International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869 (O) 2454-4698 (P), Volume-3, Issue-7, July 2015

Remote Sensing Image Retrieval using High Level Colour and Texture Features

Gauri Sudhir Mhatre, Prof. M.B. Zalte

Abstract— **The whole world has been digitized by multiple satellite remote-sensing systems. Enormous number of remote sensing images has been collected by satellites. This leads to an exponential increase in the quantity of remotely sensed images in database. These images are being used for environmental monitoring, disaster forecasting, geological survey, and other applications. Accurate and quick retrieval of satellite images from huge databases is a challenge. This paper implements a prototype system for remote sensing image retrieval system using different high level feature of images through five different methods, three were based on analysis of color feature and other two based on analysis of texture features. Experiment was carried out on test images database and results are analysed with the help of retrieval accuracy, classification, confusion matrix and precision and recall plot.**

Index Terms—**Remote sensing, Color autocorrelogram, Color moments, DWT, Gabor filter, HSV histogram.**

I. INTRODUCTION

Content-based image retrieval measures the visual similarity between a query image and database images. The retrieval result is an images similar with the query image. By extracting the feature vectors of the query image and the database images, there is need to develop similarity measures that will rank the database images by the actual distance between their vectors and the query image vector. Content Based Image Retrieval systems think that the best retrieving images are the most visually similar images to a given query image from a large collection of images. It is index visual characteristics of an image, such as its color, textures and shape to look for an explicit image in a large amount of images.

Advancement in satellite system has resulted in diverse remote sensed data which is increasing rapidly. Remote Sensing Image Retrieval is an application of CBIR techniques to remote sensing archive. Satellite images represent large amount of complex geographical data. For Spatial information retrieval consideration of low level features of images would not be sufficient.

This paper proposes a methodology to retrieve RS images by applying high level features. RS image is first segmented and over segmented regions are merged to locate region of interest and extract visual features. Then region features are extracted in the form of color and texture.

The reminder of this paper is organized as follows. In section 2, a brief discussion about the research works related to the content based image retrieval in remote sensing images is given. Section 3 introduces our proposed method. Implementation and results are shown in section 4. Conclusion of the paper is given in section 5.

II. LITERATURE REVIEW

A comprehensive study of the content based image retrieval has been done in the past few years. The survey includes 100+ papers covering the research aspects of many authors. Yong Rui, Thomas S. Huang, Shih-Fu Chang has reviewed image feature representation and extraction, multi-dimensional indexing, and system design, three of the fundamental bases of content-based image retrieval [1]. K. Vijay Kumar has provided a comprehensive survey of the recent technical achievements in high-level semantic-based image retrieval in [2].

Tingting Liu, Liangpei Zhang, Pingxiang Li and Hui Lin have implemented the approach of semantic mining [3]. Homogeneous spectral and textural characteristics are extracted from image regions, and then a uniform region-based representation for each image is built. In [13] authors have designed an object-based confusion matrix (OCM) classification accuracy assessment scheme to accurately estimate the overall and individual category classification accuracy. Neera Lal, Neetesh Gupta, Amit Sinhal in their paper [11] have compared the techniques of image classification in CBIR and also introduced classifiers like support vector machine, Bayesian classifier for accurate and efficient retrieval of images. Many authors have worked on high level color and texture feature extraction [5-10].

III. METHODOLOGY

Remote sensing image retrieval systems present mechanisms for choosing the data items that resemble most a query image among all the accessible data in a database. The proposed system will be mainly comprised of six stages. Fig.1 shows flow of retrieval process.

Gauri S. Mhatre, Department of Electronics and Telecommunication Engineering, K.J.Somaiya College of Engineering, Mumbai, India, 9819265576.
Prof. M.B. Zalte, Department of Electronics and Telecommunication Engineering, K.J.Somaiya College of Engineering, Mumbai, India, 9821132516.

