System Refinement for Content Based Satellite Image Retrieval

NourElDin Laban*, Motaz ElSaban§, Ayman Nasr * and Hoda Onsi§
*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
§ Department of Information Technology, Faculty of Computers and Information,
Cairo University, P.O.Box 12613 , Orman, Giza, Egypt

***Abstract*—We are witnessing a large increase in satellite generated data specially 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.**

I. Introduction

Satellite images have become a common component of our daily life either on the internet, in car driving and even in our hand-held 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 this content [1].

The traditional methods for retrieving images from geo-databases are: geographic location, date of acquisition and spectral /spatial properties of acquisition devices [2].Our needs 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 [3]. 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.

There are other challenges in the field of satellite images itself [4].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 orthorectifation 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.

II. Review of Related Work

During the last decade many approaches have been proposed to retrieve satellite images using their content. Ma and Manjunath 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 [5]. Li et al. have retrieved satellite images after classification into a predefined semantic concepts as cloud, water, forest, urban area, farmland, bare soil and rock using grayscale images [6], or using multispectral isolated images [7], [2], [8]. Ferecatu and Boujemaa retrieved six predefined classes as city, cloud, desert, field, forest, and sea from isolated images and ground truth database[9]. Lei Niu et al. have used multi-band isolated JPEG2000 coded images to retrieve area of interest depending on query image using hue saturation and value color model conversion [10]. Tuia et al. have adopted a satellite images classifier using active learning in high resolution hyperspectral images [11]. Blanchart et al. have developed a system which combines the auto-annotation systems and the category search engines [12].

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 based on the adjacent area to first stage candidates and irregular shape

of the required semantic.

III. System Overview

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

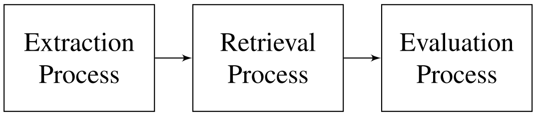


Fig. 1. System Overview

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 multi-stage 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.

A. 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 [4]. The color histograms of the satellite images are extracted using the quantization method. The wavelet-based textures were extracted using Daubechies wavelets [13], which analyze an image in the spatial-frequency domain. Let's denote a tile's feature vector by V_t described by equation 1

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

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 wavelets coefficients and L is the number of bands in the image. The feature database is thus defined as DB_x expressed in equation 2

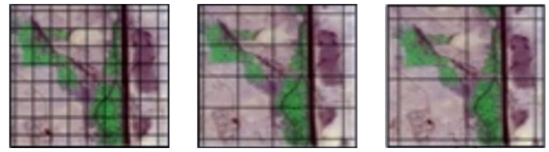
$$DB_x = [V_1, V_2, V_3, ..., V_n] \qquad (2)$$

Where DB_x is feature database of tile size x and $V_n$is feature vector of tile number n in database.

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

Equation 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 (figure 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 image for tiling, by using the georeferenced grid instead of using traditional pixel grid.



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

Fig. 2. different grid sizes used for tilling

B. Retrieval Process

Building Query: Geographic information systems (GIS) have been developed to deal with georeferenced data, which use georeference polygons to determine area of interest. We have designed our query to use these polygons to determine the area of interest which will be used later as the query data set instead of using rectangular images as shown in figure 3.

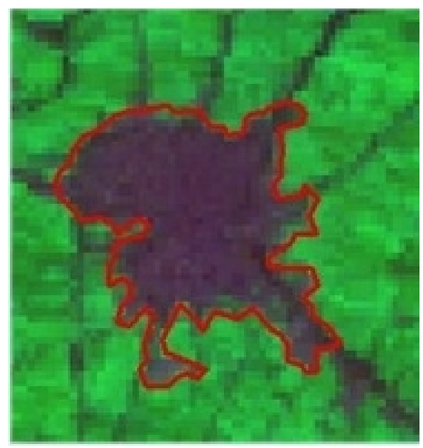


Fig. 3. Query by polygon

Using Query by polygon approach gives us the ability to exactly determine area of interest of a specific semantic concept. More specifically, this allows us to select the best tile size and feature database that improve our refinement process. For creating the polygon area of interest, the procedure is as follows: - First, we select the area of interest as a polygon, then extract a polygon regions from the satellite image mosaic. Second, we select the appropriate tile size for the selected area, then tiling the whole image area to subimages with this tile size. This results into query tiles which will be used for the first phase of the retrieval process as shown in figure 4.

