International Journal of Computer (IJC)
ISSN 2307-4523 (Print & Online)

http://ijcjournal.org/

Classification of Satellite Images Based on Color Features Using Remote Sensing

Assad H. Thary Al-Ghrairi[a*], Zahraa H. Abed[b], Fatimah H. Fadhil[c], Faten K. Naser[d]

[a]*Electronic Computer Center, Al-Karkh University of Science, Baghdad, Iraq*
[b,c,d]*Dept. Computer Science, University of Baghdad, Baghdad, Iraq*
[a]*Email: assad.thary@gmail.com*, [b]*Email: za22z22@yahoo.com*
[c]*Email: Fatema.hassan84@yahoo.com*, [d]*Email: Fatin_1984@yahoo.com*

Abstract

The aim of this paper is to classify satellite imagery using moment's features extraction with K-Means clustering algorithm in remote sensing. Clustering is the assignment of objects into groups called clusters so that objects from the same cluster are more similar to each other than objects from different clusters. In this research, the study area chosen is to cover the area of Baghdad city in Iraq taken by landsat 8. The proposed work consists of two phases: training and classification. The training phase aims to extract the moment features (mean, standard deviation, and skewness) for each block of the satellite imagery and store as dataset used in classification phase to compute the similarity measurement. The experimental result of classification showed that the image contains five distinct classes (rivers, agriculture area, buildings with vegetation, buildings without vegetation, and bare lands). The classification result assessment was carried out by comparing the result with a reference classified image achieved by Iraqi Geological Surveying Corporation (GSC). It is observed that both the user accuracy and producers' accuracy and hence overall classification accuracy are enhanced with percent 92.12447%.

Keywords: k-Means; Image features; Remote sensing; Color Moments; Satellite Image Classification; Landcover.

1. Introduction

Remote sensing is the field that uses satellite imagery to enroll data from the Earth about specific object or phenomenon without physical contact. Remotely sensed data are contributed with information of ecosystems to offer opportunity for predicting and understanding the behavior of the Earth phenomenon [1].

----

* Corresponding author.

In modern terms, Aerial sensor technologies are used to detect and classify objects from the distance of Earth by electromagnetic radiation, its surface, atmosphere and oceans by means of propagated signals [2]. RS can be divided into supervised and unsupervised classification, but mostly it is cumbersome to obtain prior knowledge due to the effect of image noise, various characteristics and complex background. The significance of Earth observation in our future decision-making processes through Remote Sensing, Pattern recognition, automatic classification, clustering, Change detection, feature extraction and parameter estimation is advantageous for economic reasons, disaster management, and high yield of crop production, deforestation, security and surveillance [3]. Land cover refers to the features of the land surface. These can be natural or man-made. The main reason for producing land cover maps is to give us an idea of what all natural and built resources exist [4]. Satellite image classification is the most significant technique used in remote sensing for the computerized study and pattern recognition of satellite information, which is the process that involves grouping the image pixel values into meaningful categories [5, 6].]. The study area chosen for research is cover the area of Baghdad city in Iraq taken by landsat 8 with an spatial resolution 30 Km and the total area covered of 185 kilometer (115 mile) wide swath. The reasons for the selection of this region are based on the area having 5 types of land and being used with support for the interpretation. These types are rivers, agriculture area, Buildings with vegetation, Buildings without vegetation, and bare lands. The geographical position of Baghdad region and its land patterns are shown in Figure 1.

Figure 1: geographical position of Baghdad.

In this paper we propose a method for classifying satellite image based on color moment features clustering. Concept of color moment is extended to obtain the features and k-means algorithm is applied to cluster and classify the images. The rest of paper is organized as follows. Section 2 provides an overview of the previous work related to satellite image segmentation and classification. Section 3 illustrates methodology about image classification techniques. Section 4 discusses the proposed approach algorithm. Section 5 describes experimental results and discussion. Finally, the conclusion is given in Section 6.

