Design and Development of an Image Classification and Recognition System for Cubesat Constellation

Jean Marie Gashayija
Cape Peninsula university of Technology
Bellville campus

Prof. Elmarie Biermann
Cape Peninsula university of Technology
Bellville Campus

ABSTRACT
The major problems with stored images in large image database are retrieval of precisely, clear images and semantic gap. Proposed methodology approach, author will arrange, classify and categorize image into database in order to solve identified problem of retrieval and semantic gap. This proposal in progress is to solve this underlying problem of semantic gap for large images databases for small satellite database (such as CubeSats constellation) and look as well effective efficiency algorithm to improve existing methods. This paper proposes a solution based on image classification and recognition methods (such k-nearest classification and support vector machine methods) to solve this underlying semantic gap problem.

General Terms

Image retrieval, content based image retrieval, image classification and recognition, Pattern recognition.

Keywords

Image recognition, image classification and pattern recognition and image retrieval systems.

1. INTRODUCTION
Low cost and reduced development time of small satellites (and especially CubeSats) have encouraged academic institutions to participate in space activities. These development platforms have motivated small satellite programs throughout the world at Universities and emerging space industries. Imaging of the earth is a common mission for most of these small satellites, as these pictures are the most convenient means of obtaining and conveying information about the earth. Pictures concisely convey information such as positions, sizes and inter-relationships of objects and portray spatial information that can be recognized as objects [13].

An imaging CubeSat is equipped with a high resolution camera on board for capturing digital images as well as a mass storage device for storing these pictures. The imager forms part of the imaging payload subsystem onboard the CubeSat and can take pictures of any point on the earth if the orbit permits. Large mass storage memory on board the satellite is required to store huge amounts of image data. In commercial satellites, hard disk drives are used, while in small satellites flash memory on SD cards are utilized. The captured images are then transmitted to the ground station and stored in a database. Advanced processing of those images need to be performed in order to identify precise and clear images so that they are usable for applications such as mapping, disaster monitoring, and weather [13].

One of the challenges for such a CubeSat mission is the formulation of a computerized image database which will enable precise retrieval of specific images for the user at low cost. An Image database is a combination of storage and processing systems that facilitates the handling of images by structuring them and allowing easier querying of the database. The focus of this project lies within the image database design to ease the retrieval, storage, and querying of images [13].

2. BACKGROUND
The French South African Institute of Technology (F'SATI) is a graduate school focusing on the training of graduate students in the broad fields of electrical engineering and information and communication technologies.

F'SATI was established in 1996, through collaboration between the French and South African Governments. A number of role players such as the National Research Foundation (NRF), the Tshwane University of Technology (TUT), and the Paris Chamber of Commerce and Industry and CPUT, committed to the establishment of this advanced research [4].

F'SATI has two branches in South Africa, namely the Pretoria branch located at the Tshwane University of Technology and the Cape Town branch on the premises of the Cape Peninsula University of Technology. They offer graduate programmes at master and doctorate level in partnership with certain universities in France such as ESIEE Paris, and the University of Paris Est Creteil (UPEC) [4].

CPUT in collaboration with the South African National Space Agency (SANSA) and the Hermanus Magnetic Observatory (HMO) are designing a 1U and 3U CubeSat. The 1U CubeSat will be completed in 2011 and is set to be launched in 2012, while the 3U CubeSat will be completed in 2013.

2.1 The CubeSat Standard
As described by [4],the CubeSat were originally proposed by Robert Twiggs from Stanford University and later developed by the California Polytechnique State University. They introduced a general standard cubesat specification: a 10cm cube shape with mass of no more than 1 kg (as seen Figure 1).