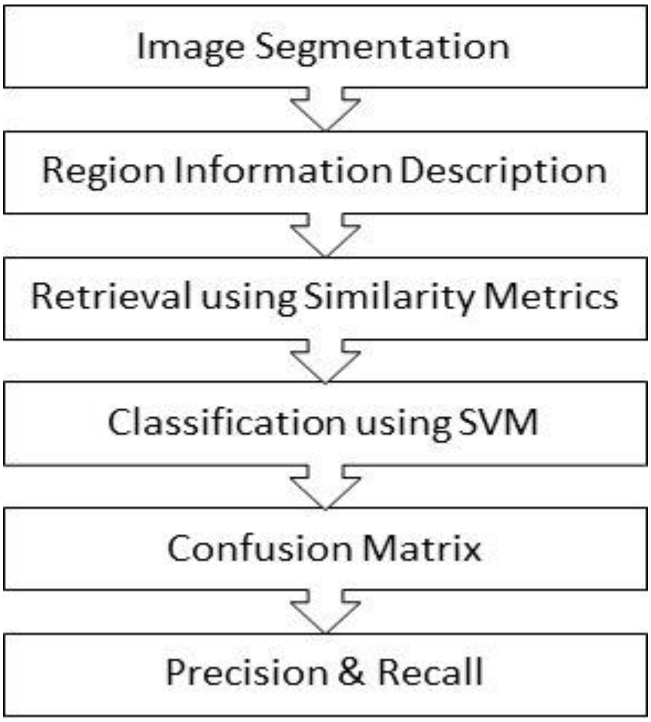

Fig 1. Methodology for image retrieval

A. *Image Segmentation*

Image segmentation is considered as a corner stone of any image retrieval system. Segmentation refers to the process of partitioning a digital image into multiple segments. Satellite images are rich in color and texture. It is difficult to identify image regions containing color-texture patterns. JSEG [4] is a region based segmentation algorithm for segmentation of color images. JSEG stands for j value segmentation. JSEG is based on the perception of region growing. It is robust system of segmenting natural images. The JSEG algorithm simplifies color and quality of images.

Here colors in the image are first quantized and after quantization, the quantized colors are assigned labels. A color class is the set of image pixels quantized to the same color. The image pixel colors are replaced by their corresponding color class labels. The new constructed image of labels is called a class-map. J images are grey scale images then formed by applying windows of different sizes to class map. Next step is spatial segmentation where region growing method is used on multi-scale J images. After segmentation last step is region merge to avoid over segmentation and we get final segmented image.

B. *Region Information Description*

The natural scene description means to describe the region, or the area, by regional shape descriptors. Regional information in RS images is described using color and texture features. Color features are extracted using Color Autocorrelogram, Color Moments and HSV Histogram whereas texture features are extracted using Gabor filters and Discrete Wavelet Transform. These features are extracted separately for each region in all images.

1) *Color Autocorrelogram*

This is the most common second order statistical measures used in image retrieval are based on the correlation function between the image pixels. Image correlogram describes the correlation of the image colors as a function of their spatial distance. The autocorrelogram is a subset of the correlogram, and it gives the probability of finding identical colors at certain distance [6].

The definition of the correlogram is the following . Let [D] denote a set of D fixed distances {d1,,…, dD}. Then the correlogram of the image I is defined for level pair (gi, gj) at a distance d.

$$\gamma^{(d)}_{(g_i,g_j)}(I) = P_{r_{P1\in I, p2\in I}}\left[\left(p_2 \in I \,\|\, p_1 - p_2 = d\right|\right)\right] \tag{1}$$

which gives the probability that given any pixel p1 of level gi, a pixel p2 at a distance d in certain direction from the given pixel p1 is of level gi. Autocorrelogram captures the spatial correlation of identical levels only:

$$\alpha_g^{(d)}(I) = \gamma_{g,g}^{(d)}(I) \tag{2}$$

It gives the probability that pixels p1 and p2, d away from each other, are of the same level gi. In this work, we considered the color autocorrelogram with distances d = [1,3,5,7] [5].

2) *Color Moments*

Color moments are measures that characterize color distribution in an image. Color moments are scaling and rotation invariant. We have computed first two color moments Mean and Deviation from each R, G, and B channel.

Mean

The first color moment can be interpreted as the average color in the image, and it can be calculated by using the following formula,

$$E_i = \sum_{j=1}^{N} \frac{1}{N} P_{ij} \tag{3}$$

where N is the number of pixels in the image and P_{ij} is the value of the j-th pixel of the image at the i-th color channel.