2. Related work

Many literatures are devoted to satellite image segmentation and classification. They differ in many aspects such as, material images, used approach, or even the application limitations. The feasibility of using color features-based classification method for satellite images is investigated in the following:

In [7], a novel method is presented for unsupervised classification in multi-temporal optical image based on DWT feature extraction and K-Means clustering is proposed. After preprocessing the optical image is feature extracted using the discrete wavelet transform. The feature extracted image feature reduction is performed using energy based selection. Finally, different K means clustering is performed and analyzed using MATLAB and ground truth data for improving classification accuracy. In [8], methodology is presented. It was provided for the Landcover classification of satellite images based on clustering of the Kohonen's self-organizing map (SOM). The implementation showed reasonable results. A segmentation and classification of remote sensing images were established in [9]. The classified image is given to k-Means algorithm and back propagation algorithm of ANN to calculate the density count. The excremental result found that k-means algorithm gives very high accuracy, but it is useful for single database at a time. Anil [10] proposed the segmentation method called Color –based K-means clustering, by first enhancing color separation of satellite image using decorrelation stretching then grouping the regions a set of five classes using K-means clustering algorithm. In [11], an efficient image classification technique for satellite images was proposed; the work done with the aid of KFCM and artificial neural network (NN). In spite of relatively long implementation time, the classification results were valued. Furthermore, a cellular with fuzzy rules for classifying the satellite image was implemented in [12].

The quality of classified image was also analyzed, and the results indicate the ability of evolutionary algorithms for classifying the satellite images. In [13] a method is proposed for area classification of Landsat7 satellite image using area clustering method, which depends on pixel aggregation after distributing some seeds in the test image. The assessment showed accurate classification result. In [14], it is presented that SVM algorithm shows better performance than the maximum likelihood estimation method for the classification of satellite images. In comparison with maximum likelihood SVM shows overall accuracy above 92%. The objective of this work is to use SVM technique for classifying multispectral satellite image dataset and comparing the overall accuracy with the conventional image classification method.

3. Methodology

To classify features in an image by using the elements of visual interpretation, an analyst identifies the homogeneous regions that represent various features. The classification algorithms are based on the assumption that the image to be processed contains one or more features, such as spectral region in remote sensing case. Each of these features belongs to distinct class [15].

Image classification can be performed using two different approaches: supervised and unsupervised. In supervised classification, the classes are specified by providing the sample set of data by analyst to supervised classification [16] while in unsupervised classification the algorithm automatically groups the pixels with similar moment features, such as mean, Standard deviation, and others into unique clusters based on some statistically determined criteria. The analyst then relabels and combines the spectral clusters into information classes [17]. In this paper, the K-means clustering algorithm based on color moments are used to extract features and classify satellite images. The following sections describe the concepts of the used features: color features and clustering based on K-Means.

3.1 Preprocessing

A preprocessing stage seeks to improve and enhance the visual image appearance of the image under consideration. This stage is relied on the intensity of pixels with no effect of the correlation of adjacent pixels. Such enhancement leads to improve the distinguishing between image features, which can be achieved by applying the following relation on the image:

$$I_e(x,y) = round[(\frac{I_o(x,y) - l}{h - l}) \times 255] \qquad \ldots (1) \qquad \ldots (1)\ldots.$$

Where, $I_e(x,y)$ represents the new enhanced image, $I_o(x,y)$ is the original image, x, and y are indices of the pixel in the image. l represents the bottom 1% of all pixel values of original image, and h represents the top 1% of all pixel values of original image [18].

3.2 Satellite Image Partitioning

Uniform image partitioning is a process of dividing the image into square blocks of uniform size. This process is not concerned with the spectral distribution of the image; it is just geometrical partition.

In this paper, the size of block is 4×4 depends on the amount of spatial resolution of the image.

It is taken into account that low resolution image is divided into a number of blocks is less than that of higher resolution image. This is for credit enough information are containing in each block.

3.3 Color Features

Color is a widely used important feature for image representation. This is quite important as it is invariant with respect to scaling, translation and rotation of an image. Color space, color quantification and similarity measurement are the key components of color feature extraction [19]. A basic color image could be described as three layered image with each layer as Red, Green and Blue and a as shown in Figure 2.