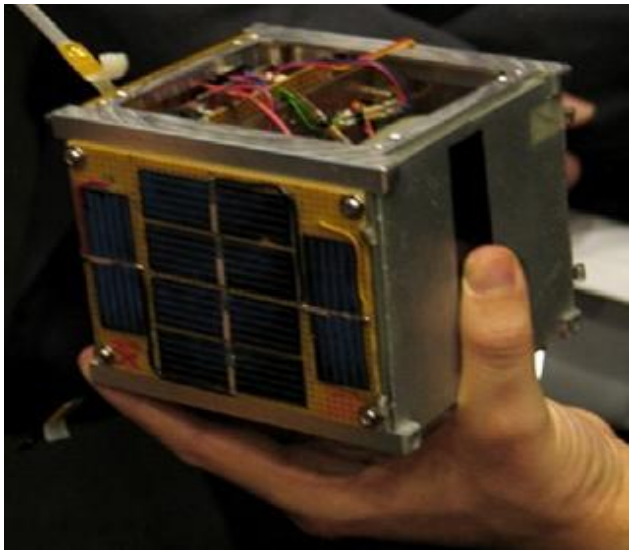

Figure 1: CubeSat model (SwissCube, 2011)

Most CubeSats are being utilised for various missions such as military, scientific as well communications purposes. The standard 1U CubeSat has five subsystems namely *power*, *communication* which assist in transmitting and receiving of the data to/from the ground station, and on-board *computer* (OBC), *attitude determination and control system* that maintains CubeSat stabilization and orientation of the desired direction and position as well to enable the camera to take the images at certain position and location [12].

[12] Describes that satellites are normally classified according to their mass. For example small satellites weigh less than 500kg; medium satellites weigh between 500kg-1000kg and large satellites weighing more than 1000kg. Furthermore, small satellites are divided into five subcategories according to their weight. Mini satellites weigh between 100kg - 500kg, micro satellites between 10kg-100kg, nano between 1kg - 10kg, pico between 0.1kg - 1kg and femto less than 0.1kg.

The Cubesat falls within the pico-satellite category and can have different types of units such as a one unit or three units. CubeSat units are defined as standard 10cm cubed Cubesat stacked together to form larger versions of CubeSat. A two unit Cubesat is thus two 10 cm cubes stacked one on top of the other. The maximum number of CubeSats that can be stacked and still fit into the Poly Picosat Orbital Deployer(P-POD) is three or 3U CubeSat [4].

A CubeSat ground station is currently being constructed on the Bellville campus of the CPUT. The ground station will be used by students to conduct radio frequency propagation experiments, track and communicate with various orbiting satellite as well as eventually commission and operate FSATI's 1U and 3U CubeSats once launched [4].

These CubeSats will be launched into polar orbit at altitude between 450km and 650km.Two CubeSats missions are being developed as part of the student training programme. Currently a one unit CubeSat (weighing 1 kg) and a three unit CubeSat (weighing less than 4kg) are being developed. The one CubeSat carrying a five megapixel matrix camera for earth observation and an advanced three axis simple magnetic attitude determination control system(ADCS) will be launched in 2012 and the three units Cubesat will carry a beacon transmitter which will be used to characterize ionospheric radar antennas in Antarctica [4].

2.2 Image Retrieval Systems

Two broad approaches for image retrieval exist, namely text based image retrieval and content based image retrieval.

Text-based image retrieval can be traced back to the late 1970s. A very popular method of image retrieval then was to first annotate the image with text and then use a text based database system to perform retrieval [2].

Most existing image retrieval systems are text based, yet images frequently have little or no accompanying textual information. Problems with text-based image retrieval systems such as Eurovision, Ciquest, mediGIFT, imageCLEF have prompted increasing interest in the development of image-based solutions [6].

Content-based image retrieval (CBIR) has been a major research area within computer vision and related application fields. The CBIR approach of image retrieval is based on image features rather on textual annotations to search for an image. Content-based image retrieval relies on the characterisation of primitive features such as color, shape, and texture that can be automatically extracted from the images [6]. [8] discuss in detail the various technologies for image indexing and retrieval based on shape, color, texture, and spatial location.

