APPLICATION OF BINARY DIVISION ALGORITHM FOR IMAGE ANALYSIS AND CHANGE DETECTION TO IDENTIFY THE HOTSPOTS IN MODIS IMAGES

Harish Kumar G R* and Dharmendra Singh
(hargrdec@iitr.ernet.in, dharmfec@iitr.ernet.in)
Department of Electronics and Computer Engineering,
Indian Institute of Technology Roorkee,
Roorkee, India.

Commission VI, Working Group IV/3

KEY WORDS: Binary Division Algorithm, Change Detection, Hotspot, MODIS images, Multi-resolution fusing

ABSTRACT:

In this paper, an approach based on Binary Division Algorithm, to detect the Hotspots using Multi-resolution fusion of band 1, band 2, band 31 and band 32 of Moderate Resolution Imaging Spectroradiometer (MODIS), NASA Satellite, for the Jharia (India) Region with the aid of image analysis and change detection technique has been proposed.

INTRODUCTION:

The management of mine fires in coal mining region is of much concern in India. Mine fires apart from economic aspects; give rise to devastating environmental effects. Most of the fires in Indian coalfield, which cause a local rise in the surface temperature, take place due to spontaneous heating of coal, which depends on various mining, geological and coal factors. The main objective of this paper is to use operational satellite data which is freely available and detect the hotspots which may help fire managers with the spatial allocation of coal fire prevention or fire extinction resources. The use of satellite remote sensing has made it possible to detect coal fires and has recently been considered as an effective and economic alternative for fire monitoring given that a number of factors involved in fire susceptibility estimation may be derived at least partially, from satellite data. Satellite observations can provide both the extent of fire scares and the expansion of ongoing fires in time and space.

Jharia coalfield in Jharkhand is the richest coal bearing area in India which contains a large number of coal mine fires which have been burning for several decades. The Jharia Coalfield is located in the Dhanbad district of the state of Bihar and is named after the main mining area of Jharia. The Jharia Coalfield is confined between latitudes 23°38' N and 23°50' N and longitudes 86°07' E and 86°30' E. The maximum extent of the coalfield is about 38 km from East to West and 19 km from North to South.

Hotspots are defined as image pixels whose brightness temperatures exceed a pre-defined threshold value. Threshold values of 316-320 K have been reported in scientific literature. Different temperature threshold will result in different hotspot counts. The lower the threshold, the higher the number of hotspots detected. MODIS-series satellites has become the most widely used satellite data set for regional fire detection and monitoring because of its availability, spatial resolution, spectral characteristics, and low costs.

* Corresponding Author. Phone: 09837483542.

The Terra and Aqua MODIS instrument provides high radiometric sensitivity (12 bit) in 36 spectral bands ranging in wavelength from 0.4 µm to 14.4 µm. The responses are custom tailored to the individual needs of the user community and provide exceptionally low out-of-band response. Two bands are imaged at a nominal resolution of 250 m at nadir, with five bands at 500 m, and the remaining 29 bands at 1 km. A ±55-degree scanning pattern at the EOS orbit of 705 km achieves a 2,330-km swath and provides global coverage every one to two days. In This paper, the Terra MODIS, band 1 with resolution of 250m and bandwidth 620-670 nm, band 2 with resolution of 250m and bandwidth 842-876nm and band 31, 32 with resolution of 1km and bandwidth 10780-11280nm and 11770-12270nm respectively.

Change detection is the process of identifying differences in the state of a feature or phenomenon by observing it at different times. There are many remote sensor system and environmental parameters that must be considered whenever performing change detection. Failure to understand the impact of the various parameters on the change detection process can lead to inaccurate results. Ideally, the remotely sensed data used should be acquired by a remote sensor system that holds temporal, spatial, spectral, and radiometric resolutions constant.

In this paper the MODIS images of May 2005 and Nov 2005 have been considered. Binary Division Algorithm, which is one of the efficient clustering methods, has been applied on these images for detecting hotspots in Jharia region. Clustering gives us significant information about data distribution in multidimensional feature space. There have been a lot of clustering methods reported. They are roughly divided into three groups: agglomerative methods, partitioning methods, and agglomerative-partitioning methods. The agglomerative methods have high efficiency but low accuracy. On the other hand, the partitioning methods are accurate and efficient, but they require huge memory, and much time. The agglomerative-partitioning methods**,** improve the accuracy by using iterative processing, which makes the method inefficient. In this paper, a partitioning method, binary division Algorithm, for improving accuracy and efficiency has been proposed. An ISODATA clustering Algorithm for agglomerative-partitioning methods also exists, whose performance to Binary Division Algorithm is not of that accuracy and efficiency, hence Binary Division Algorithm, have been used to detect the hotspots

Principles and Procedures

In this paper, the band 1 and band 2 of MODIS (MOD09A1), and band 31 and band 32 of MODIS (MYD11A2) of May 2005 and Nov 2005 have been considered. The MOD09A1 have a resolution of 500 m and MYD11A2 have a resolution of 1000m. The MYD11A2 resolution was increased to 500m by the application of bilinear interpolation technique. The MOD09A1 provides an estimate of the surface spectral reflectance and MYD11A2 provides an estimate of the surface emissivity. So the MOD09A1 is normalized with respect to the MYD11A2 data. On these images the Binary division Algorithm is applied.