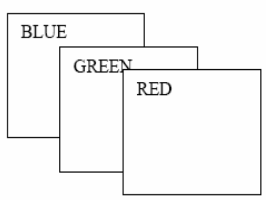

Figure 2: Image Component.

A) Color Features Extraction

Feature extraction is a sort of dimensionality reduction that efficiently represents a region of an image as a compact feature vector or the process in which certain features of interest within an image are detected and represented for further processing. Color moments are measures that can be used to differentiate images based on their features of color. The basis of color moments set in the assumption is that the distribution of color in an image can be interpreted as a probability distribution.

Probability distributions are characterized by a number of unique moments (e.g. Normal distributions are differentiated by their mean and variance). It therefore follows that if the color in an image follows a certain probability distribution, the moments of that distribution can then be used as features to identify that image based on color.

The mean, standard deviation, and Skewness of an image are known as color moments used to extract features for each block of the three color component. Following equations that definine the mean, standard deviation, and Skewness of an image of size W × H [20]:

MOMENT 1 – Mean:

$$\mu = \frac{1}{WH}\sum_{i=0}^{w-1}\sum_{j=0}^{h-1} F_{ij} \qquad \dots (2)$$

Where μ is the mean of the image or sub block of image, $F_{ij}$is the value of the ij-image pixel, *H* is the height of the image, and *W* is the width of the image. It can be understood as the average color value in the image or sub block of image.

MOMENT 2 – Standard Deviation:

The standard deviation is the square root of the variance of the distribution.

$$\text{SD} = SQRT(\frac{1}{WH}\sum_{i=0}^{W-1}\sum_{j=0}^{H-1}(F_{ij} - \mu)^2) \qquad \dots (3)$$

MOMENT 3 – Skewness:

$$\text{S} = (\frac{1}{WH}\sum_{i=0}^{W-1}\sum_{j=0}^{H-1}(F_{ij} - \mu)^3)^{1/3} \qquad \dots (4)$$

Skewness can be understood as a measure of the degree of asymmetry in the distribution.

3.4 K Means Clustering Algorithm

K-means is one of the simplest unsupervised learning algorithms in which each point is assigned to only one particular cluster. Feature extraction is the most important step of any recognition and classification system. The goal of k-means is to reduce the variability within the cluster.

The procedure follows a simple, easy and iterative way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The procedure consists of the following steps [9]:

Step 1: Sets the number of cluster k.

Step 2: Determines the centroid coordinate.

Step 3: Determines the distance of each object to the centroids.

Step 4: Groups the object based on minimum distance.

Step 5: Continues from step 2 until convergence where no object moves from one group to another.

4. Proposed approach

The generic structure of the proposed work is shown in Figure (3). It is shown that a method is designed to consist of two phases: the training and classification. The training is responsible for collecting sample image classes to be stored as in database. However, the classification is responsible for verifying the test image pixels in comparison with the trained models found in the database. Both phases are composed of three stages including: image preprocessing, image partitioning, and features extraction mentioned in sections 3.1, 3.2, and 3.3 respectively.

4.1 Training Phase

The training of dataset is an important step in the image classification. It is used for determining the image classes as shown in figure 3.

The color moment's features (mean, standard deviation, and skewness) are extracted for each image block as a vector for each feature and stored in 2D array, and then K-Means algorithm is used for grouping these features and then determining the best clusters (centroids) within the resulted features. The image part belongs or closes to each centroid are stored in database to be used in the classification phase.

4.2 Classification Phase

The classification phase is carried out after performing the training phase. The classification method depends on the comparison of spectral value of each pixel with the established database, and it depends on the proximity of each pixel into the available classes in the database.

