Available online at www.elixirpublishers.com (Elixir International Journal)

Advanced Engineering Informatics

Elixir Adv. Engg. Info. 38 (2011) 4161-4163

ISSN: 2229-712X

Comparative analysis of glacier classification for land remote sensing satellite images

T. Sree Sharmila[1] and K. Ramar[2]

[1]Anna University of Technology Tirunelveli, Tirunelveli, Tamilnadu, India.
[2]Einstein College of Engineering, Tirunelveli, Tamilnadu, India.

ARTICLE INFO

Article history:
Received: 28 June 2011;
Received in revised form:
22 August 2011;
Accepted: 27 August 2011;

Keywords
Change Detection,
Fuzzy Classification,
LANDSAT.

ABSTRACT
Geospatial information gathered through different sensors and geographic objects is generally indistinct, vague and uncertain. The ambiguity turns out to be obvious due to the multi-granular formation of the multisensory satellite images and that directs to error accumulation at every stage. The main aim of this paper is to compare the K-Means and Fuzzy C-Means classification algorithm and find out the change detection in glacier classification by processing images taken over different time frames. The LANDSAT images correspond to the Himachal Pradesh region, one dated June 2005 and the other dated June 2010. To estimate the quality of remote sensing data the non-linear objective assessment parameters are used. Though the classification of glacier cover calculation, by improving the accurate geological classification, might be in a crude form but when projected on a larger scale, it can act as a great tool for research and analysis on a particular geographical location. The environment related bodies around the globe are deeply benefited from the valuable images provided by satellite imagery and their analysis help strategize different methods for environment protection in general and curb global warming in specific.

Introduction

This work describes the glacier classification of Land Remote-Sensing Satellite (LANDSAT) images taken from different time frames from the Himachal Pradesh region. The K-Means algorithm allows a particular data point belong to either one or the other cluster in contrast to the Fuzzy C-Means algorithm which allows data points belong to more than one cluster.

The objective of this paper is to show the superiority of unsupervised algorithms by processing LANDSAT images taken from different time frames. The image processing algorithms are applied on the images and the glacier classification is determined (Yao 2006).

The advent of global warming and thus the glacier meltdown has been a serious issue amongst the various environmental bodies of the world (Aizen and Kuzmichenok 2007). The efficient analysis of LANDSAT images can provide valuable results and information which can help strategize the preventive steps taken by the environmental bodies (Krishnan and Swoboda 2010). Here the K-Means and Fuzzy C-Means clustering techniques are widely used for analyzing satellite images to find the change in glacier content, the results are quite accurate. Furthermore, the comparative analysis of the two algorithms will highlight some of the key features of the algorithm with respect to the context of image processing and this will act as a basis for improved satellite image processing algorithms.

The major constraint of any clustering algorithm is the clustering efficiency (Thanh, Ron and Lutgarda 2005), which is directly related to the input image and where it is derived from. The satellite images, if processed accurately, can yield a lot of valuable results. The classification of glaciers and determining the percentage of change in glaciers over a time frame can give an accurate idea of the repercussions of global warming and thus strategize environmental protection plans for the future (Noferini and Macaluso 2009).

Usage of unsupervised glacier classification over supervised glacier classification can lead to reduction in manual labor and automation of research (Yang, Gong and Gao 2009). The Major functionalities include: Comparing the K-Means and Fuzzy C-Means approach for classification of LANDSAT images

System Design for K-Means Clustering

K-means is used to cluster the pixel intensities in an image into K number of clusters. This offers the segmentation of an image into K segments similar intensities and it provides more automation than manual thresholding of an image (Ravichandran and Anandhi 2009).

Assume the size of an image is m x n, convert to a vector with (m x n) rows and 1 column. With the images, a 3-dimensional feature vector per pixel is used. K-means puts pixels into K number of clusters based on similarities of intensity values.

Finally got a set of different segmented image, where each segment is relatively homogeneous in terms of pixel intensities. K-means clustering is used for a simple technique for image classification. The following are the main three steps of K-Means clustering until convergence.

Repeat the steps until convergence.

- Determine the centroid coordinate.
- Determine the distance of each object to the centroids.
- Cluster the object based on minimum distance.

