The Egyptian Journal of Remote Sensing and Space Sciences (2012) **15**, 91–97

National Authority for Remote Sensing and Space Sciences

The Egyptian Journal of Remote Sensing and Space Sciences

www.elsevier.com/locate/ejrs
www.sciencedirect.com

System refinement for content based satellite image retrieval

NourElDin Laban [a,*], Motaz ElSaban [b], Ayman Nasr [a], Hoda Onsi [b]

[a] *Department of Image Processing and its Applications, Data Reception and Analysis Division, National Authority for Remote Sensing and Space Sciences (NARSS), P.O. Box 1564, Alf Maskan, Cairo, Egypt*
[b] *Department of Information Technology, Faculty of Computers and Information, Cairo University, P.O. Box 12613, Orman, Giza, Egypt*

Received 19 January 2012; revised 30 April 2012; accepted 1 May 2012
Available online 29 August 2012

KEYWORDS

Satellite images;
Content based image retrieval;
Query by polygon;
Retrieval refinement

Abstract We are witnessing a large increase in satellite generated data especially in the form of images. Hence intelligent processing of the huge amount of data received by dozens of earth observing satellites, with specific satellite image oriented approaches, presents itself as a pressing need. Content based satellite image retrieval (CBSIR) approaches have mainly been driven so far by approaches dealing with traditional images. In this paper we introduce a novel approach that refines image retrieval process using the unique properties to satellite images. Our approach uses a Query by polygon (QBP) paradigm for the content of interest instead of using the more conventional rectangular query by image approach. First, we extract features from the satellite images using multiple tiling sizes. Accordingly the system uses these multilevel features within a multilevel retrieval system that refines the retrieval process. Our multilevel refinement approach has been experimentally validated against the conventional one yielding enhanced precision and recall rates.

1. Introduction

Satellite images have become a common component of our daily life either on the internet, in car driving and even in our handheld mobile handsets. There is new image and video content appearing every second through multiple competing television and internet channels. Manual interaction with this large volume of data is becoming more and more inappropriate, which creates an urgent need for automatic treatment to store, organize and retrieve required content (Martins et al., 2002).

The traditional methods for retrieving images from geodatabases are: geographic location, date of acquisition and spectral/spatial properties of acquisition devices (Li and Bretschneider, 2007). Our need from the satellite scenes are specific contents. Therefore we need to retrieve images that contain our intended contents. The content based image retrieval (CBIR) approach challenge is how to fill the gap between the low level features that describe the scenes and our human understandable semantic concepts which is called the semantic gap (Gondra and Heisterkamp, 2008). In addition, these semantic concepts themselves may be defined differently, e.g. each one of us interprets what he sees from his point of view.

* Corresponding author.
E-mail address: nourlaban@gmail.com (NourElDin Laban).

Peer review under responsibility of National Authority for Remote Sensing and Space Sciences.

Production and hosting by Elsevier

1110-9823 © 2012 National Authority for Remote Sensing and Space Sciences. Production and hosting by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejrs.2012.05.001

There are other challenges in the field of satellite images itself (Li and Bretschneider, 2004). These images are georeferenced images; this means that all images form in reality a huge continuous image covering the entire earth surface. It is not always proper to deal with such content as isolated images. Moreover the concept of defining our areas of interest by rectangular shapes (as in traditional images) becomes limiting for Geographic information systems (GIS) to include georeferenced elements (points, lines and polygons) which define globally the absolute position of pixels. Satellite images have different bands according to the scanning system used in these satellites. There are many parameters that affect the scanning process such as: satellite altitude, angle of acquisition and climate circumstances. Moreover the scanned images should have georeference and orthorectification to adapt with unfolded earth surface according to the specific georeference system used as each pixel gets its absolute position. Besides, the scanned images must have radiometric balancing to compensate, as much as possible, for the different effects of undesired climate circumstances. In addition the satellite scene itself is more complicated than traditional images, i.e. it could be interpreted only by persons with a higher degree of expertise.

