Volume 3, Issue 9, September 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Classification of Remote Sensing Image Areas Using Surf Features and Latent Dirichlet Allocation

Ms. Chandrakala.M*[1] , Mrs. R.Amsaveni[2]
Department of Computer Science[1],
Department of Information Technology[2],
PSGR Krishnammal College for Women
Bharathiar University, Coimbatore, Tamilnadu, India.

Abstract— In this paper we are interested in classifying the satellite images into 3 different regions as water, urban and greenland. The process involves two steps, first training the class images and second classifying the testing image which consist of all the classes based on training image. Speed Up Robust Features (SURF) are used to enhance the performance over low level feature like mean and standard deviation. Topic modelling concept is used to obtain Bag of Features (BoF) with Latent Dirichlet Allocation (LDA) algorithm. Threshold value for each class is obtained from BoF and compared with testing image feature values in order to classify it. Experiments are conducted on LANDSAT 7 images obtained from google earth.

Keywords— SURF, LDA, Remote sensing, Satellite Images, Image Classification

I. INTRODUCTION

Remote sensing is a process of gathering information about an object or area without being in direct contact with it. Thus the sensors are not in direct contact. The information needs a physical medium to travel from the objects/areas to sensors. The electromagnetic radiation is used as an information carrier in remote sensing. In general sense Remote Sensing refers to acquisition of earth information. The output of a remote sensing system is an image representing the scene being observed usually called as Remote Sensing Images.

Image classification is an important part of the remote sensing, image analysis and pattern recognition. Image classification is used to collect data about earth, produces a map like image as a result. Thus image classification is an important tool for examination of the satellite images. Image Classification can be divided into two broad categories as Supervised Classification and Unsupervised Classification. In Supervised Classification, training data is selected with known identities and corresponds to the classes. The algorithm classifies the image to the classes depending on the training data. In Unsupervised Classification, the image is segmented into classes based on natural groupings based on the data found in that image. After the formation of groups it is given with class name [1]. Section II explains a detailed background study for this paper.

Section III gives a review on Literature Survey. Section IV describes the proposed work and Section V describes experiments and results. Section VI and VII furnish conclusion and Future enhancement.

II. BACKGROUND STUDY

This section deals with the background study for this work.

A. Image Registration

Image registration is a discipline under digital image processing that studies how to map similar points of source image to reference image taken at different times or different angles or different sensors. Accurate registration algorithms are essential for creation of mosaics satellite images, change detection on the planet's surface over time, image classification, environmental monitoring and weather forecasting. There are many image registration methods, which condense it into the following eight categories: image dimensionality, registration based, geometrical transformation, degree of interaction, optimization procedure, modalities, subject, and object [2]-[5]. In recent years, many attempts are developed for automatic registration algorithm for remote sensing images. Scale Invariant Feature Transform (SIFT) based on a scale restriction and Speeded Up Robust Features (SURF) algorithm with a higher correct matching rate are used for registration of satellite images.

B. SIFT

Scale-Invariant Feature Transform (or SIFT) is used to detect and describe local features in images. There are two steps in detection of feature, first finding the interest points of the image which consist of meaningful structures. The Difference of Gaussian (DoG) is calculated for all location in the image with different scales and compared with each other. Second is to calculate scale invariant descriptor on each interest point. For rotation invariant a rectangle which is proportional to the scale of interest point is assigned and reduced to 4X4 grid. Interest point descriptor is constructed by subtracting gradient or absolute value of gradient from sub square.

C. SURF

Speed Up Robust Features (or SURF) is a scale and rotation invariant interest point detector and descriptor. The SURF feature is a speed up version of SIFT. It uses an approximated DoG and the integral image to increase the speed. An integral image is calculated as each pixel value is the sum of all the original pixel values left and above it. Block process can be computed between any 2 blocks of an integral image with 6 calculations. Thus SURF features is faster than the SIFT features.

D. LDA

Latent Dirichlet Allocation (LDA) is simple and most popular topic modeling algorithm. D is a collection of documents which consist of words. LDA is a generative model which describes how each document obtains its words. By assuming K topics in document with V elements. Let β_i be the multinomial for the i-th topic, where the size of β_i is V: |β_i | = V. By this, the LDA generative process is given as follows:

1. For each document:
    a. Choose a random distribution over topics
    b. For each word in the document:
        i. One of the K topics from the distribution over topics obtained in (a) is taken using probability, say topic β_j.
        ii. One of the V words from β_j is taken using probability.