After the tiling process, we end up with two types of tiles, Type A which is fully contained within the area of interest and Type B which is partially contained. For the first phase of retrieval process we use Type A tiles (pure semantic tiles) as shown in figure 5(a). Whereas tiles for the second stage will be the partially contained Type B tiles as shown in Figure 5(b).

Retrieval procedure: Our system is divided into two stages, where the output of the first stage is fed to the second stage

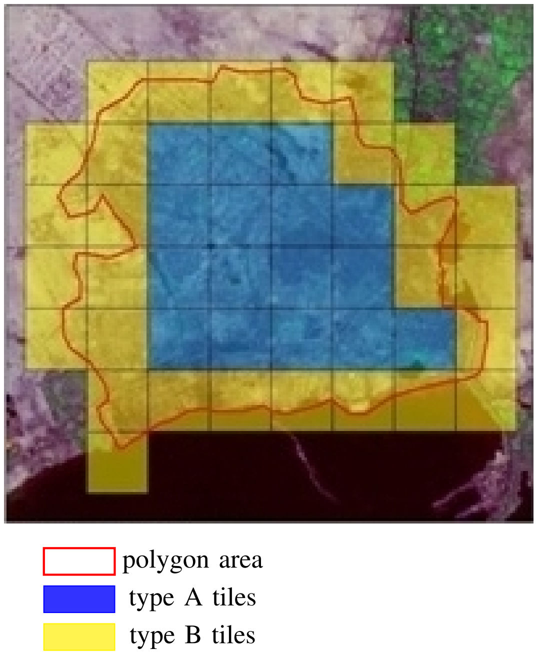
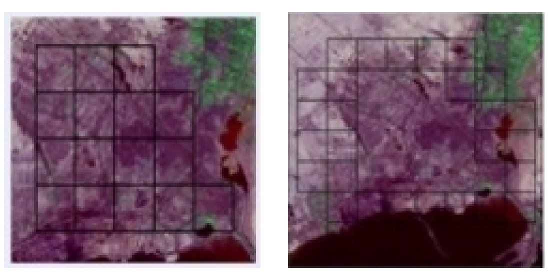


Fig. 4. Tiles types according to intersection



(a) Type A tiles (b) Type B tiles

Fig. 5. different types of tiles

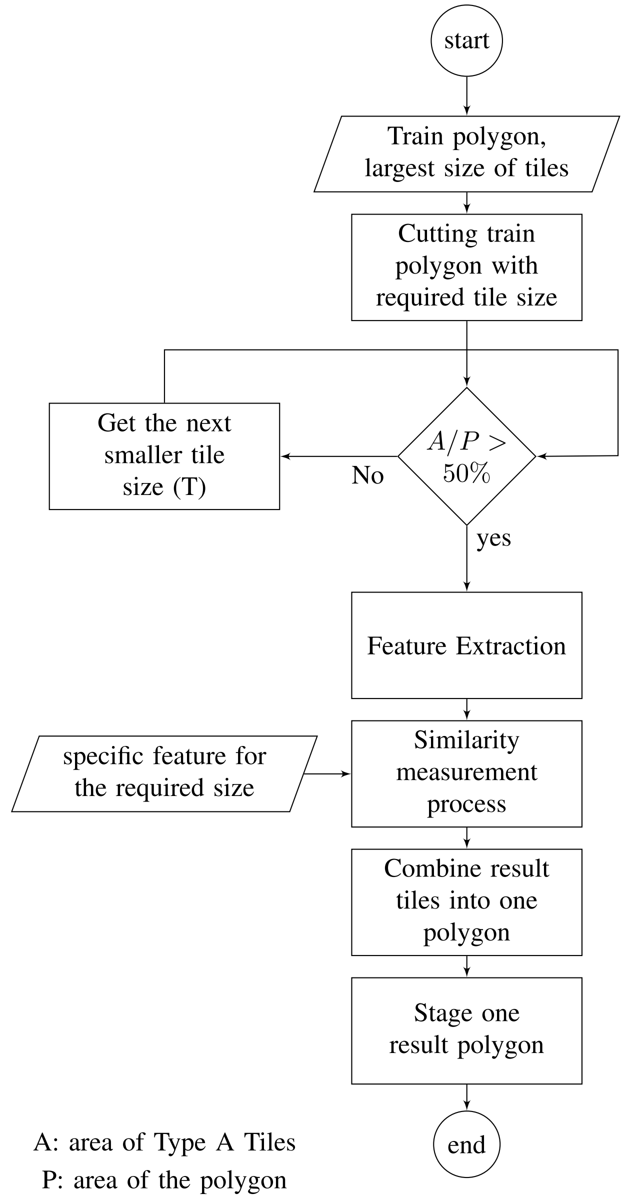


Fig. 6. stage one of retrieval process

as follows:

1) First stage(candidates selection): Initially,as shown in figure 6 the preferred polygons that represent the required semantic area are determined. We select the tile size that can perfectly enhance the refinement process. Firstly; we tile the query polygon area using the greatest tile size in the feature database as shown in figure 7.