Standard Deviation

The second color moment is the standard deviation, which is obtained by taking the square root of the variance of the color distribution.

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N}(P_{ij} - E_i)^\wedge 2\right)} \tag{4}$$

where E_i is the mean value, or first color moment, for the i-th color channel of the image.

3) *HSV Histogram*

A color histogram is the proportion of the number of different types of colors, regardless of the spatial location of the colors. During this step following actions are performed: Color Space Conversion, Color Quantization and Compute Histogram.

First RGB color space is converted into HSV color space. In the HSV color space, hue is used to distinguish colors, saturation is the percentage of white light added to a pure color and value refers to the perceived light intensity. In color quantization, for every image in the database, colors in the HSV model are quantized, to make later computations easier. Color quantization reduces the number of distinct colors used in an image. Each component is quantized with non-equal intervals: H: 8 bins; S: 2 bins and V: 2 bins. Finally we concatenate 8X2X2 histogram and get 32-dimensional vector. If we use direct values of H, S and V components to represent the color feature, it requires lot of computation. So it is better to quantify the HSV color space into non-equal intervals. At the same time, because the power of human eye to distinguish colors is limited, we do not need to calculate all segments. Unequal interval quantization according the human color perception has been applied on H, S, and V components [8].

4) Gabor Filters

The extraction of texture of an image is accomplished by using a set of Gabor Filters. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. Gabor filters are a group of wavelets capturing energy at a specific frequency and a specific direction. The expansion of a signal using this basis provides a localized frequency description, therefore, capturing local features/energy of the signal. Texture features can thus be extracted from this group of energy distributions [7].

For a given image I(x,y) with size PXQ, its discrete Gabor wavelet transform is given by a convolution:

$$G_{mn}(x,y) = \sum_s \sum_t I\,(x-s, y-t)\varphi^*_{mn}(s,t) \tag{5}$$

Where, s and t are the filter mask size variables, and φ^*_{mn} is a complex conjugate of φ_{mn} which is a class of self-similar functions generated from dilation and rotation of the following mother wavelet:

$$\varphi_{mn}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right].\exp(j2\pi Wx) \tag{6}$$

where W is called the modulation frequency. The self-similar Gabor wavelets are obtained through the generating function:

$$\varphi_{mn}(x,y) = a^{-m}\varphi\,(x,y) \tag{7}$$

Where m and n specify the scale and orientation of the wavelet respectively, with m=0,1,……….M-1,n= 0,1,……….N-1.

After applying Gabor filters on the image with different orientation at different scale, we obtain an array of magnitudes:

$$E\,(m,n) = \sum_x \sum_y |G_{mn}\,(x,y)| \tag{8}$$

These magnitudes represent the energy content at different scale and orientation of the image. The main purpose of texture-based retrieval is to find images or regions with similar texture. It is assumed that we are interested in images or regions that have homogenous texture, therefore the following mean μ_{mn} and standard deviation σ_{mn} of the magnitude of the transformed coefficients are used to represent the homogenous texture feature of the region:

$$\mu_{mn} = \frac{E\,(m,n)}{PXQ} \tag{9}$$

$$\sigma_{mn} = \frac{\sqrt{\sum_x\sum_y(|G_{mn}(x,y)| - \mu_m)\text{^}2}}{PXQ} \tag{10}$$

A feature vector is created using μ_{mn} and σ_{mn} as the feature components. Four scales and Six orientations are used in common implementation and the feature vector of length 48 is given by:

$$f_g = \{\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots \dots, \mu_{35}, \sigma_{35}\} \tag{11}$$

5) Discrete Wavelet Transform

DWT can be performed by iteratively filtering a signal or image through the low - pass and high - pass filters, and subsequently down sampling the filtered data by two [10]. This process will decompose the input image into a series of sub band images.

The DWT decomposition of the image is applied up to third level. Wavelet Statistical Features for each level, for each sub band (High-High, High-Low, Low-High, Low-Low) are calculated. First two moments [9] of wavelet coefficients i.e. Mean and Standard coefficient are to be considered.