The classification is carried out by computing the similarity measure (S_k) between each pixel in the enhanced image F_{ij} and the means μ as given in equation (5). The maximum value of S_{m_K} refers to the class to which image pixel is belonging.

$$S_k = 1 - |\mu - F_{ij}| \qquad \dots (5)$$

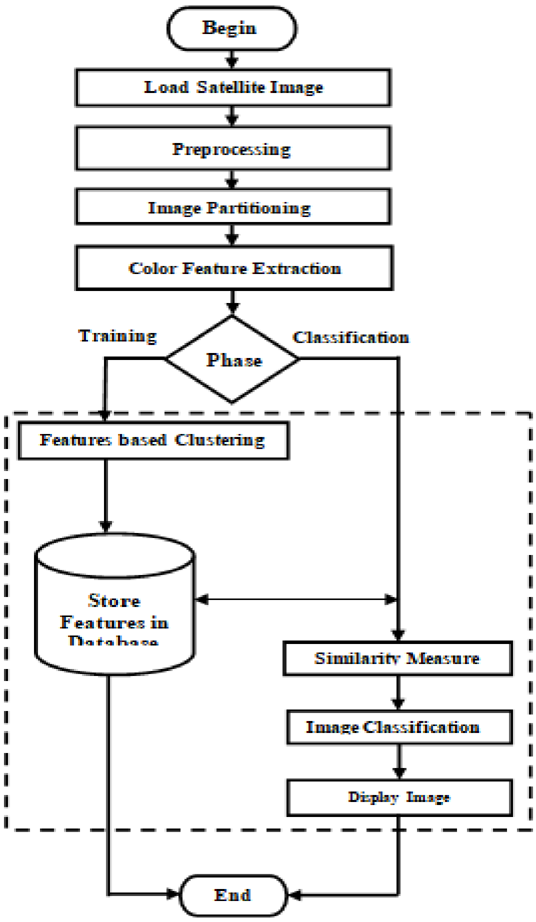

Figure 3: Block diagram of proposed classification method.

5. Experimental results and discussion

The Landsat 8 ETM$^+$ data has been used to identify the features in area of interest (AOI). The resolution of input satellite image is *1024x1024 pixels*, which carried acceptable range of informatics details about the image of consideration as shown in figure 4. It can be classified into five categories: rivers (C1), agriculture area (C2), buildings with vegetation (C3), buildings without vegetation (C4), and bare lands (C5).

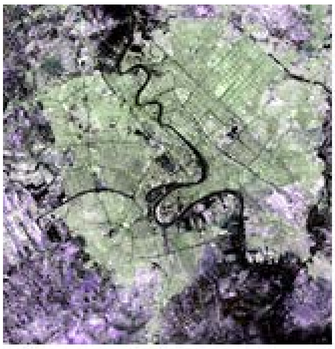

Figure 4: Color satellite image.

The first phase of the proposed work computes color moments for each block of the three color components. Each block yields a feature vector of three elements as discussed in section 3.3 i.e. mean, standard deviation and skewness. The database contains moment's features values of image blocks corresponding to the final best centroids resulted from applying the K-Means on the image blocks as discussion in section 3.4 and 4.1, where the number of iterations needed to get convergence is 12. Each of the best five centroids represents a feature for a specific block. The mean of each block shown in Table (2).

Table 1: five classes with the best Centroids.

Classes	Best Centroids
Rivers	25
Agricultural area	83
Buildings with vegetation	115
Buildings without vegetation	184
bare lands	240

The K means Clustering algorithm using the results of color features clustering is applied and Classification obtained is shown in Figure.5. What is obtained is that the proposed approach was able to sense the small variation found in some image regions, truly classifying the fine details of that regions. In the classified image, rivers are shown in blue, agricultural area in green, building with vegetation in pink, and bare land in yellow color. Based on commonly used performance measures in satellite image classification, a reference image of the same area study is classified by Iraqi Geological Surveying Corporation (IGSC) used for purpose of comparison. This image is classified by Maximum Likelihood Method using ArcGIS software version 9.3. Each individual class covers an area and it has a number of pixels shown in Table 2.

