enables fast and accurate retrieval by mapping high dimensional feature vectors into binary hash codes and later into hash tables for easy search Two Kernel based Methods (KULSH AND KSLSH) are used and comparative analysis between both methods are carried out .After evaluation it is observed that both methods sharply reduces the image retrieval time and storage cost [6].

Begum Demir et.al has further researched on hashing based method for fast retrieval in huge remote sensing data archives. Image retrieval is based on hamming distances of image hash codes. The paper extends the LSH methods by designing hash functions for cases where the similarity function is a kernel function .For future work , hashing methods can be applied for the retrieval of long time series of remote sensing images [7].

Min Wang et.al puts forth a remote sensing image retrieval scheme by using scene semantic mining .The paper presents a variety of technical schemes for semantic based retrieval of RS Image scenes .It includes a process of conversion of Image low level VF's into high level spatial semantics followed by spatial SS modeling .Finally a prototype system is also used for validation purpose. Results prove that the system provides practical solutions for scene matching in SBRSIR [8].

In the paper titled Image Retrieval using Hash Code and Relevance Feedback Technique, Sapana Prakash Mali and Nitin N. Patil have put forth the importance of Relevance feedback in the area of image retrieval. Relevance describes a

method where the system changes the present query with the feedback given back by the user based on relevance of the given images. The above research provided better results when applied on the dataset of Flickr images [9].

Erchan Aptoula work presents the results of applying global morphological texture descriptors for content-based RS image retrieval. It also introduces new descriptors which exploits Fourier power spectrum of the input query. The proposed approach displays good results even though there were shortcomings related to short vector lengths [10].

Houria Sebai et.al proposes a CBIR approach based on 3D-Local Binary Pattern and Histogram of Orientated Gradients features. The proposed work aims to improve the performance by optimizing the image features. It also decreases the complexity by providing better matching between the images The feature are extracted using both shape and texture information [11].

T. Liu et.al proposes a region level method based on semantic mining in order to solve the issues of semantic gap. A uniform depiction for each image was constructed by segmenting the images by regions .Moreover the EM method was also applied to mine the minute semantic features. The proposed work gave better segmentation results along with good precision and recall when experimented on a dataset of thousand satellite-images [12].

3. METHODOLOGY

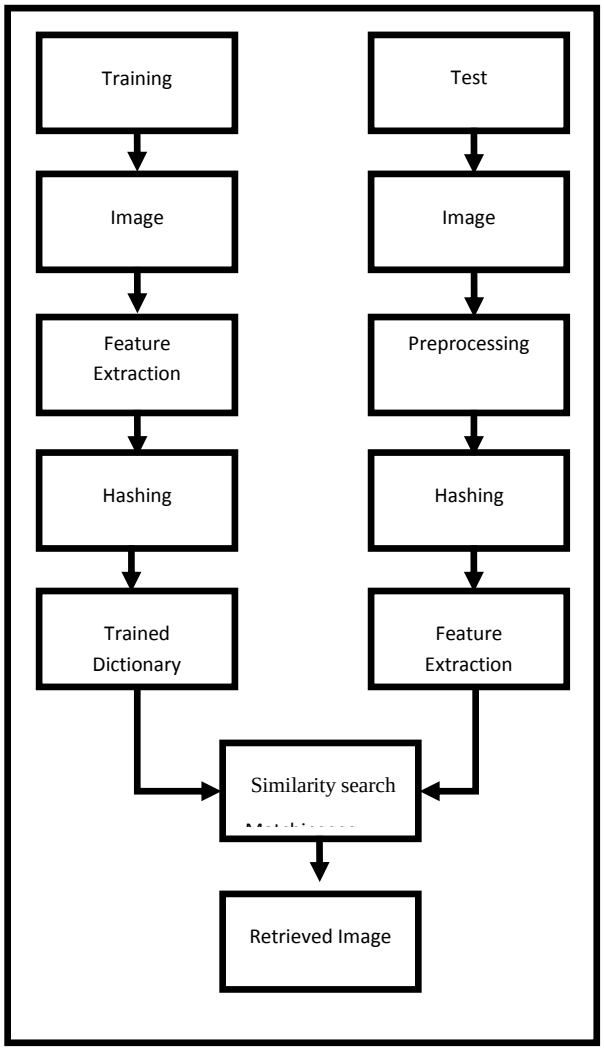

Figure 1: Methodology of the proposed work

STEP 1: CREATING AN HASH DATABASE

In the first step, an hashing-based technique for feature extraction takes place. The features are extracted in an unsupervised manner. The hashing aims at converting high-dimensional image feature vectors into distinct binary hash codes, which are indexed into a hash table that enables real-time search and accurate retrieval. Thus the images are hashed on the basis of hash codes obtained and stored in the hash database.

STEP 2: INPUT A QUERY IMAGE

In the second step, we input the desired query image by browsing it from the system which we want to search. The reason behind in putting the query image is to find the images from the database which are most similar to the query image.

STEP 3: PREPROCESSING AND FILTERING OF THE INPUT IMAGE

In the third step, the test image is converted into gray scale and an enhanced image is obtained using adaptive histogram equalization. For better results, the image is also filtered using Gaussian filter approach and sent for feature calculations.