Then we calculate the summation of type A tiles area and calculate type A percent as in equation 4

$$.Type\ A\% = \frac{type\ A\ tiles\ area}{query\ polygon\ area} \times 100 \tag{4}$$

As we have different tile size in the DB_{global} , it is important to determine the most appropriated size ax that will describe our query area as whole but with the least number of tiles. This procedure guarantees that query tiles contain as much as possible of the query polygon during the first stage. After computing the query tiles with an appropriate size we compute the Euclidean distance D_{ij} between the query features vectors and DB_{global} as in equation 5.

$$D_{ij} = \sqrt{\sum_k X_{ik} - X_{jk}} \tag{5}$$

Where D_{ij} is the distance between query feature vector X_i and database feature vector X_j and k is the corresponding element number in the feature vector. The distance vector D_j is formed as in equation 6.

$$D_j = [D_{1j}, D_{2j}, D_{3j}, ..., D_{nj}] \tag{6}$$

Where n is the number of query tiles. Next, we get D_q that represents the minimum distance of vector D_j. Then we

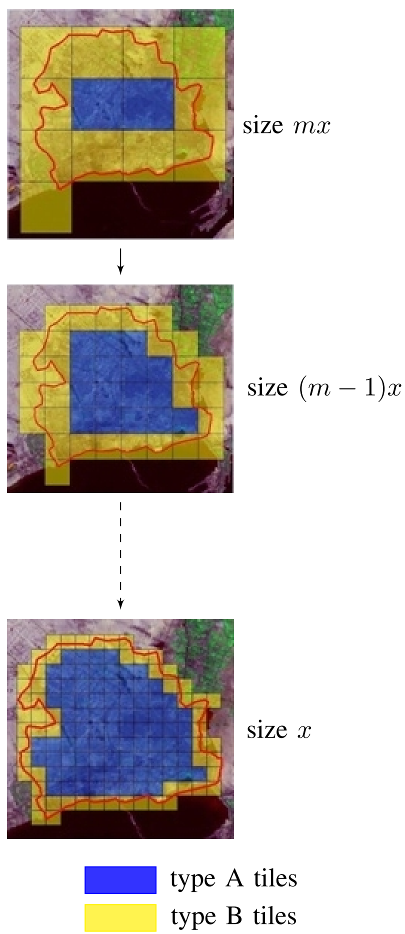


Fig. 7. tile size selection

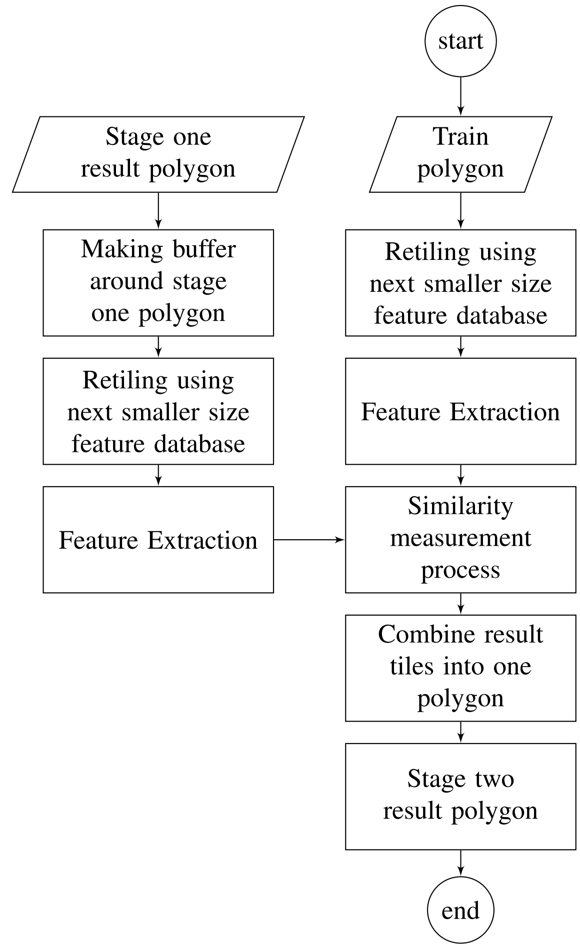


Fig. 8. stage two of retrieval process

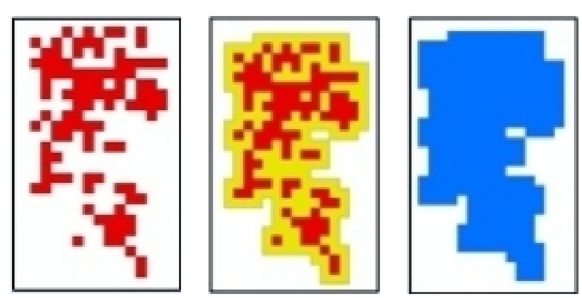
ascending order database distance vector D as in equation 7.

$$D = [D_{q1} < D_{q2} < D_{q3} < .... < D_{qm}] \tag{7}$$

where m is the number of tiles in the database. The nearest tiles are selected and merged in one polygon and we call it first stage polygon.

2) *Second stage(refinement candidates):* Starting from the first stage polygon as shown in figure 8, we append a buffer zone using the same tile size around the first stage polygon. The result buffer is combined with the first stage polygon as shown in figure 9.

