RESEARCH ARTICLE OPEN ACCESS

A Novel Approach of Content Based Image Retrieval Improved Segmentation Algorithm for Satellite Based Images

Ms. D.N.Rajeswari[1], Ms.R.Deepthi[2]
[1]*Andhra Loyola Institute of Engineering and Technology*
[2]*Andhra Loyola Institute of Engineering and Technology*

Abstract
Now a days, satellite images plays a vital role in environmental monitoring, geological survey, and disaster forecasting and other applications. Due to the demand presence regarding remotely sensed images, many image based satellites have been launched. Each of the image based satellite produces different set of categorized images of which we require the image with Complete JPEG or high definition based image format. So as to process such type of images, we use a technique called image retrieval technique CBIR (content –based image retrieval) which gives semantic gap between low level features and high level concepts. Regional and semantic features of this type are combined to produce narrow semantic gap. In order to understand image content by region, proposal of a region level semantic based image retrieval system approach is used along with improved segmentation algorithm. Then uniform region based representation for each image is built. Based on region level features, a probabilitistic relationship among image, region and hidden semantic feature is developed. Expectation maximization method is used to mine the hidden features. Experimental results for proposed improved segmentation algorithm give better segmentation and retrieval precision than earlier multiple future fusion based algorithm.

Keywords-- Satellite based Images, Content Based Image Retrieval (CBIR), Improved Segmentation Algorithm.

I. INTRODUCTION

As satellite images are widely used, number of countries operates satellites which collect images of earth for commercial purposes. The term Satellite simply refers to a body in orbit around another body. Remote sensing is a module which focuses on Earth observation from airborne or space borne platforms. Radiation reflected or emitted from earth surface is converted to signal .The reflectance from a feature depends on the atmospheric conditions, seasons, time of a day, and physical and chemical characteristics of the feature. Selection of a remote sensing system is a compromise between spectral characteristics of surface being sensed and spectral characteristics of surface being sensed. Remote sensing used in many applications like Land-use mapping ,Forest and agriculture applications, Telecommunication planning, Environmental applications, Hydrology and coastal mapping and Urban planning. Therefore, how to retrieve an useful image fastly and accurately from a huge database becomes a challenge.

At the time many researches taken place ,there are different image query techniques used to retrieve images by matching keywords, such as geographic location, sensor type, and time of acquisition but content of image which has more priority than attributes such as width, height, number of frames is not considered in these techniques. So to overcome the difficulties present in traditional techniques content based image retrieval has become important tool for image exploration. Content-based image retrieval, a technique which uses visual contents to search images from large scale image databases according to user's interests.

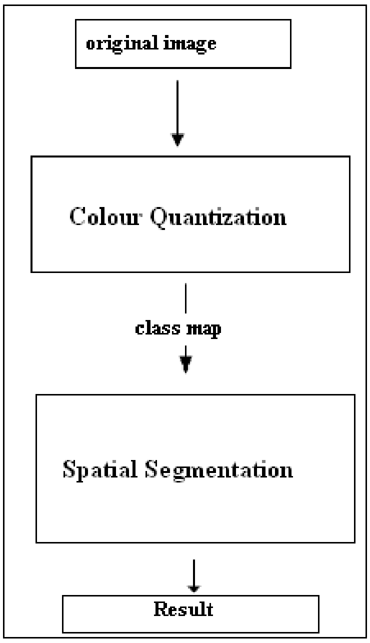
In CBIR system low-level features are used to represent image content and retrieve image form database but they cannot easily utilized to describe users perception of image. The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation. User seeks semantic similarity, but the database can only provide similarity by data processing.

Semantic feature mining is essential and performed based upon low level feature extraction and high level semantic feature extraction. Low level feature based on pixel characteristics. Li and Narayanan[1] identified ground coverage information based on spectral characteristics using supervised classification and extracted textural features by characterizing spatial information using wavelet coefficients. Pixel level textural information to extract global information does not give understanding of image thus it is often replaced with region level description of visual information.

A group of connected pixels with similar properties taken as region. For correct interpretation, image must be partitioned into regions that correspond to objects or parts of an object. Partitioning into regions done often by using gray values of the image pixels. By far, datcu et al.[2] and dascheil and datcu[3] ,aksoy et al[5] proposed pixel based Bayesian framework for narrow the semantic gap between low level features and high level concepts. So, to address the limitations in previous concepts, a novel approach is proposed to achieve region level semantic feature mining. Here, steps need to be follow for remotely sensed image retrieval is performed using region level is, identifying the region level image content to facilitate users perception then give a probability relationship among image, region, and hidden semantic features. Then, Expectation maximization(EM) method is used to mine the hidden semantic features.