2. Review of related work

During the last decade many approaches have been proposed to retrieve satellite images using their content. Ma and Manjunath (1997) have designed region-based image-retrieval systems where the similarity between two images was measured based on individual region-to-region similarity and later extended to image-to-image similarity based on all segmented regions within the scenes. Li et al. have retrieved satellite images after classification into predefined semantic concepts as cloud, water, forest, urban area, farmland, bare soil and rock using grayscale images (Li and Bretschneider, 2003), or using multispectral isolated images (Li and Bretschneider, 2006, 2007; Li and Dong, 2011). Ferecatu and Boujemaa (2007) retrieved six predefined classes as city, cloud, desert, field, forest, and sea from isolated images and ground truth database. Niu and Ni (2005) have used multi-band isolated JPEG2000 coded images to retrieve an area of interest depending on the query image using hue saturation and value color model conversion. Tuia et al. (2011) have adopted a satellite images classifier using active learning in high resolution hyperspectral images. Blanchart et al. (2011) have developed a system which combines the auto-annotation systems and the category search engines.

Although all the previous work has taken satellite image as the matter of interest, it did not take into account:-

- the continuous nature of satellite cover and the geospatial relationships between different satellite images.
- The multiple hierarchy of frame work is based on the adjacent area to first stage candidates and irregular shape of the required semantic.

3. System overview

The components of our proposed system consist of three consequent processes as shown in Fig. 1.

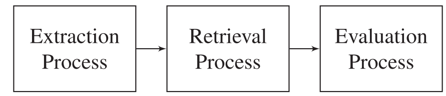


Figure 1 System overview.

Firstly, through the extraction process we extract feature vectors according to multiple tile sizes to establish the feature database. We use the special characteristics of satellite images for tiling, by using the georeferenced grid instead of using traditional pixel grid. Secondly, in the retrieval process we formulate our query using a QBP paradigm through a multistage process. We coarsely retrieve the most adequate area of interest, and by using the resulted area and its adjacent regions we refine it to get our final results. Finally, in the evaluation process we make different comparisons before and after the refinement process.

3.1. Feature extraction process

The choice of features to describe the imagery's contents depends on the type of data. Our approach decomposes the utilized multispectral scenes into non overlapping sub images. Two different kinds of features were exploited, namely color histograms and wavelet-based textures for each band of the multispectral image to form the feature vector (Li and Bretschneider, 2004). The color histograms of the satellite images are extracted using the quantization method. The wavelet-based textures were extracted using Daubechies wavelets (Boggess and Narcowich, 2002), which analyze an image in the spatial-frequency domain. Let us denote a tile's feature vector by V_t described by Eq. (1)

$$V_t = [V_t(1), V_t(2), \ldots, V_t(N_h + N_w)] \times L. \quad (1)$$

where V_t is feature vector of tile number, N_h is the number of color histogram bins and N_w is the number of bins for Daubechies wavelet coefficients and L is the number of bands in the image. The feature database is thus defined as DB_x expressed in Eq. (2)

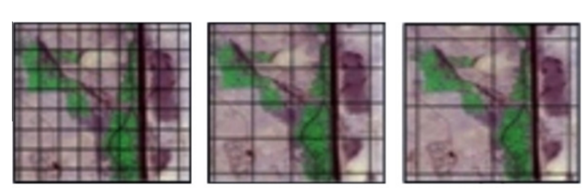
$$DB_x = [V_1, V_2, V_3, \ldots, V_n]. \quad (2)$$

where DB_x is the feature database of tile size x and V_n is the feature vector of tile number n in database.

$$DB_{global} = [DB_{1x}, DB_{2x}, DB_{3x}, \ldots, DB_{mx}] \quad (3)$$

Eq. (3) express global feature database where DB_{global} is the global feature database and DB_x is the feature database of size number m and initial tile size x (Fig. 2).

According to this approach we have m databases with different tile size databases which will be used for our refinement system. We use the benefits of georeferencing the satellite im-



(a) 1km grid (b) 2km grid (c) 4km grid

Figure 2 Different grid sizes used for tiling.