The new polygon is retiled using size $(a-1)x$ to form second stage tiles set. Feature vectors are obtained using candidate feature database $DB_{candidate}$, the new tiles community. On the other hand, the query polygon is also retiled using the same tile size as shown in figures 10(a) and 10(b). Here we use all tiles either A or B types in features extraction process and query process as first stage. Then we do our similarity measurement process using k percent nearest neighbor model and $DB_{candidate}$. The nearest tiles are then combined to form the final result polygon which will be used in the evaluation process.



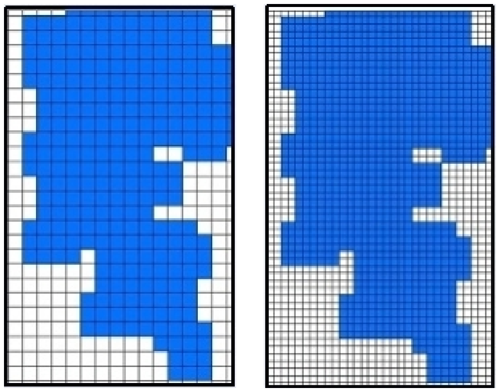
(a) (b) (c)

Fig. 9. Buffering process (a) stage one polygon (b) polygon with buffer (c) new polygon

C. Evaluation Process

Our evaluation process is carried out in terms of recall (equation 8) and precision (equation 9) using relevant areas in the database.

$$recall = \frac{correctly\ retrieved\ area}{target\ area} \tag{8}$$



(a) ax size (b) $(a-1)x$ size

Fig. 10. stage two retiling

$$precision = \frac{correctly\ retrieved\ area}{all\ retrieved\ area} \quad (9)$$

We use the map coordinates (e.g Latitude and Longitude) instead of use of file coordinates(pixels). As the map coordinates is universal and continuous where the file coordinates is file specific and dedicated. the global coordinate system is independent from the pixel size whatever the scanning satellite or stored file .

IV. Experimental Results

To evaluate the proposed system, we have used 6 bands multispectral Landsat scene which approximately covers an area of $21000 km^2$ eastern Nile Delta acquired on 2003 with a variety of landscape types as shown in figure 11. This area forms 5323 tiles with $2km$ tile size and 21098 tiles with $1km$ tile size and 84118 tiles with $0.5km$. Prior to the retrieval experiments, we have determined the different semantic classes using the polygon shapes namely, water, sand dunes, desert, vegetation, rock and urban. After selecting the semantic areas, we divided them into two types; the query and the target area. We use the k nearest neighbor(KNN) percent Euclidean similarity measure to get the nearest tiles to our query polygon, and we use the target area to evaluate our results.