Binary division Algorithm

In the binary division, it is essential to determine the cluster to be divided, the subset of the feature space to be used for the division, and the division threshold in the subspace.

A. Selection of Division Cluster:

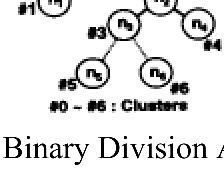

Figure 1. Binary Division Algorithm

* Corresponding Author. Phone: 09837483542.

The figure 1 shows an example of the binary division Algorithm. In this example, image data are divided into clusters 1 and 2 at node n_0, and those in cluster 2 into clusters 3 and 4 at a node n_2. Cluster 3 is selected for the next division among all terminal clusters 1**,** 3, And 4**.** The principle is described using this example.

Clustering in Binary division Algorithm is regarded as a minimization process of the total within-group sum of squares (WGSS) of the image data. In this Algorithm, image data are clustered so that the maximum reduction is obtained in the intragroup sum of squares (IGSS) among child clusters produced by the division. It is known that a mixed cluster c_0 of clusters c_1 and c_2 has the variance

$$\mathrm{Var}(c_0) = \mathrm{Var}(c_1) + \mathrm{Var}(c_2) + [\mathrm{mean}\ (c_1) - \mathrm{mean}\ (c_2)]^2 \qquad (1)$$

Using the same manner, we obtain the relation for the sum of squares in N-dimensional feature space as

$$S_{Parent} = S_{childl} + S_{child2} + S_{between} \qquad (2)$$

Where S_{Parent} is an N x N matrix for the WGSS of a parent cluster, S_{childl} and S_{child2} are those for child clusters produced by the division, and **&,** $S_{between}$ means the IGSS between the child clusters. We select the division cluster among all terminal clusters so that the WGSS efficiently falls to the minimum. We define a matrix D for evaluation of the reduction of the WGSS, as

$$D = S_{Parent} - (S_{childl} + S_{child2}) \qquad (3)$$

The reduction ***D,*** however, depends on the sum of squares of the parent cluster. So we define the reduction rate A as a normalized index

$$A = \mathrm{trace}(S_E^{-1/2} D S_E^{-1/2}) \qquad (4)$$

* Corresponding Author. Phone: 09837483542.

Where S_E is the total WGSS of all terminal nodes clusters, and is $S_1 + S_3 + S_4$ in the example shown in Figure 1. Thus we select the next cluster among all terminal clusters so that a maximum reduction rate is obtained. Next, we will describe the algorithm for selecting the optimal subspace to divide the cluster into two subclusters and for determination of the threshold in the subspace.

B. Subsets of Feature Space and Boundary Search: Search areas in which the optimal boundary will be selected become larger as the dimension (the number of spectral bands) increases. In order to achieve higher efficiency, we decrease the dimension by projecting image data onto a small number of the feature spaces. We adopt the first two canonical components p_1 and p_2 to form the two-dimensional subspace as canonical correlation analysis is useful in data compression and in noise reduction. We use the following six projection functions defined on a two-dimensional subspace.

$$\begin{array}{ll} f_1(p_1), & f_2(p_2) \\ f_3(p_1 + p_2), & f_4(p_1 - p_2) \\ f_5(p_1 \times p_2), & f_6(\tan^{-1}(p_1/p_2)) \end{array} \qquad (5)$$

Let V_T be the variance-covariance matrix of data in a cluster, and V_e be that of noise. We obtain the eigenvector matrix B which makes BV_TB' and BV_eB' diagonal matrices. The first two canonical components p_1 and p_2 are derived from eigenvectors related to the larger two eigenvalues. An element σ_{ij}^2 of the matrix V_e is estimated from differential values

$$\sigma_{ij}^2 = \frac{1}{20\{(s-2)(t-2)-1\}} \sum_k \sum_l u_{kl}^i u_{kl}^j \qquad (6)$$

$$u_{kl}^i = d_{k-1,l}^i + d_{k,l+1}^i + d_{k+1,l}^i + d_{k,l-1}^i - 4d_{k,l}^i \qquad (7)$$