II. REGION –LEVEL IMAGE REPRESENTATION USING JSEG ALGORITHM

Regional level image representation has these components: image segmentation, regional information description and codebook extraction. Region-based methods mainly rely on the assumption that the neighboring pixels within one region have similar value. The common procedure is to compare one pixel with its neighbors. If a similarity criterion is satisfied, the pixel can be set belong to the cluster as one or more of its neighbors. The selection of the similarity criterion is significant and the results are influenced by noise in all instances.



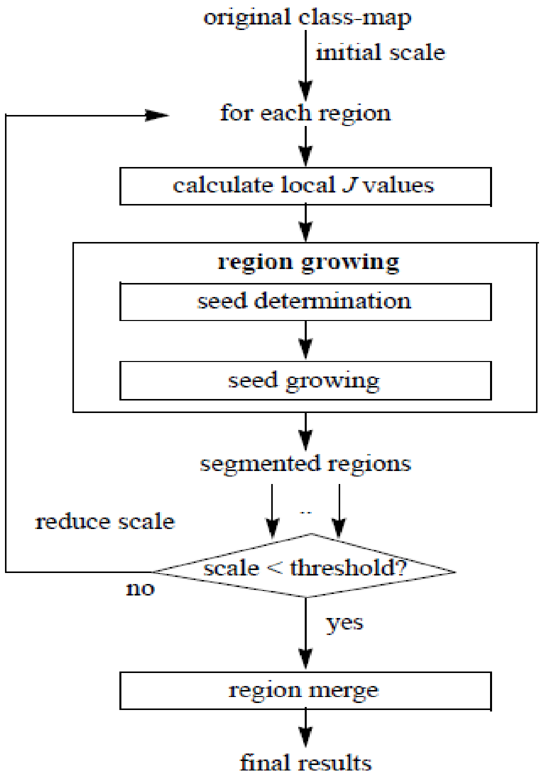


Fig 1: Original JSEG algorithm schematic diagram

Image segmentation based on region level performed by following two segmentation methods: JSEG and GrabCut.JSEG is based on the concept of region-growing. But Grab-Cut is interactive foreground/background segmentation in image. In this paper, we focused on JSEG algorithm.

The concept of the JSEG algorithm is to separate the segmentation process into two portions, color quantization and spatial segmentation. The color quantization quantizes colors in image into several representative classes that can differentiate regions in the image. The process of quantization is implemented in the color space without considering the spatial distribution of the colors. The corresponding color class labels replace the original pixel values and then create a class-map of the image.

Step 1: Calculate local j values: A color quantization algorithm is applied to image. Each pixel is assigned its corresponding color class label.

a) Estimate region by *J* value: J=(ST –SW)/SW ; Where S_T *and* S_W are an variance.
b) Total variance: ST=∑ $||Z\text{-}m||^2$ where *z* is coordinate and *m* is mean of coordinate.
c) Mean of variance of each class S_W=∑ S_i=∑∑$||Z\text{-}m||^2$ where m_i is the mean coordinate of class Z_i.
d) Segmented class-map and value J=1/N ∑ $M_K J_K$

Step 2: Seed Determination

a) Compute the average and the standard deviation of the local *J* values.
b) Set threshold $TJ=\mu_J + \alpha\sigma_J$

c) Pixels with local J values less than T_J are set as candidate seed points.

d) Associate candidate seed points as seed area if its size larger

Tabular representation for finding seed:

scale	window (pixels)	sampling (1 / pixels)	region size (pixels)	min. seed (pixels)
1	9 x 9	1 / (1 x 1)	64 x 64	32
2	17 x 17	1 / (2 x 2)	128 x 128	128
3	33 x 33	1 / (4 x 4)	256 x 256	512
4	65 x 65	1 / (8 x 8)	512 x 512	2048

Table 1: Finding min. seed and region size

Step 3: Seed Growing
a) Remove "holes" in the seed areas.
b) Compute the average of the local J values in the remaining un-segmented part of the region.
c) Connect pixels below the average to compose growing areas.
d) If a growing area is adjacent to one and only one seed, we merge it into that seed.
e) Compute local J values of the remaining un-segmented pixels at the next smaller scale and repeat region growing.
f) At the smallest scale, the remaining pixels are grown one by one.
Step 4: In spatial segmentation step, region growing method is used to segment image based on J-image, in which a threshold controls region growing result. In this, 0.5 is chosen as an empirical value.