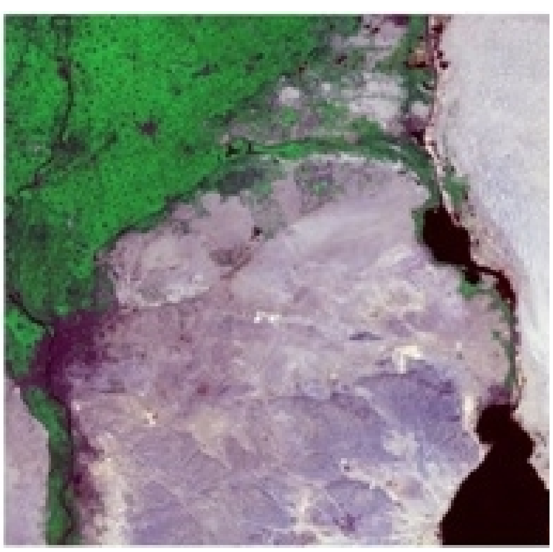


Fig. 11. Satellite image used in experimentation

V. Results Analysis

For each semantic concept we have used two polygon shapes one used as query polygon and the other used for evaluation as shown in figure 12. Taking water semantic as an example.

To evaluate our procedure we have done two experiments using our two stage procedure and the traditional one stage procedure. We determine the two polygons areas, one for query and other for the target for each semantic concept used. For the two stage procedures, 2 km tile size has been used to be the first candidate size from DB_{global} and the results are shown in table I.

The refinement stage results, using $1km$ tiles size for $DB_{candidate}$ and the middle 4%, 6% and 8% from table 1, are shown in tables II, III and IV respectively. For comparison we have performed an experiment using $1km$ tiles size to be DB_{global} and the output results are shown in table V.

The refinement process has improved the recall and precision as compared to the first stage that use $2km$ tile size as DB_{global}. In the same time it refines the precision and recall in comparison with the one stage procedure that uses $1km$ as DB_{global}. Furthermore, it reduces the retrieval time as it uses $1km$ $DB_{candidate}$ which use less than 15% of DB_{global} considering buffer area. This refinement is attributed to decreasing the tile size and utilizing a buffer which makes it more likely to find the similar adjacent semantic tiles. On the other hand; employing irregular polygon shapes guarantees a more tight covering for the required semantic type.

TABLE I
First stage recall(R) and precision (P) using 2km tile size

K%	Desert		Rock		Sand,		veget,		water	
	R	P	R	P	R	P	R	P	R	P
2	12	74	17	88	27	99	8	100	43	100
4	23	70	29	76	53	97	16	100	80	94
6	33	66	40	69	71	87	24	99	83	65
8	42	63	48	63	76	70	32	99	83	49
10	48	58	56	58	80	59	40	99	83	39

TABLE II
Second stage recall(R) and precision(P) of nearest 4% first stage

K%	Desert		Rock		Sand,		veget,		water	
	R	P	R	P	R	P	R	P	R	P
20	35	71	40	72	59	94	23	99	89	92
40	45	68	50	69	65	93	30	98	92	85
60	52	64	61	68	72	92	37	98	92	77
80	58	59	69	65	77	90	44	98	93	71
100	63	55	73	59	78	85	50	96	93	66

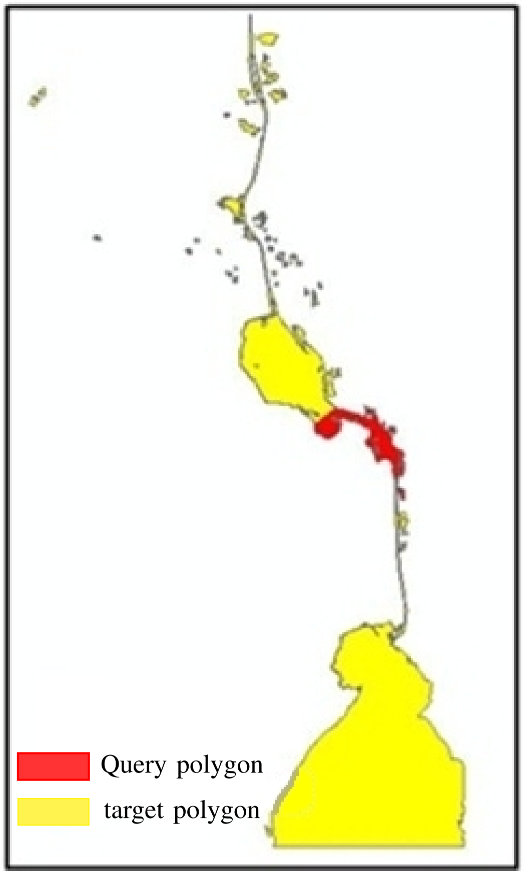


Fig. 12. Query and Target polygons

TABLE III
SECOND STAGE RECALL(R) AND PRECISION(P) OF NEAREST 6% FIRST STAGE

K%	Desert		Rock		Sand,		veget.,		water	
	R	P	R	P	R	P	R	P	R	P
20	47	68	53	66	78	84	31	98	92	57
40	57	64	66	64	83	80	38	98	92	48
60	64	59	76	61	86	74	45	98	92	41
80	69	54	81	55	88	70	53	98	92	36
100	75	51	84	49	90	65	59	96	93	33