2.3 Image classification

Image classification is the process of finding a model (or function) that describes and distinguishes data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose class label is unknown. The derived model is based on the analysis of a set of training data (i.e. data objects whose class label is known)[8].

Image classification is a critical step in image understanding. To extract information from images, classification is performed first to separate an image into regions of different types. Image classification is now widely used in different fields, such as computer vision. The overall objective of image classification procedures is to automatically categorize all pixels in an image into classes or themes.

3. CONTENT BASED IMAGE RETRIEVAL SYSTEMS

Content based image retrieval systems have been developed to search and retrieve images of interest by content such system will be described below.

- **Query by Image Content (QBIC):** This system was designed by IBM for retrieving images in large databases, relying on two parts such as database construction and database query. The QBIC uses also primitive features such as color, textures, shape as well the sketch. During the database construction, the visual features of images are

extracted and stored in image database as index keys. During the database query, a graphical user interface is provided for users to composite an example image or graphical user interface is provided for users to acquire an example image or graphical query. The system can support complex multiple visual features queries. The QBIC was first designed as a pure image query system and later extended to include video data. The QBIC system extracts the relevant visual features from the example image or graphical query to match the corresponding visual features in the databases for similar images[10].

- **Virage system** is a content based image retrieval system based on an image search engine developed at Virage Inc. Similar to QBIC and supports visual queries based on colour, composition, texture and structure[1].

- **WebSEEK Systems** is a visual feature search engine and web based oriented text/image search engine, which was developed at Columbia University. The visual features used in these systems are colour set and wavelet transforms based texture. Uses binary tree based indexing algorithms for speeding up retrieval processes [14].

- **The VisualSEEK system** was developed at the Columbia University Center for Telecommunication Research. The system supports retrieval of still images and video based on visual features and spatial layout. The VisualSEEK is a World Wide Web oriented text/image search engine. It supports queries based on both keywords and visual content[14].

- **The Netra system** was developed at the University of California; Santa Barbara (UCSB).The system supports features of color, texture, shape and spatial information of segmented image regions to region-based search. Images are segmented to homogenous regions. Using the regions as the basic unit, users can submit queries based on features that combine regions of multiple images [17].

- **RetrievalWare system** is a content based image retrieval engine. Its more recent search engine uses colour, shape, texture, brightness, colour layout, and aspect ratio of the image as the query features. Allows combinations of several visual query features, whose weights are specified by the users [10].

All this literature review systems were designed to solve certain problems which were facing the textual based system as well the contented based retrieval systems. But the major problems still exists for example of such problem let say different people may supply different textual annotation for the same image. This makes it extremely difficult to answer user queries reliably. In content based such as QBIC, Virage, WebSEEK, VisualSEEK and the Netra system described above they eliminated certain problem of textual based image retrieval. We will seen in the next following section what most systems problems did or still experiencing [16].

4. PROBLEM STATEMENT

The main problem in image retrieval is semantic gap. As described by [16] the semantic gap is the lack of correlation between the semantic categories that a user requires and the low-level features that CBIR systems offer. The semantic gap between the low-level visual features (such as color, shape, texture, etc.) and semantic concepts identified by the user remains a major problem in content based image retrieval. Other sub-problem is locating a desired image in large and varied collection in the database. As stated by Enser (1995), the problems of image retrieval are becoming widely recognized.

The classifying image of larger amount into database still challenging problem in content based image retrieval system such as QBIC and RetrievalWare e.t.c. once image classification can be achieved will enhance the performance of CBIR systems. This rather difficult problem has not been addressed in current content based image database retrieval system. The effectiveness of all current CBIR system limited in a way they only operate in low level features only [18].

The problem involves entering an image as a query into a software application that will be designed to employ CBIR techniques in extracting visual properties, and matching them. This is done to retrieve images in the database that are visually similar to the query image.