III. PROPOSED SYSTEM ARCHITECTURE

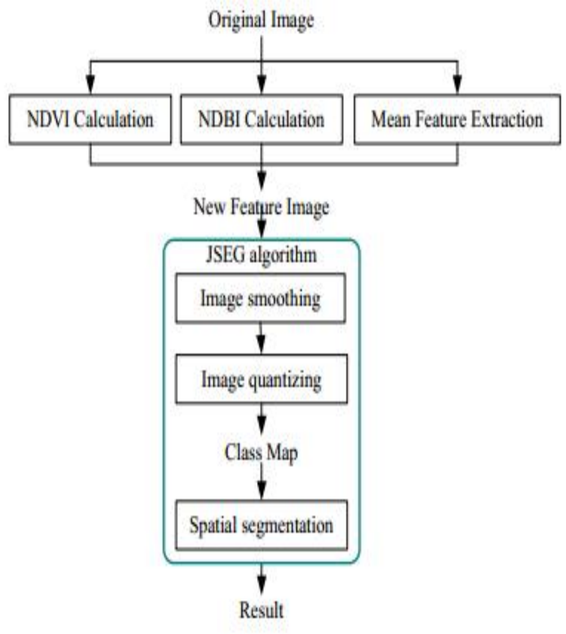


Fig 2: Flow chart of improved image segmentation algorithm

A. Image: sensors present in satellite, measures the amount of variability of the the light source. It filters the light coming to it along the path way.What is actually being measured in sensor systems is spectral radiance or the radiant energy from the target. . According to the law of energy conservation, a part of the radiant energy that arrives to the object is absorbed. The radiosity represents the light that leaves a diffuse surface (Forsyth & Ponce, 2002). This way an image is acquired

B. NDVI Calculation: It is normalized difference vegetation index, used to measure green biomass (Tucker, 1979).NDVI is calculated from the visible and near-infrared light reflected by vegetation. Healthy vegetation absorbs visible light and reflects a large portion of the near-IR light. Unhealthy or sparse vegetation reflects more visible light and less near-IR light. A typical reflectance spectrum of a vegetation can be subdivided into 3 parts, visible (0.40 – 0.70 µm), near infrared NIR (0.701 – 1.3 µm) and middle-infrared (1.301 – 2.5 µm). NDVI is calculated using an equation

$$NDVI=\frac{(NIR-VIS)}{(NIR+VIS)} \tag{1}$$

Where VIS, NIR stands for spectral reflectance acquired in the visible (red) and near infrared region. For land targets the index ranges from values close to 0 for arid or barren areas to ~ 1 for densely vegetated areas. Negative values of NDVI usually correspond to urban areas. The NDVI over water surfaces is very close to -1 due to their very low reflectance in the NIR band.

C. NDBI Calculation: Normalized Difference built up Index used to compare urban areas with built up areas between satellite images, it can be calculated as

$$NDBI = \frac{(MIR-NIR)}{(MIR+NIR)} \tag{2}$$

Where MIR and NIR stands for spectral reflectance measurements acquired in the middle infrared and near infrared regions. The middle-infrared region (1.301µm - 2.5µm) contains information about the absorption of radiation by water, cellulose and lignin and several other biochemical constituents.

Texture reflects the local variability of grey level in the spatial domain and reveals the information about the object structures in the natural environment [26]. Grey-Level Co-occurrence Matrix (GLCM), which is commonly applied in statistical procedure for interpreting texture. Finally, the pixels in original image can be represented as Equation

$$f = \{fNDVI, fNDBI, ftexture\} ; \tag{3}$$

D. GLCM: Texture is used to estimate the boundaries of objects. Since the repetitive local arrangement of intensity determines the texture, to analyze neighborhoods of pixels, to measure texture properties and texture gradient is used to estimate the orientation of surfaces.

Region information divided into two. Spectral feature is the original pixel value and textural feature is extracted using GLCM. These two features are extracted separately for each region in all images.

Step1: To compute such a matrix, we first separate the intensity in the image into a small number of different levels.

Step 2: Then we choose a displacement vector d = (dx, dy).

Step 3: The gray-level co-occurrence matrix P(a, b) is then obtained by counting all pairs of pixels separated by d having gray levels a and b.

Step 4: Afterwards, to normalize the matrix, we determine the sum across all entries and divide each entry by this sum.

Step 5: This co-occurrence matrix contains important information about the texture in the examined area of the image.