Mean: The mean is measurement of average intensity level in that sub band.

$$\text{Mean} = \frac{1}{N^2}\sum_{i,j=1}^{N} C(i,j) \tag{12}$$

Where C(i,j) is the transformed value in (i,j) for any subband of size N x N

Standard Deviation: The standard deviation of the image gives a measure of the amount of detail in that sub band.

$$\text{SD} = \sqrt{\frac{1}{N^2}\sum_{i,j=1}^{N}[C(i,j) - m]\text{^}2} \tag{13}$$

Where C(i,j) is the transformed value in (i,j) for any subband of size N x N and m is mean.

C. Retrieval using Similarity Metrics

To perform the actual image retrieval we investigated a number of vector distance measures to discover which gave the most accurate and perceptually correct result. Similarity metrics used are L1, L2, Standardized L2, Cityblock, Minkowski, Chebyshev, Cosine, Correlation, Spearman, Normalized L2 and Relative deviation.

D. Classification using SVM

Support vector machine is a supervised learning technique that analyzes data and identify pattern used for classification. It is different type of classifier that performs classification by generating a hyper plane. The optimal hyperplane separates data in two categories. The aim of support vector machine is to find the optimal hyperplane that is used to distinguish group of vectors in a way that one category of the required variables is on one side of hyper plane and the other class of variables are on the other side of plane. Vectors nearer to the hyper plane are called as support vectors [11]. The optimum hyper plane can be defined as the linear classifier with the maximum margin for a given set of variables. SVM is one of the best known methods in pattern classification and image classification.

It is designed to separate of a set of training images two different classes, (x1, y1), (x2, y2), .., (xn, yn) where xi in Rd, d-dimensional feature space, and yi in {-1,+1}, the class label, with i=1..n [12]. SVM builds the optimal separating hyper planes based on a kernel function (K).

E. Confusion Matrix

A confusion matrix, also known as an error matrix, visualizes the performance of an algorithm, typically a supervised learning one [13]. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. The name stems from the fact that it makes it easy to see if the system is confusing two classes (i.e. commonly mislabeling one as another). It is a table with n rows and n columns that reports the number of false positives, false negatives, true positives, and true negatives where n is number of classes. This allows more detailed analysis than mere proportion of correct guesses (accuracy).

F. Precision & Recall

In information retrieval, precision is the fraction of retrieved instances that are relevant, while recall is the fraction of relevant instances that are retrieved. Both precision and recall are therefore based on an understanding and measure of relevance.

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total retrieved}}$$

$$\text{Recall} = \frac{\text{Total number of relevant images}}{\text{Number of relevant images retrieved}}$$

IV. Implementation and Results

Above algorithm is tested on the database of 400 images. Images in the database are collected from different open sources available on internet. All these images are classified into 4 different classes namely Land & Water, Beach, Mountain, and Dainos.

Fig 2 below shows retrieval results for Land & Water class. Fig 3 shows classification result for query image in fig 2. Fig 4 is the confusion matrix for query image. Fig 5 is Precision & Recall plot of proposed system.

Table I summarizes Retrieval accuracy of different classes using various distance measures.

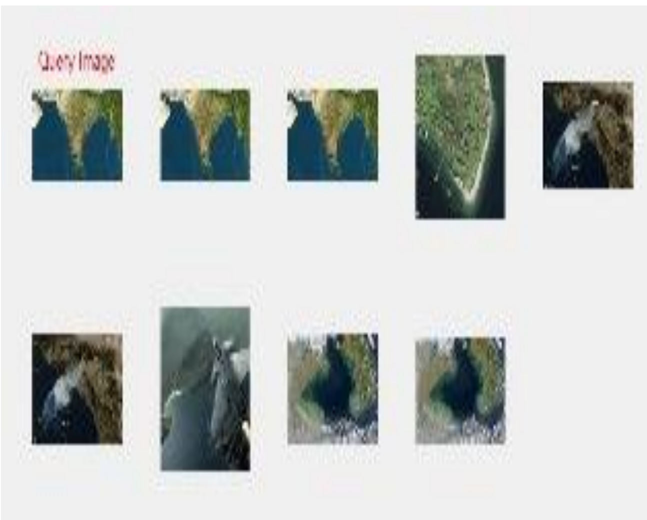

Fig.2 Image retrieval results

SVM (1-against-1):
Accuracy = 91.60%
Confusion Matrix:

41	4	0	0	5	0	0	0	0	0
1	44	0	0	5	0	0	0	0	0
2	0	48	0	0	0	0	0	0	0
0	0	0	50	0	0	0	0	0	0
3	15	0	0	32	0	0	0	0	0
0	1	0	0	0	49	0	0	0	0
1	1	0	0	2	0	46	0	0	0
0	0	0	0	0	0	0	50	0	0
0	0	0	0	0	0	0	0	50	0
2	0	0	0	0	0	0	0	0	48

Predicted Query Image Belongs to Class = 1

Fig 3. Classification Results

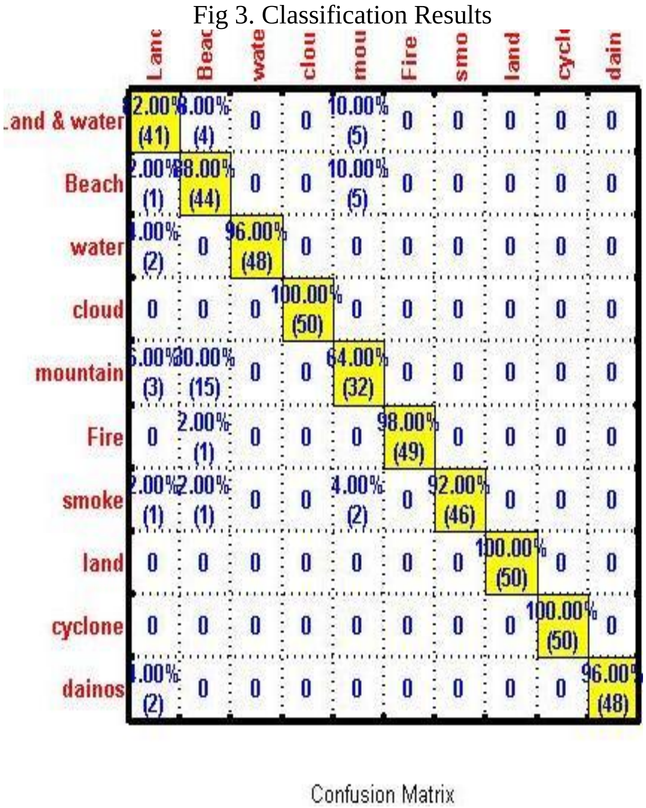

Fig 4. Confusion Matrix

Table I. Retrieval accuracy

	Land & Water	Beach	Mountain	Dainos
L1	89.40%	92.00%	92.00%	92.00%
L2	91.80%	91.00%	90.60%	91.80%
Std.L2	91.60%	89.20%	92.20%	91.40%
Cityblock	90.40%	93.00%	90.80%	91.40%
Minkowski	91.60%	91.40%	91.00%	91.80%
Chebyshev	90.80%	91.40%	90.60%	92.40%
Cosine	88.80%	91.20%	90.60%	90.60%
Correlation	90.20%	91.00%	89.80%	91.80%
Spearman	92.00%	89.80%	91.60%	91.40%
Norm.L2	90.60%	92.00%	91.80%	92.40%
Relative Deviation	91.80%	90.6%	91.80%	93.00%