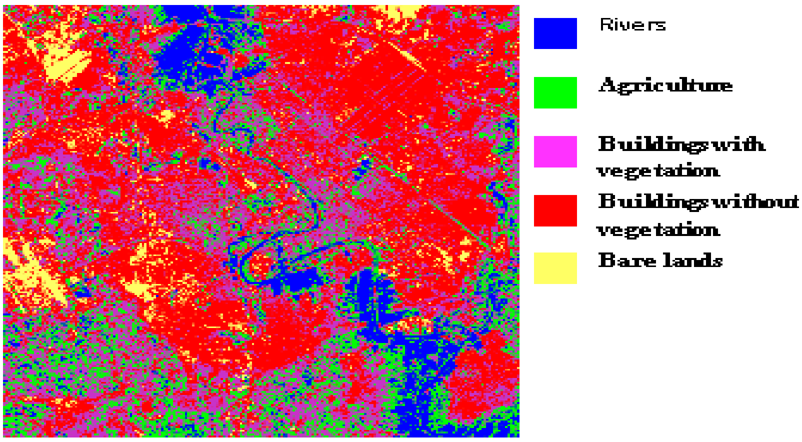

Figure 5: Classified imagery using color feature clustering

Table 2: No. of pixel with cover area for each class in classified image.

Classes	**No. Pixels**	**Area Covers (m^2)**
Rivers	96089	2882670
Agricultural area	152278	4568340
Buildings with vegetation	292375	8771250
Buildings without vegetation	453595	1.360785E+07
Bare lands	54239	1627170

Generally speaking, classification algorithm evaluation is conducted by the Confusion Matrix. The Confusion Matrix is composed of the number cmij of elements from the class i classified in the class j. The data produced in a classification scheme are the counts of the correct and incorrect classifications from each class and this information is displayed in a Confusion Matrix, which is a form of contingency table showing the differences between the true and predicted classes for a set of labeled data [21]. A matrix representing classification accuracy test (%) for the classification has been neatly constructed as shown in Table 3 that shows Confusion matrix accuracy, with the overall performance and classification for accuracy percentage of as 92.12447 %, where each class we refer are .

Table 3: Confusion Matrix/ Precision-Recall values for satellite imagery classification.

		True Classes					***Producer Accuracy (Precision)%***
		C_1	***C_2***	***C_3***	***C4***	***C_5***	
Classes /Classifier results	***C_1***	96089	0.0	0.0	0.0	0.0	100
	C_2	0.0	152278	0.0	0.0	0.0	64.83805
	C_3	0.0	82581	209794	0.0	0.0	100
	C_4	0.0	0.0	0.0	453595	0.0	100
	C_5	0.0	0.0	0.0	0.0	54239	100
User Accuracy (Recall) %		100	100	71.75511	100	100	
Overall Accuracy (OA):		**92.12447 %**					

6. Conclusion

The purpose of feature extraction is to take the important characteristics of the satellite image and classify the overall image using this small set of information. The selection of features directly effects the classification operation. Good features results have been obtained by using this procedure. The result of the classification is much cleaner so post processing is avoided. The procedure is also interactive and has few tuning parameters. K

–means clustering algorithm based on moment features classification approach is suitable when classes are to be determined by spectral distinctions that are inherent in the data .It is observed that both the user accuracy and producers' accuracy and hence overall classification accuracy are enhanced with percent 92.12447%.

References

[1] Chijioke, G. E. " Satellite Remote Sensing Technology in Spatial Modeling Process: Technique and Procedures", International Journal of Science and Technology, Vol. 2, No.5, P.309-315, May 2012.

[2] Schowengerdt D.andA.Robert.“Remote sensing: models and methods for image processing”, Academic Press (3rd ed.) ISBN 978-0-12-369407-2, 02Pp, 2007.

[3] Liu.S., L. Bruzzone. F. Bovolo, M. Zanettiand P. Du., “Sequential Spectral Change Vector Analysis for Iteratively Discovering and Detecting Multiple Changes in Hyperspectral Images,” Transactions on Geoscience and Remote Sensing, IEEE, Vol. 53, 8:4363–4378, 2015.