For a given co-occurrence matrix P(a, b), we can compute the following eight important characteristics:

$$\text{ENERGY}=\sum_{a,b}^{2} P(a,b) \quad (4)$$

$$\text{ENTROPY}=\sum_{a,b} p(a,b)\log_2 p(a,b) \quad (5)$$

$$\text{MAXIMUM PROBABILITY}=\max_{a,b} p(a,b) \quad (6)$$

$$\text{CONTRAST}=\sum_{a,b} |a-b|^{k} p(a,b), \text{usually } k=2, \lambda=1 \quad (7)$$

$$\text{INVERSE DIFFERENCE MOMENT}=\sum_{a,b} \frac{P(a,b)^{\lambda}}{|a-b|^{k}} \quad (8)$$

$$\text{CORRELATION}=\frac{\sum_{a,b} |(ab)P(a,b)| - \mu_x \mu_y}{\sigma_x \, \sigma_y} \quad (9)$$

IV. JSEG ALGORITHM

According to the characteristics of the J values, the modified region growing method can be applied to segment an image. The algorithm starts the segmentation at the largest scale. Then it repeats the same process on the newly segmented regions at the next lower scale. After finishing the final segmentation at the smallest scale, the region merging operation follows region growing to derive the final segmentation result.

V. CODEBOOK EXTRACTION

Calculation of similarity between regional features for all pairs of regions is time-consuming task. Most of time regions on different images with similar spatial and spectral features are identified. So, GLA(Generalized Lloyd Algorithm)is used to classify the low-level features into a set of codes based on which a codebook will be generated.GLA for codebook generation is that for power law distributed data, more cluster centers are generated around high density areas Jurie & Triggs (2005). Our experiments also confirmed such phenomenon, this can cause the codewords to be overpopulated in the high density area while the sparse area not properly represented. To solve this problem, we propose to merge these very close cluster centers ,first generate more centers than what we expect, and then we merge these centers into the number of centers we desire. Through merging close centers, we can avoid cluster centers to crowd into high density areas.

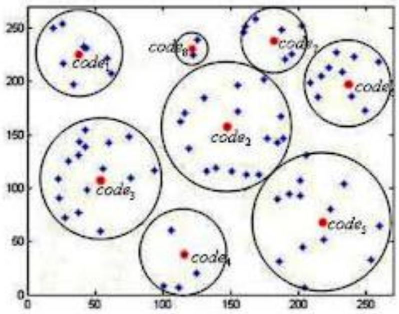


Fig 3: Schematic diagram of codebook extraction

VI. SEMANTIC FEATURE EXTRACTION

EM is used to compute maximum likelihood estimates given incomplete samples. The EM algorithm estimates the parameters of a model iteratively. Starting from some initial guess, each iteration consists of an E step (Expectation step) and an M step (Maximization step).

First, various parameters are defined as follows:

a) Image data: dj is an image in the database, dj ={d1,...,dM}; M is the total number of images.

b) Regional feature data: ri is the ith region feature in the feature codebook, ri = R = {r1,...,rN}, where N is the total number of regional features.

c) Hidden semantic features: sk is the hidden semantic feature, sk; S = {s1,...,sK}, where K is the total number of semantic features.

P(dj) denotes the probability that an image will occur in a particular image database. P(ri|sk) denotes the class conditional probability of region ri given the hidden semantic feature sk. P(sk|dj) denotes the class-conditional probability of the hidden semantic feature sk given a particular image dj. Dj and ri are independently defined on the state of the associated hidden semantic feature.

Step: Initialize K clusters: $C_1, \ldots, C_K$ (μ_j, Σ_j) and $P(C_j)$ for each cluster j.

Iteration Step: Estimate the cluster of each data point. Then Re-estimate the cluster parameters.

Joint probability of di and ri = P(ri,di)=P(di)P(ri/di)

Then apply total probability form: P(di)P(ri/di)= $P(di)\sum_{k=1}^{k} P(ri/sk)\, P(sk/dj)$

Using Bayesian formula ,class conditional probability

$$P(sk|ri,di)= \frac{P(dj|sk)P(ri|sk)P(sk)}{p(dj)\sum_{K=1}^{K} p(ri|sk)P(sk|dj)} \quad (10)$$

We have reduced the problem of selecting the best model to that of selecting the best parameter and want to select a parameter p which will maximize the probability that the data was generated from the model with the parameter p plugged-in. The parameter p is called the maximum likelihood estimator. The maximum of the function can be obtained by setting the derivatives of the function ==0 and solving for p.