TABLE IV
SECOND STAGE RECALL(R) AND PRECISION(P) OF NEAREST 8% FIRST STAGE

K%	Desert		Rock		Sand,		veget,		water	
	R	P	R	P	R	P	R	P	R	P
20	56	63	61	60	85	64	39	98	92	43
40	64	58	72	56	86	56	47	98	92	36
60	69	53	78	51	88	49	54	98	92	30
80	75	49	81	45	89	44	61	98	92	27
100	80	46	83	41	91	41	67	95	93	24

TABLE V
ONE STAGE RECALL(R) AND PRECISION (P) USING 1KM TILE SIZE

K%	Desert		Rock		Sand		veget.,		water	
	R	P	R	P	R	P	R	P	R	P
2	13	80	14	71	27	97	8	99	43	100
4	25	76	27	71	53	95	16	99	86	100
6	36	74	38	66	73	88	24	99	88	68
8	46	70	49	63	80	73	32	99	88	51
10	54	65	58	60	83	60	40	99	88	41

VI. CONCLUSIONS

In this paper, a new approach was developed to refine the process of satellite image retrieval using two stages; candidate selection stage and refinement stage. The refinement stage is based on the results of the candidate selection stage including the adjacent regions, where irregular polygon shapes covering the semantic types area are used. The capability of the developed system was tested using a wide area of satellite images and assessed in terms of precision and recall rates. Results show that the developed system enhanced the precision and recall with respect to the traditional approach.

REFERENCES

[1] M. Martins, L. Guimaraes, and L. Fonseca, "Texture feature neural classifier for remote sensing image retrieval systems," *XV Brazilian Symposium on Computer Graphics and Image Processing*, pp. 90 – 96,, 7-10 Oct. 2002.

[2] Y. Li and T. Bretschneider, "Semantic-sensitive satellite image retrieval," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, pp. 853 – 860, April 2007.

[3] I.Gondra and D. Heisterkamp, "Content-based image retrieval with s normalized information distance," *Computer Vision and Image Understanding*, vol. 111, pp. 219 – 228, August 2008.

[4] Y. Li and T. Bretschneider, "Semantics-based satellite image retrieval using low-level features," *IEEE International Geoscience and Remote Sensing Symposium*, vol. 7, pp. 4406 – 4409, 2004.

[5] W. Y. Ma and B. S. Manjunath, "Netra: A toolbox for navigating large image databases," *in Proc. IEEE Int. Conf. Image Process.*, pp. 568 – 571, 1997.

[6] Y. Li and T. Bretschneider, "Supervised content-based satellite image retrieval using piecewise defined signature similarities," *Geoscience and Remote Sensing Symposium 2003.*, vol. 2, pp. 734 – 736, 21-25 July 2003.

[7] Y. Li and T. Bretschneider, "Remote sensing image retrieval using a context-sensitive bayesian network with relevance feedback," *Geoscience and Remote Sensing Symposium*, pp. 2461 – 2464, August 2006.

[8] Y. Li and X. Dong, "Boosting the semantics sensitive satellite image retrieval using a voting algorithm," *Spatial Data Mining and Geographical Knowledge Services*, pp. 335 – 339, 2011.

[9] M. Ferecatu and N. Boujemaa, "interactive remote-sensing image retrieval using active relevance feedback," *Transactions on Geoscience and Remote Sensing, IEEE* ,, vol. 45, pp. 818 – 826, April 2007.

[10] L. Niu and L. Ni, "A method of remote sensing image retrieval based on roi," *Third International Conference on Information Technology and Applications Volume 2*, pp. 226 – 229, 4-7 July 2005.

[11] D. Tuia, E. Pasolli, and W. Emery, "Using active learning to adapt remote sensing image classifiers," *Remote Sensing of Environment*, vol. 115, pp. 2232 – 2242., 5 September 2011.

[12] P. Blanchart, M. Ferecatu, and M. Datcu, "Mining large satellite image repositories using semi-supervised methods," *Geoscience and Remote Sensing Symposium*, pp. 1585 – 1588, 24-29 July 2011.

[13] A. Boggess and F. J. Narcowich, *A First Course in Wavelets with Fourier Analysis*. Pearson Education, 2002.