Finally, to minimize the following objective function the squared error function is used.

Tele:
E-mail addresses: sreesharmilat@ssn.edu.in

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} \left\| x_j^i - c_i \right\|^2 \qquad (1)$$

where $\left\| x_j^i - c_i \right\|^2$ is a distance measure between a data point x_j^i and the cluster center c_i.

The design for implementing K- Means clustering technique is shown in Fig 1.

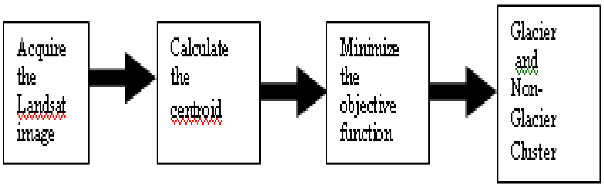

Fig 1. System Design for Implementing K-Means Clustering

System Design for Fuzzy C-means Clustering

Fuzzy C-Means (FCM) clustering algorithm permits a data belong to more than two clusters (Ruiand Donald 2005). This method is frequently used in change detection, pattern recognition and classification. It is mainly aims to minimize the following objective function.

$$A_m = \sum_{i=1}^{N} \sum_{j=1}^{C} a_{ij}^m \left\| x_j - c_i \right\|^2, \quad 1 \le m < \infty \qquad (2)$$

where $m \ge 1$, a_{ij} is the degree of membership of x in the cluster j, x_j is the j^{th} dimensional measured data, c_i is the cluster center, and $\left\| x_j - c_i \right\|$ expressing the similarity between measured data and the cluster center.

The objective function of the optimization yields the fuzzy partitioning shown above, with the updation of membership function a_{ij} and the cluster centers c_j.

$$a_{ij} = \frac{1}{\sum_{m=1}^{c} \left[\frac{\left\| x_j - c_i \right\|}{\left\| x_j - c_m \right\|} \right]^{\frac{2}{m-1}}} \qquad (3)$$

The iteration will stop until $\max_{ij} \left\{ \left| a_{ij}^{k+1} - a_{ij}^{k} \right| \right\} < \in$, $0 < \in < 1$, k is the number of steps to repeat the iteration.

The following Fig 2 depicts the system design for implementing Fuzzy C-Means.

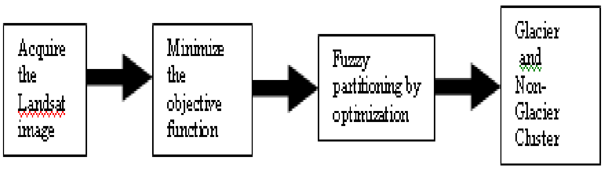

Fig 2. System Design for Implementing Fuzzy C-Means

Fuzzy C-Means Implementation Steps

- The FCM function is called upon which takes the pixel array and the number of clusters as input.
- The FCM function returns the objective function values, membership grades and the cluster centers as output.
- Based on the degree of membership the data points are moved into different clusters.
- Given a data set, there are n training samples. The values for the parameter $X = x_1, x_2, \ldots x_n$, denoted as $Y = \text{sort}[X] = y_1, y_2, \ldots y_n$. The parameter Y indicates the ascending order of the parameter X. Once the X values are sorted easily find out the association between adjacent values.
- Next determine the difference between the set of training data:

$$d_i = Y_{i+1} - Y_i, \quad i = 1, 2, \ldots. n-1. \qquad (4)$$

where Y_i and Y_{i+1} are adjacent values.

- Calculate the adjacent value similarities.

$$S_i = \begin{Bmatrix} 1 - \frac{d_i}{CP \times \sigma_s}, & d_i \le CP \times \sigma_s \\ 0 \quad , & \text{otherwise} \end{Bmatrix} \qquad (5)$$

where d_i is the difference, σ_s is the standard deviation of d_i and CP is the control parameter.

- A threshold value σ_s divides adjacent values into classes. Determination of clusters can be recapitulated by a rule. This can be expressed as a formula,

$$\begin{aligned} &\text{IF}(s_i > \sigma)\text{THEN } Y_i, Y_{i+1} \in CP_i \\ &\qquad \text{ELSE } Y_i \in CP_i, Y_{i+1} \in CP_{i+1} \end{aligned} \qquad (6)$$

where CP_i and CP_{i+1} denote two dissimilar classes for the input or output parameter.