STEP 4: FEATURE CALCULATIONS AND RETRIEVAL OF RESULTS

The filtered test image is again hashed in order to obtain the hash codes of the same and compared with the trained ones. Retrieval of results take place by matching similar hash codes and the best image is selected with respect to the highest correlation-coefficient.

4. ANALYSIS

The performance of an RS image retrieval system depends on its efficiency and effectiveness. Performance comparison of image retrieval systems can be measured using many methods such as Precision and Recall, Average normalized modified retrieval rank (ANMRR), Sensitivity and Specificity .In our research work, Precision and Recall are two standard measures that will be used in order to analyze the performance of RS image retrieval success**.**

5. EXPERIMENTAL RESULTS

5.1 Dataset Description

The dataset included satellite images of various areas such as coastal, desert and forests. The trained data set consist of 300 satellite images, 100 from each specific area.

5.2 Implementation

The project work is divided mainly into two parts .1) Creation of hash database and 2) Detection of Images. The work has been partially implemented in MATLAB R2014a.

5.2.1 Main graphical user interface

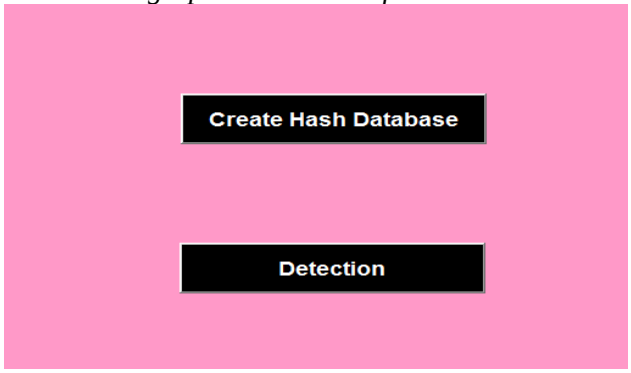

Figure 2: Main GUI of the research work

5.2.2 Creating Hash Database

Figure 3: Creating a hash database for given satellite images

5.2.3 Detection of images

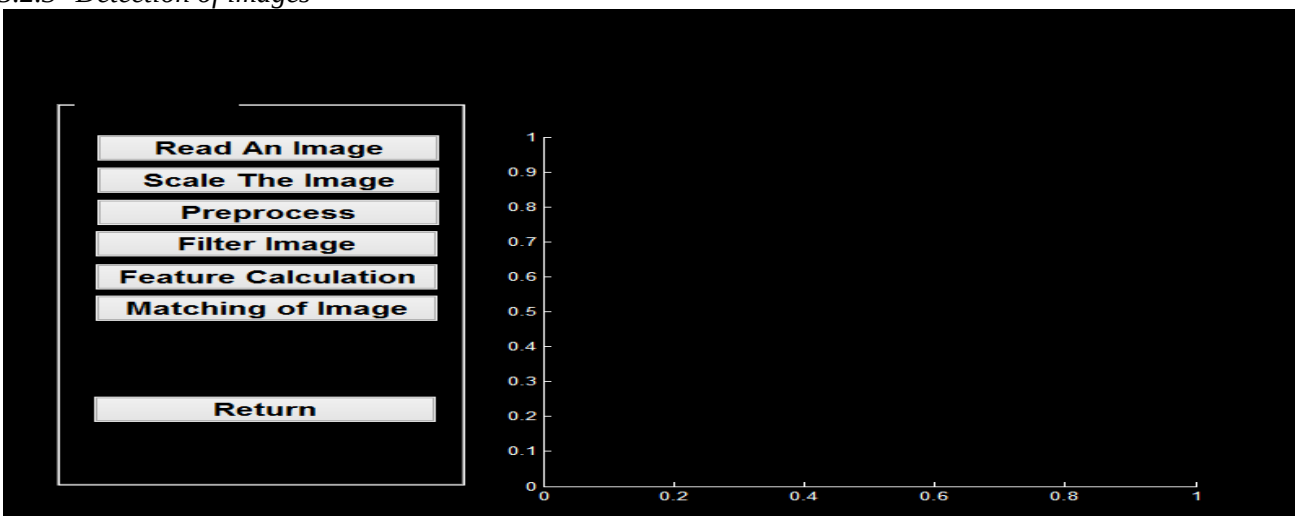

Figure 4: Detection of images using various functions

5.2.4 Giving a query image and retrieval of similar result

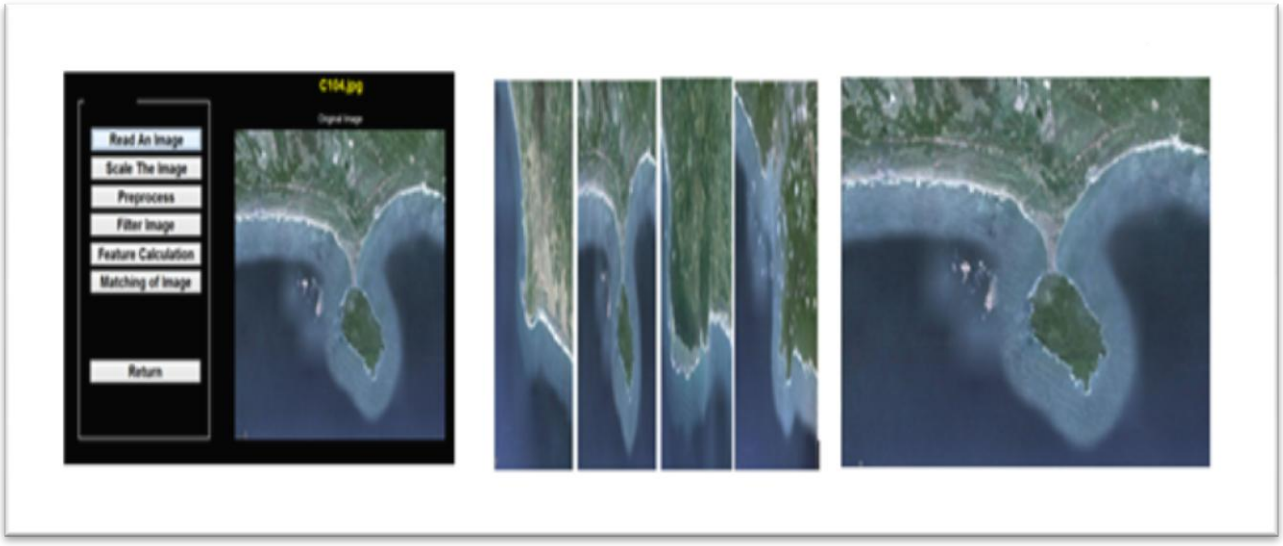

Figure 5: Input a query image and retrieve similar result with the best match

6. CONCLUSION

RS images contain valuable information that can be used in different applications. Due to the huge quantity of the image data, the storage and the management of the images are necessary. The creation of satellite image database requires efficient tools for searching and browsing the image database. Feature extraction is an important part in RSIR approach. For this research work, three features such as gray level, texture and shape are used to retrieve images. The accuracy of the image retrieval can be improved by using the other features such as spatial relation etc. The performance of the system can be better evaluated by the standard precision and recall graphs, with the increasing number of images. This system can be further improved and used as a tool to compare the images in real time, for other applications as forest fires, dust storms and etc. In this research, two different methods i.e. Semantic mining and Hashing have been combined in order to obtain better results with accurate precision and recall.