The major problem which will be addressed in this research is of classifying of larger amount of images into small classes to improve effectiveness the system by using comparative and efficiently best algorithm to retrieve store images from database. This rather difficult problem has not been adequately addressed in current image database systems. The focus will be to improve existing methods and implement this on CubeSat applications.

In this research, we intend to evaluate past and current image database retrieval systems. Arrange, classifying and categorizing image into database keeping in mind problem will want to solve.

5. RESEARCH AIMS AND OBJECTIVES

The main objective of this research is to design an image classification and image recognition and retrieval system for a CubeSat constellation that will enable the efficient categorizing and characterizing, storing aswell as fast retrieving of images. Essential to this is the design and implementation of data mining and implementation and pattern recognition algorithms. The aim is to choose suitable methods, algorithms and software programs in order to retrieve image into database using content based methods. The sub objectives that need to be accomplished are as follows:

- To uncompress the received images
- To characterise and categorise images as received from the CubeSats constellation.
- To design and develop a suitable storage system for images through the 3U CubeSat constellation.
- To study and data mining and pattern recognition algorithm in CubeSat constellation
- To design and implement image retrieval system for images obtained from 3U CubeSat

- To design and develop a suitable software interface system that will be implemented within a distributed network
- To deploy a designed image retrieval application on the web
- To study what will be the method of calculating the similarity between images
- To study how can the images be ranked and retrieved
- To study what will be the method to evaluate the performance of the CBIR system
- Implementation and integration of the system

6. METHODOLOGY

The proposed solution approach will be to extract the primitive feature of a query image and compare them to those of the database image. The image feature under consideration will be colour; texture and shape. However, by using matching and comparison algorithms the colour, texture and shape features of the one image shall be compared and matched to the corresponding features of the similar image. The comparison shall be performed using colour, texture and shape features. To retrieve stored database images will depend on what similarities the images have against user queries. In order to achieve this certain methodologies shall be adopted:

- Literature review on CubeSat imaging missions: whereby information on CubeSat environment will be collected from related books, newly published journals, previously done work and internet search engines.. This will entail clear understanding of all subsystems especially the imaging payload.

- Literature review on image database systems and data mining techniques.

- Literature review on image retrieval techniques. Detailed study on all possible methods and ultimately choosing the most appropriate one for this research. This includes a study on image retrieval systems and what problems they faced in general.

- Experimental method: Simulate and retrieve image in standalone designed application before implementation on distributed site.

- Evaluation of proposed design and implementation of the system

7. DESIGN AND IMPLEMENTATION

The system to be developed will use certain software such as a compiler to execute the source code program such as graphical interface which will be developed, i.e. Matlab to generate colour histogram, Eclipse to compile and develop an interface to interact with users. Consider classification and characterization of the image stored in the database. In order to carry out this research the following will be considered

- Designing of CBIR system architecture
- Choice of language to develop the CBIR system
- Choice of database to store images
- Design of the interface for user interaction
- Test bed for evaluating performance of the application

8. SYSTEM PERFORMANCE AND EVALUATION

In content-based image retrieval system, efficiency is important due to the large data size of digital images. There is a need of retrieval effectiveness which is ability to retrieve relevant items and ability to discard irrelevant items; this becomes one of the most important parameter to measure the performance of the content based image retrieval systems (Huiskes & Lew, 2008)

Evaluation has been used since beginning of 1950s. Many researchers have paid attention to understanding and developing criteria, measures and methods for evaluation of image retrieval systems and most research community insist on high quality evaluation. Several standard criteria for evaluation have been developed or suggested for image retrieval systems which include effectiveness, efficiency, usability, satisfaction and cost effective [7].