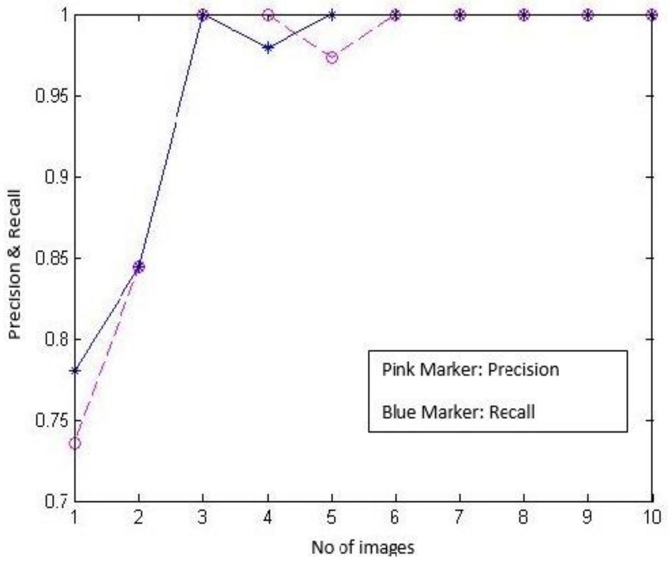

Fig 5. Precision & Recall Plot

V. Conclusion

In this paper Remote Sensing Image Retrieval is proposed and implemented for accurate image search as per user interest from digital image database. Images are first segmented by using JSEG color segmentation method and high level color and texture features are extracted from each region. During retrieval phase feature vector of query image is compared with that of database images using distance metrics. All the distance metrics give more than 90% accuracy on an average. Retrieval accuracy for different classes using different similarity matrices is compared. Precision-Recall plot for proposed system shows good retrieval accuracy once system has been trained after few initial retrievals. The idea in this paper is most applicable and shows good retrieval results.

References

[1] Yong Rui, Thomas S. Huang, Shih-Fu Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues," Journal of Visual Communication and Image Representation 10, 39–62 (1999).

[2] K. Vijay Kumar, "Content Based Image Retrieval System Consume Semantic Gap", International Journal of Computer Science and Information Technologies, Vol. 3 (5), 2012, 5231-5235.

[3] Tingting Liu, Liangpei Zhang, Pingxiang Li and Hui Lin, "Remotely sensed image retrieval based on region-level semantic mining,"Liu et al. EURASIP Journal on Image and Video Processing 2012.

[4] Yining Deng, B. S. Manjunath and Hyundoo Shin, "Color Image Segmentation,"Department of Electrical and Computer Engineering University of California, Santa Barbara, CA.

[5] Iivari Kunttu, Leena Lepistö, Ari Visa, "Image Correlogram in Image Database Indexing and Retrieval."

[6] Huang, J. Et Al. (1997). Image Indexing Using Color Correlograms. Proc. Ieee Computer Society Conference On Computer Vision And Pattern Recognition, San Juan, Puerto Rico, P. 762-768.

[7] S. Mangijao Singh , K. Hemachandran, "Content--Based Image Retrieval using Color Moment and Gabor Texture Feature," IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012.

[8] Deepak John, S.T. Tharani, Sreekumar K, "Content Based Image Retreival using HSV-Color Histogram and GLCM," International Journal of Advance Research in Computer Science and Management Studies, Volume 2, Issue 1, January 2014.

[9] D. A. Doshi, A M. Kothari, D G. Kamdar, "Feature Extraction for Texture classification – An Approach with Discrete Wavelet Transform," International Journal Of Darshan Institute On Engineering Research & Emerging Technologies Vol. 2, No. 1, 2013.

[10] Marcin Kociołek, Andrzej Materka, "Discrete Wavelet Transform – Derived Features For Digital Image Texture Analysis," Proc. of Interational Conference on Signals and Electronic Systems, 18-21 September 2001, Lodz, Poland, pp. 163-168.

[11] Neera Lal, Neetesh Gupta, Amit Sinhal, "A Review of Image Classification Techniques in Content Based Image Retrieval," International Journal of Computer Science and Information Technologies, Vol. 3 (5) , 2012,5182 – 5184.

[12] Lei Zhang, Fuzong Lin, Bo Zhang, "Support vector machine learning for image retrieval,"IEEE, 2001.

[13] Lina Yi, Guifeng Zhang, "Object-Oriented Remote Sensing Imagery Classification Accuracy Assessment Based on Confusion Matrix," IEEE,2012.

Gauri S. Mhatre is Master of Engineering student in Electronics and Telecommunication Engineering from K.J. Somaiya College of Engineering, Mumbai.

Prof. M. B. Zalte is currently working as Associate Professor in Electronics and Telecommunication Engineering department of K.J. Somaiya College of Engineering, Mumbai.