7. REFERENCES

[1] Harpreet Kaur, Kiran Jyoti,2013:“ Survey of Techniques of High Level Semantic Based Image Retrieval ”, International Journal of Research in Computer and Communication technology, Volume 2, Issue 1, January 2013 .

[2] Reshma Chaudhari and A. M. Patil ,2012:“ Content Based Image Retrieval Using Color and Shape Features”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering , Vol. 1, Issue 5, November 2012

[3] Zobeir Raisi, Farahnaz Mohanna and Mehdi Rezaei ,2014:“Applying Content-Based Image Retrieval Techniques to Provide New Services for Tourism Industry”, Int. J. Advanced Networking and Applications Volume: 6 Issue: 2 , ISSN : 0975-029, 2014

[4] Vishal S. Kore, B. A. Tidke and Pankaj Chandre,2015:“ Survey of Image Retrieval Techniques and Algorithms for Image-Rich Information Networks”, International Journal of Computer Applications (0975 – 8887) Volume 112 – No. 6, February 2015

[5] Hui Hui Wang, Dzulkifli Mohamad,and N.A Ismail ,2009: “ Image Retrieval: Techniques, Challenge, and Trend”, Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:3, No:12, 2009

[6] Begüm Demir,and Lorenzo Bruzzone:“ Hashing-Based Scalable Remote Sensing Image Search and Retrieval in Large Archives”, IEEE Transactions On Geoscience And Remote sensing, 2014.

[7] Begüm Demir,and Lorenzo Bruzzone :“ Kernel-Based Hashing For Content-Based Image Retrieval In Large Remote Sensing Data Archives ”, IEEE Transactions On Geoscience And Remote sensing,2015.

[8] Min Wang and Tengyi Song:“ Remote Sensing Image Retrieval by Scene Semantic”, IEEE Transactions On Geoscience And Remote sensing,Vol. 51, No. 5, May 2013.

[9] Sapana Prakash Mali, Nitin N. Patil :“ Image Retrieval using Hash Code and Relevance Feedback Technique”,International Journal of Computer Applications (0975 – 8887).

[10] Erchan Aptoula :“Remote Sensing Image Retrieval With Global Morphological Texture Descriptors”, IEEE Transactions On Geoscience And Remote sensing, Vol. 52, No. 5, May 2014

[11] Houria Sebai ,Assia Kourgli :“ An adaptive CBIR system for remote sensed data” 978-1-4799-3990- 2014 IEEE

[12] T. Liu, L. Zhang, P. Li, and Hui Lin “Remotely sensed image retrieval based on region-level semantic mining,” EURASIP Journal on Image and Video Processing, pp. 1687–5281, April 2012.