9. CONCLUSION

This work introduced the proposed image database and retrieval system for a CubeSat mission. The system forms an important part of the ground station segment of imaging missions. There exist several solutions for storage and retrieval systems and there is no single solution to every mission or system. A significant limitation of current CBIR technology is the problem of efficiency retrieving the set of stored image similar to a given query. And another problem in image retrieval is semantic gap. The semantic is the lack of coincidence between the information that one extract from the visual data and the interpretation that the same data have for a user in a given situation. This research project will aim in reducing this semantic gap.

10. ACKNOWLEDGMENTS

Thanks are due to my supervisor Prof.Bierman for guidance in times where I was in douts, without her this journal could not have been a written success. Thanks also due FSATI for financial support.

11. REFERENCES

[1] Bach,R.J.,Fuller.C.,Gupta,A.,Hamparur,A.,Horowitza,B.,Ja in,C.R.&Shu,C.R.1996.*Virage image search: an open framework for image management.*Proc SPIE, 2670:76-87.

[2] Chang, C. Y., Wang, H.J & Li ,C.F.2009.Semantic analysis of real world images using support vector machine. Expert systems with application: An international journal, 36(7): 10560-10569

[3] Datta, R., Joshi, D., Jia, L. & Wang, J. 2008. *Image retrieval: ideas Influences and trends of the new age.* ACM computing surveys, 40(2):1-64, April.

[4] French South Africa Institute of Technology.2011. http://www.cput.ac.za/fsati/. [10 August 2011]

[5] Goodrum, A. & Spink, A.1999. Visual information seeking:A study of image queries on the World Wide Web.Proceedings of the 1999 annual Meeting of the

American Society for Information Science.,October 31-Nov.4,1999,Washington,DC:665-674.

[6] Goodrum, A. 2000. *Image information retrieval: an overview of current research.* Information Science Research, 3 (2):63-67.

[7] Huiskes, J.M & Lew, S.M.2008. Performance evaluation of relevance feedback method.CIVR'08, ACM: 239-248.July 9.

[8] Han, J.&Kamber,M.2006.Data mining:concepts and techniques.2nd ed.Hamilton:Morgan Kaufmann. March

[9] Idris, F. & Panchathan, S.1997.Review *of image and video indexing techniques*. Journal of visual Communication and image representation, 8(2):146-166, June.

[10] Mishra, J., Sharma, A. & Chaturvedi, K. 2011 .An Unsupervised cluster based image retrieval algorithm using relevance feedback methods. International journal of managing information technology, 3(2):9-16, May.

[11] Ray, A.K. & Acharya, T.2004.Information technology: principles and applications. New Delhi, Prentice hall.

[12] Rui, Y., Huang, T.S. & Chang, S.F.1999. *Image Retrieval: current techniques, promising directions and open issues*. Journal of Visual Communication and Image Representation, 10(1):39-62.March 1.

[13] Schor, D., Scowcroft, J., Nichols, C. & Kinsner, W.2009.*A command and data handling unit for pico-satellite missions.*Electrical and Computer Engineering, 2009.CCECE'09. IEEE Xplore Canadian Conference on, 3-6 May 2009. St.John, NL: 874-879.

[14] Sivakumar M.V.K, Roy P.S., Harmsen K., Saha S.K.2004.*Satellite remote sensing and GIS application in agricultural meteorology.* Proceedings of the training workshop on,7-11 July,2003.Dehra Dun:1-23.

[15] Smith, J.R, and Chang, S.F.1996. Local color and texture extraction and spatial query. IEEE International conference on image processing, 3:1011-1014.

[16] Stricker, M. &Orengo, M.1995.Similirity of color images, in Storage and Retrieval for image and video Database III, Proc SPIE 2420:381-392.

[17] Wang, Z.J.2001.*Integrated region-based retrieval.Massachusetts*.Kluwer academic publishers group.

[18] Yang, C.2004.Content based image retrieval: a comparison between query by example and image browsing map approach. Journal of information science, 30(3):254-267.January 7.

[19] Zhang, H. & Zhong, D.1995.*A scheme for visual feature based image indexing*.Proc.SPIE, 2420:36-46.