Thus the goal of topic modeling is to discover the topics automatically from a collection of documents.

III. LITERATURE REVIEW

Teodor Costachioiu, Rodica Constantinescu, Bashar AlZenk and Mihai Datcu uses LDA for an unsupervised analysis of satellite image time series. This method is used at pixel level, for preserving the spatial resolution of the original time series [6]. Tang Yingjun has proposed a model to learn and recognize natural scene category with more precise prior of topic distribution by doing double inference [7]. Diane J. Hu uses the LDA model for text, image and music. For text, topics and word dependencies are accurately determined which increases the performance of document classification. The LDA is extended to model the images and used for object categorization and object localization. It is also used in harmonic analysis. The musical notes that occur within the same segment come under the same topic [14]. Pradheep K Elango and Karthik Jayaraman uses the LDA model for modeling the inter-segment relationships and inter-image relationships are modeled in an unsupervised manner and clustered the image [15]. Hong Tang, Yingfeng Qi, Li Shen, Yang Shu proposed a framework to cluster multispectral remote sensing images. A three-band false color image is used for experiments [16]. Marie Liénou, Henri Maître, and Mihai Datcu, proposed text analysis tools to semantically annotate satellite images, using LDA concepts. The algorithm classifies the image into classes using latent topics. Simple features mean and standard deviation are used [20].

IV. PROPOSED WORK

In this section, it describes the approach for classifying satellite images, using the LDA model. We had used SURF feature for images to improve the performance of the LDA models and, thus, to enhance the annotation task.

Obtain the training images for each class and pre-process it to enhance the image. Apply SURF algorithm to extract the features of training images. Since images are assumed to be independent in the LDA, a same pixel in the image may be allocated different topic labels when it resides in different images. Thus K means clustering is applied to combine similar features. Create a Bag of Features. LDA is applied for the training images of each class to obtain a threshold value. The testing image which consists of all the type of classes is acquired and pre-processed to remove the noise. Extract the SURF features and applies K means to combine similar features. Compare the threshold values of each class with the feature values of testing image for identifying the class it belongs to. Then each class is colored for better visualization. The classified image of proposed method and existing method is compared with the inputted testing image for image similarity. It shows the proposed method image have less difference when compared to existing method. The steps involved for training and testing images are shown in Fig. 1 and Fig. 2.