Results and Discussions

Input Dataset Imageries

The Fig 3 and Fig 4 depict the input LANDSAT images of Himachal Pradesh region during June 2005 and June 2010.

Fig 3. June 2005 image

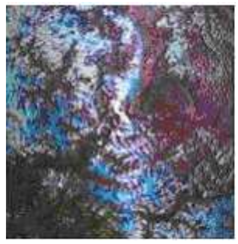

Fig 4. June 2010 image

Results of K-Means Clustering

The input images have been subjected to K-Means clustering. The Fig 5 and Fig 6 shows the results of K-Means clustering that gives the change of glacier from non-glacier in the given input image. Using the Figure 5 and Figure 6 easily find out the percentage of change of glacier from non-glacier.

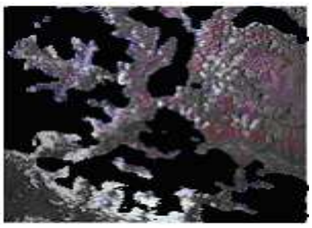

Fig 5. Objects identified by K-Means during June 2005

Fig 6. Objects identified by K-Means during June 2010

The percentage of glacier and non-glacier in June 2005 image is 36% and 63% and June 2010 is 29% and 71%. So, the change in coverage area of glacier is a significant decrease over five years in the given image by 7% and vice-versa.

Results of Fuzzy C-Means Clustering

The Fig 7 and Fig 8 portray the results of June 2005 and June 2010 image after applying Fuzzy C-Means clustering. The percentage of change of glacier in June 2005 image over June 2010 image is 8%. In accordance with the real time data, it gives a confidence to indicate that the work to be performed will yield a fruitful result.

Fig 7. June 2005 image after Applying Fuzzy C-Means

Fig 8. June 2010 image after Applying Fuzzy C-Means

Conclusion

The glacier images are classified using the unsupervised algorithms, ones which do not need pre determined cluster definitions, namely the K-Means and the Fuzzy C-Means algorithms. The glacier classification using the Fuzzy C-Means approach yields a superior result compared to K-Means approach. Using the K-Means approach there are a number of unclassified glacier and so the classification accuracy is low compared to the Fuzzy C-Means approach. According to the ground truth the Fuzzy C-Means clustering derives the better result compared to the K-Means clustering. This work will intend to define a prediction mechanism so as to aid the long-term planning by predicting the preventable losses and to generalize this scheme for global warming.

References

[1] V. Krishnan, J. Swoboda, C. E. Yarman and B. Yazici, Multistatic Synthetic Aperture Radar Image Formation, IEEE transactions on image processing, 2010, vol. 19, no. 5, pp. 1290-1296.

[2] K. S. Ravichandran and B. Ananthi, Color Skin Segmentation Using K-Means Cluster, International Journal of Computational and Applied Mathematics, 2009, vol. 4, no. 2, pp. 153-157..

[3] Rui Xu and Donald Wunsch, Survey of Clustering Algorithms, IEEE, 2005, vol. 16, no.3, pp. 645-678.

[4] Thanh N. Tran, Ron Wehrens, Lutgarde M. C. Buydens, Clustering Multispectral Images: a Tutorial, Chemometrics and Intelligent Laboratory Systems, 2005, pp. 3-17.

[5] Aizen V.B, Kuzmichenok V.A., Glacier changes in the Tien Shan as determined from topographic and remotely sensed data, Global and Planetary Change, 2007, 56(3), pp.328-340.

[6] Noferini L, Mecatti D, Macaluso G., Monitoring of Belvedere Glacier using a wide angle GB-SAR interferometer, Journal of Applied Geophysics, 2009, 68(2), pp.289-293.

[7] Yang Xiaomei, Gong Jianming, Gao Zhenyu, The research on extracting method of microscale remote sensing information combination and application in coastal zone. Acta Oceanologica Sinica, 2009, 31(2), pp. 40-48.

[8] Yao Tandong, The Response of Environmental Changes on Tibetan Plateau to Global Changes and Adaptation Strategy, Advance in Earth Sciences, 2006, 21(5), pp.459-464.