Maximum likelihood function: L=log(P(R,D,S))

$$=\sum_{i=1}^{N}\sum_{i=1}^{M} P(sk|ri,di) \quad \log[P(sk,dj)P(ri|sk)]$$

Standard procedure for maximum likelihood estimation depends on E-step which can be interpreted as mining the relationship between current estimates of the parameters and the latent variables by computing posterior probabilities .M-step can be interpreted as updating parameters based on the so called expected complete data –likelihood.

$$P(ri|sk) = \frac{\sum_{j=1}^{M} n(ri,dj)\, P(sk|ri,dj)}{\sum_{m=1}^{N} \sum_{j=1}^{M} n(rm,dj)P(sk|rm,dj)}$$

$$P(sk|dj)= \frac{\sum_{i=1}^{N} n(ri,dj)P(sk|ri,dj)}{\sum_{n=1}^{N} n(rn,dj)} \quad (11)$$

VII. EXPERIMENTAL RESULTS

Different scenes of TM images are utilized, Each image is split into small size of 256*256, and total number of small size images is 2500. After preprocessing, each image and its features are stores in the image database. In this image database, each image is classified into seven concepts: cloud, sea, river, mountain, urban area, farm land, bare soil.

A. Image segmentation

JSEG: Color quantization threshold - specify values 0-600, leave blank for automatic determination. The higher the value, the less number of quantized colors in the image. For color images, try 250. If you are unsatisfied with the result because two neighboring regions with similar colors are not getting separated, please try a smaller value say 150a) Number of scales - The algorithm automatically determines the starting scale based on the image size and reduces the scale to Refine the segmentation results. If you want to segment a small object in a large-sized image, use more number of scales. If you want to have a coarse segmentation, use 1 scale only .Region merge threshold specify values 0.0-0.7, leave blank for default value 0.4. If there are two neighboring regions having identical color, try smaller values to avoid the merging.

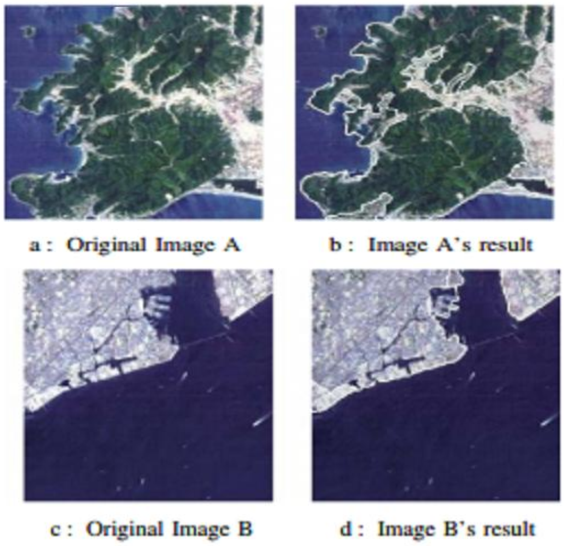


Fig 4: Images A,B and their segmentation results using JSEG algorithm

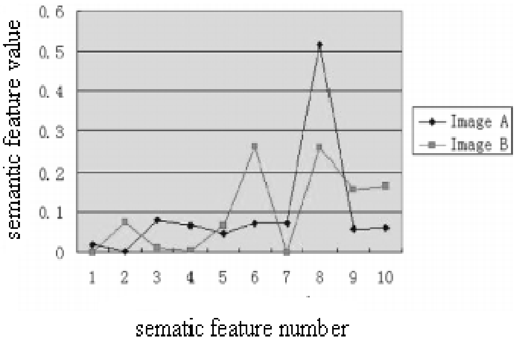


Fig 5: Semantic feature extraction

(a)

(b)

Fig 10 : a) image covers mountain and urban area b) top 20 results.

$$\text{precision}=\frac{\text{Irelevent n Iretrieved}}{\text{Iretrieved}}$$

$$\text{recall}=\frac{\text{Irelevent n I retrieved}}{\text{Irelevent}}$$

Where Irrelevant is the total number of relevant images, and I retrieved is the total number of retrieved images. When the numbers of total retrieved images are 10, 20, 30, and 40, respectively. Although the precision and recall of first experiment are much higher than those of second experiment.