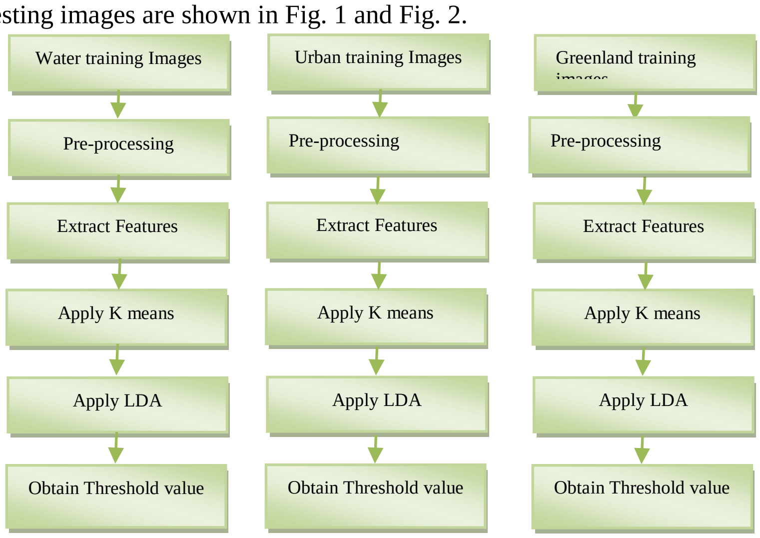

Figure 1: Steps involved for training images

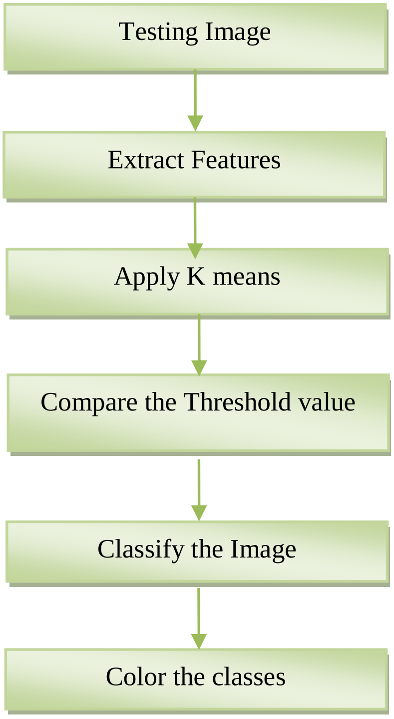

Figure 2: Steps involved in testing images

V. EXPERIMENTS AND RESULTS

A. Dataset

The LANDSAT system of remote sensing satellites is currently operated by the EROS Data Center of the United States Geological Survey. There have been seven LANDSAT satellites; in 1972 the first satellite was launched and the seventh LANDSAT was launched in April, 1999. The LANDSAT 7 images are obtained for google earth. Google Earth is a virtual globe, map and geographical information system. The earth was superimposition to map by images obtained from satellite imagery, aerial photography and GIS 3D globe. Google Earth, a free version with limited function. Google Earth can be used for research purpose. In our work we have taken three types of classes as Water, Urban and Greenland. 50 training images of LANDSAT 7 are collected from google earth for each class. Similarly 50 testing images which consist of all the type of classes are collected.

B. Experiments

Our experiment is implemented in Matlab 2012. The proposed method is applied for the training images in order to classify the regions of testing image as water, urban and Greenland. SURF features are extracted from the training images of each class as shown Fig. 3, Fig. 4 and Fig 5.

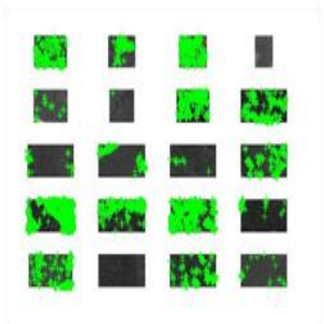

Fig 3 SURF Features for Greenland

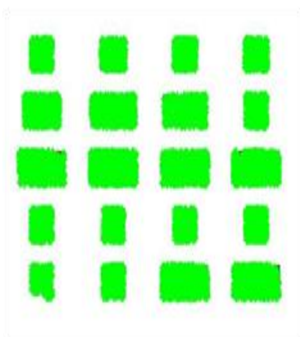

Fig 4 SURF Features for Urban

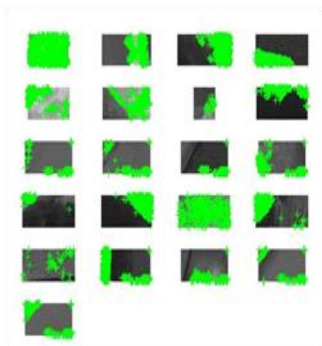

Fig 5 SURF Features for Water

Threshold values of the training images are compared with testing image and classified accordingly. Fig 6 shows the testing images acquired over a time period from 2003 to 2013. Fig 7 shows the result of image classification using mean and standard deviation as features. Fig 8 shows classified image using SURF features.