[4] Quanfang Wang, Haiwen Zhang, Hangzhou Sun, “New Logic for Large-scale Land Cover Classification Based on Remote Sensing”, Geoinformatics17th international conference,ISBN 978-1-4244-4562-2,Pg 1-5,2009.

[5] Baboo, Capt. Dr.S S., and Thirunavukkarasu,"Image Segmentation using High Resolution Multispectral Satellite Imagery implemented by FCM Clustering Techniques", IJCSI International Journal of Computer Science Issues, ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784, vol. 11, Issue 3, no 1, S., May 2014.

[6] Sunitha A., and Suresh B. G., " Satellite Image Classification Methods and Techniques: A Review", International Journal of Computer Applications, pp. 0975 – 8887, Volume 119 – No.8, June 2015.

[7] Venkateswaran. K., N. Kasthuri. K. Balakrishnan. and K. Prakash, “Performance Analysis of K-Means Clustering For Remotely Sensed Images,” International Jour.of Computer Applications (0975–8887) Vol.84, 12, (2013).

[8] Márcio L. Gonçalves1, Márcio L.A. Netto, and José A.F. Costa, "A Three-Stage Approach Based on the Self-organizing Map for Satellite Image Classification", Springer-Verlag Berlin Heidelberg, pp. 680–689, 2007.

[9] Sathya, P., and Malathi, L., "Classification and Segmentation in Satellite Imagery Using Back Propagation Algorithm of ANN and K-Means Algorithm", International Journal of Machine Learning and Computing, vol. 1, no. 4, October 2011.

[10] Anil Z Chitadeand Dr. S K. Katiyar,”Color based image segmentation using K-Means Clustering”, Department of civil Engineering, International Journal of Engineering Science and Technology, Vol.

2(10, 5319-5325), 2010.

[11] Ankayarkanni and Ezil S. L., "A Technique for Classification of High Resolution Satellite Images Using Object-Based Segmentation", Journal of Theoretical and Applied Information Technology, Vol. 68, No.2, and ISSN: 1992-8645, 2014.

[12] Harikrishnan.R, and S. Poongodi, "Satellite Image Classification Based on Fuzzy with Cellular Automata", International Journal of Electronics and Communication Engineering (SSRG-IJECE), ISSN: 2348 – 8549, volume 2 Issue 3, March 2015.

[13] Thwe Z. P., Aung S. K., Hla M. T., " Classification of Cluster Area For satellite Image", International Journal of Scientific & Technology Research Volume 4, Issue 06, 2015.

[14] S.Manthira Moorthi, Indranil Misra, “Kernel based learning approach for satellite image classification using support vector machine”, Recent Advances in Intelligent Computational Systems (RAICS), IEEE, DOI: 10.1109/RAICS.2011.6069282, ISBN: 978-1-4244-9478-1, 03 November 2011.

[15] A. M. Chandra and S. K . Ghosh “Remote Sensing and Geographical Information systems”, Narosa publishing House, New Delhi, 2007.

[16] Mohammed S. Mahdi Al-Taei, Assad H. Thary Al-Ghrairi,"Satellite Image Classification Using Moment and SVD Method", International Journal of Computer (IJC) Volume 23, No 1, pp 10-34, 2016.

[17] Jain., A. K. “Data clustering: 50 years beyond K-Means,” Journal of Pattern Recognition, Vol. 31,651-666, Elsevier Science Inc. New York, NY, US, 2010.

[18] Mahdi, M. S. and Abdul Hassan, A. A., Satellite Images Classification in Rural Areas Based on Fractal Dimension, Journal of Engineering, vol. 22, no. 4, pp. 147-157, 2016.

[19] Chang S.K., and Hsu A. (1992), “Image information systems: where do we go from here?” IEEE Trans. On Knowledge and Data Engineering, Vol. 5, No. 5, pp. 431-442.

[20] M.Stricker and M.Orengo, “Similarity of color images”, Storage and Retrieval for Image and Video Databases III (SPIE) 1995: 381-392.

[21].J. R. Landis, G.G. Koch, "The Measurement of Observer Agreement for Categorical Data. Biometrics", 2012, 33(1):159- 174.