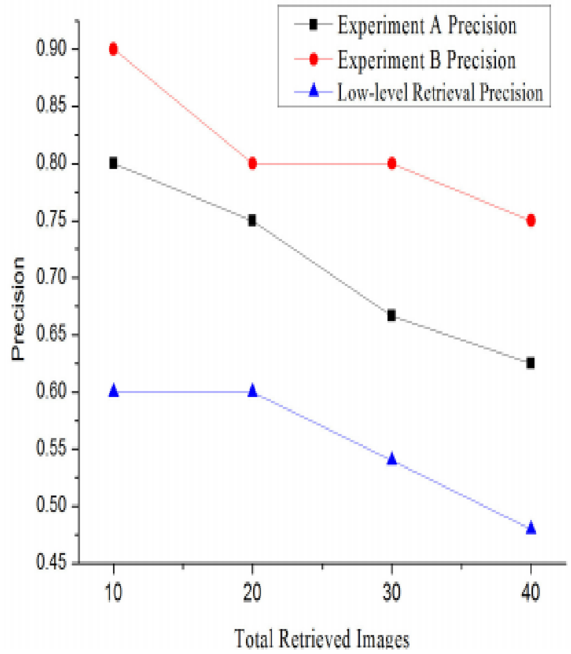


Fig 11:The precision results for different number of retrieved images taken above.

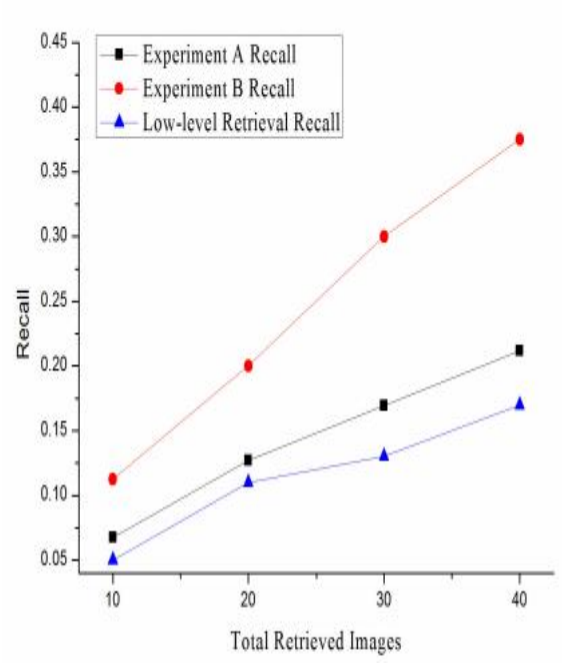


Fig 12: The recall results for different number of retrieved images taken above.

VIII. CONCLUSION

Different image query techniques are used to retrieve images, to overcome the difficulties present in traditional retrieval techniques content based image retrieval has become tool for image exploration. In CBIR system, cannot easily utilized to describe users perception of image. Semantic gap occur due to the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data given for a user. Regional and semantic features are combined to narrow semantic gap. Improved segmentation algorithm along with Expectation maximization method gives better precision results.

REFERENCES

[1] Li J, Narayanan RM: Integrated spectral and spatial information mining in remote sensing imagery. IEEE Trans Geosci Remote Sens 2004,42(3):673-685.

[2] Datcu M, Daschiel H, Pelizzari A, Quartulli M, Galoppo A, Colapicchioni A,Pastori M, Seidel K, Marchetti PG, D'Elia S: Information mining in remote sensing image archives: system concepts. IEEE Trans Geosci Remote Sens2003, 41(12):2923-2936.

[3] Daschiel H, Datcu M: Design and evaluation of human-machine communication for image information mining. IEEE Trans Multimedia 2005, 7(6):1036-1046.

[4] Deng YN, Manjunath BS: Unsupervised segmentation of color-texture regions in images and video. IEEE Trans Pattern Anal Mach Intell 2001,23(8):800-810.

[5] Aksoy S, Koperski K, Tusk C, Marchisio G, Tilton JC: Learning Bayesian classifiers for scene classification with a visual grammar. IEEE Trans GeosciRemote Sens 2005, 43(3):581-589.

[6] Bimbo AD: Visual Information Retrieval Morgan 19. Wang ZY, Boesch R: Color- and texture-based image segmentation for improved forest delineation. IEEE Trans Geosci Remote Sens 2007,45(10):3055-3062.

[7] Chang Y, Lee D, Wang Y: Color-texture segmentation of medical images based on local contrast information. Proc IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology Honolulu, HI; 2007, 488-493.

[8] Han JW, Ngan KN: Automatic segmentation of objects of interest in video: a unified framework. Proceeding of 2004 Intelligent Signal Processing and Communication Systems (ISPACS) Seoul, South Korea; 2004, 375-378.