Fig 6 Testing Images acquired over a time period

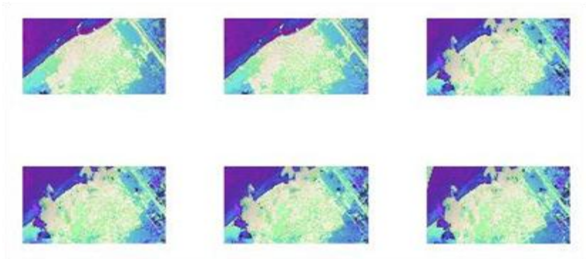

Fig 7 Classification result for Mean and Standard deviation features

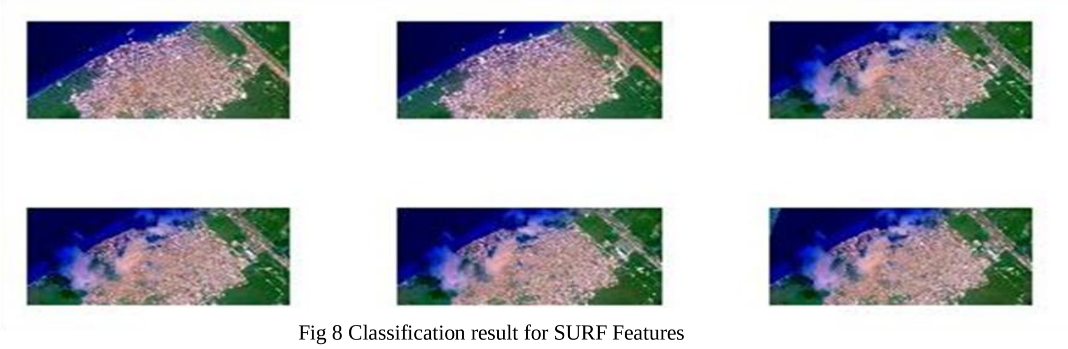

Fig 8 Classification result for SURF Features

C. Results

The resulted classified image of mean and standard deviation feature and SURF feature are compared with the inputted testing image for image similarity. Fig 9 shows that the SURF featured images have less difference when compared to mean and standard deviation feature. Thus SURF feature is best over low level feature.

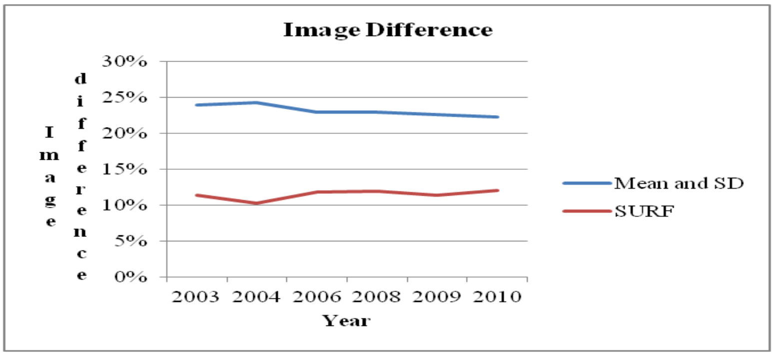

Fig 9 Line graph which show the image similarity between inputted image and classified images

Percentage for each region in the testing image is calculated. As an application of this method environmental changes such as change in any of the region is identified with help of the percentage calculated for each regions. For instance the image of a place is acquired over a time period is used to identify the environmental changes. In this work we had calculated the percentage of water, urban and greenland for the testing image which acquired over a time period from 2003 to 2013. The change detection over this time period is shown in Fig 10.

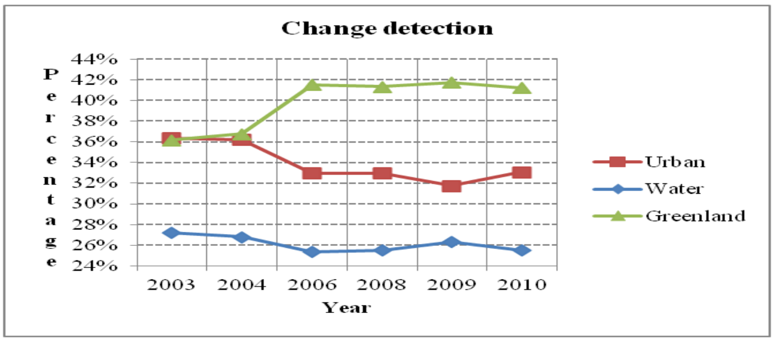

Fig 10 Changes detected for the inputted image

As a result of classification the some areas are not colored in the resulted image which indicates that parts are not appeared under trained image features. Some of the features of are overlapped with each other classes which results in mixing of colors. In certain cases the shallow regions in the image are misclassified as water region.

VI. CONCLUSIONS

In this paper we used the SURF features to enhance the performance of LDA for classification of satellite images. The classified image which uses SURF as retains most similar to the testing image than classified image which uses mean and standard deviation as feature. Thus SURF is better over low level features. An automated system is proposed to classify the satellite images and to identify the changes in the environment. Based on the images collected over a period of time, comparisons are made and analyzed over it. In future, the unidentified regions of images can be identified by increasing the number of training classes i.e. the urban area can be further classified as buildings and roads, similarly greenland can be divided into forest and vegetation land. Overlapping of feature values can be refined by improving the features extraction process and extending the LDA algorithms as spatial LDA.

REFERENCES

[1] Aykut AKGÜN, A.Hüsnü ERONAT and Necdet TÜRK,. "*Comparing Different Satellite Image Classification Methods: An Application In Ayvalik District,Western Turkey*".
[2] Aknksha Balyan, Shyam Ji Gupta, Shriniwash Singh, "*An Efficient Approach to Image Registration: Genetic algorithm*", International Conference on Recent Trends in Engineering & Technology (ICRTET2012).
[3] Barbara Zitova, Jan Flusser, "*Image registration methods: a survey*", Image and Vision Computing 21 (2003) 977–1000
[4] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, "*SURF: Speeded Up Robust Features*".
[5] Edouard Oyallon, Julien Rabin, "*An analysis and implementation of the SURF method, and its comparison to SIFT*".
[6] Teodor Costachioiu, Rodica Constantinescu, Bashar AlZenk and Mihai Datcu, "*Semantic Analysis Of Satellite Image Time Series*", 20th European Signal Processing Conference (EUSIPCO 2012) Bucharest, Romania, August 27 - 31, 2012.
[7] Tang Yingjun, "*The Variant Of Latent Dirichlet Allocation For Natural Scene Classification*", Computing and Informatics, Vol. 30, 2011, 311–319.
[8] Jiakai Liu, Rong Hu, Meihong Wang, Yi Wang and Edward Y. Chang ,"*Web-scale Image Annotation*".
[9] Duangmanee Putthividhya, Hagai T. Attias, Srikantan S. Nagarajan "*Topic Regression Multi-Modal Latent Dirichlet Allocation for Image Annotation*".
[10] Drew Schmitt, Nicholas McCoy, "*Object Classification and Localization Using SURF Descriptors*", December 13, 2011.
[11] Deren LI , Kaichang DI, Deyi LI, "*Land Use Classification Of Remote Sensing Image With GIS Data Based On Spatial Data Mining Techniques*", International Archives of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B3. Amsterdam 2000.
[12] Ntana Nkanza, "*Image Registration and its Application to Computer Vision: Mosaicing and Independent Motion detection*" , Cape Town, June 2005.
[13] Rochdi Bouchiha and Kamel Besbes, "*Automatic Remote-Sensing Image Registration Using SURF*", International Journal of Computer Theory and Engineering, Vol. 5, No. 1, February 2013.
[14] Diane J. Hu "*Latent Dirichlet Allocation for Text, Images, and Music*".
[15] Pradheep K Elango and Karthik Jayaraman, "*Clustering Images Using the Latent Dirichlet Allocation Model*", University of Wisconsin, Madison, Dec 2005.
[16] Hong Tang, Yingfeng Qi , Li Shen, Yang Shu ,"*An object-oriented clustering algorithm for Multi-spectral images using the latent dirichlet allocation*", Proceedings of the 4th GEOBIA, May 7-9, 2012 - Rio de Janeiro - Brazil. p.257.
[17] Sang Rok Lee, "*A Coarse-to-Fine Approach for Remote-Sensing Image Registration Based on a Local Method*", International Journal on Smart Sensing And Intelligent Systems, vol. 3, no. 4, December 2010.
[18] Manjusha P. Deshmukh & Udhav Bhosle, "*A Survey Of Image Registration*", International Journal of Image Processing (IJIP), Volume (5) : Issue (3) , 2011
[19] David M. Blei, Andrew Y. Ng, Michael I. Jordan, "*Latent Dirichlet Allocation*", Journal of Machine Learning Research 3 (2003) 993-1022.
[20] Marie Liénou, Henri Maître, Member, IEEE, and Mihai Datcu ,"*Semantic Annotation of Satellite Images Using Latent Dirichlet Allocation*", IEEE Geoscience and Remote Sensing letters, vol. 7, no. 1, January 2010.