Where we assume the image has s columns and t lines (s x t = M), and $d^i_{k,l}$ means spectral density of a pixel at the position of the k^{th} column and the l^{th} line in an image of the i^{th}

band. Next is the boundary search in the subspaces. Data projected onto the subspaces are reassigned to integer values from 0 to 255 for the following processing. We use the histogram of these values to search a candidate for the optimal boundary. We employ an index Q, a kind of pixel density, in a cluster for boundary search as

$$Q = \frac{L}{(V+1)} \tag{8}$$

Where L means the number of pixels in a cluster, V is variance of data in the cluster, and constant 1 is added to avoid the divergence of Q. The index changes from Q_{parent} to Q_{childl} + Q_{child2} with division. As the division decreases $S_{between}$ in (3), the density increases with division as

$$Q_{parent} \leq Q_{childl} + Q_{child2} \tag{9}$$

The equality holds when data in the parent cluster are uniformly distributed. The candidate for the optimal boundary is selected among all possible boundaries so that the sum of densities of child clusters has the maximum, that is

$$Q_1(X) + Q_2(X) = \max_{x}\{Q_1(x) + Q_2(x)\} \tag{10}$$

Where, we suppose cluster 0 is divided into clusters 1 and 2, x is a boundary number, and X the selected boundary. The optimal boundary is selected among these candidates obtained from all the subspaces. We use a normalized density G for comparison as

$$G_i = \frac{Q_1(X_i) + Q_2(X_i)}{Q_{0i}} \quad (i=1\sim6) \tag{11}$$

Where ***i*** specifies the subspace. The boundary X_j in the j^{th} subspace is the optimal boundary for division of a terminal cluster when

$$G_j = \max_{i}\{G_i\} \tag{12}$$

* Corresponding Author. Phone: 09837483542.

C. Procedures: The procedures of Binary Division Algorithm are summarized below. As Algorithm has no stopping rule yet, the algorithm is stopped when the number of clusters is equal to a specified number N_{ic}.(The no of clusters got by applying the ISODATA algorithm on these data)

1) Specify the number of clusters to be obtained.
2) Apply procedure 3) for all terminal clusters, and select a cluster for the next division.
3) Apply canonical correlation analysis to a terminal cluster and project the first two canonical components to the subspaces. Select the optimal boundary among all possible boundaries in the spaces.
4) Divide the cluster at the boundary.
5) Repeat procedures 2)-4) until the number of clusters is equal to N_{ic}
6) Calculate the mean vectors of all terminal clusters, and replace the density of pixels with the mean vector.
7) Yield the resultant image.

The clustered image, is segmented to isolate hotspots and change detection is applied to this clustered image and the MOD14A2(most confident fire) of May 2005 image

RESULTS

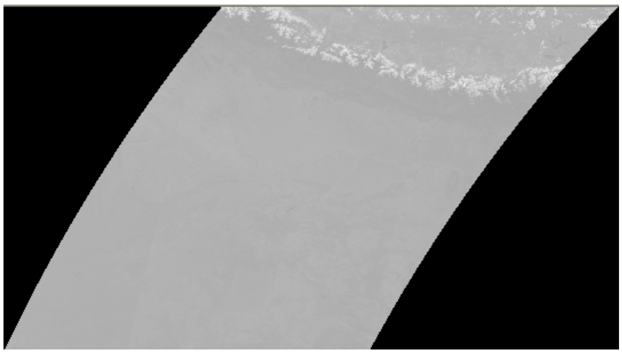

Figure 2. MODIS Band 1 May 2005 Image

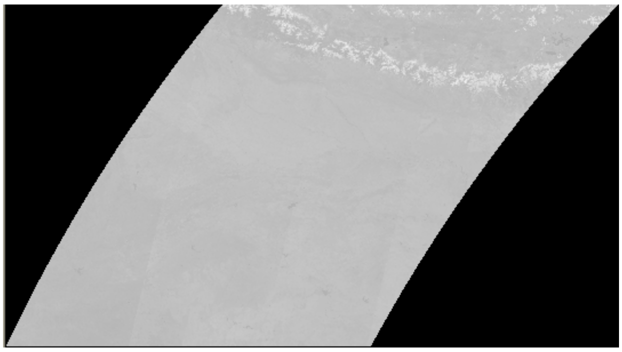

Figure 3. MODIS Band 2 May 2005 Image

Figure 4. MODIS Band 31 May 2005 Image

Figure 5. MODIS Band 32 May 2005 Image

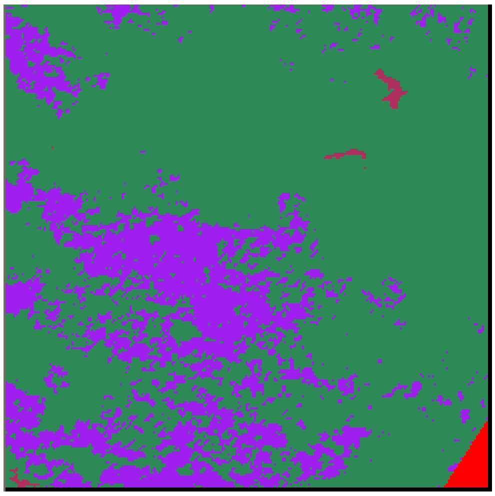

Figure 6. BDA Applied image

Figure 7. Segmented image

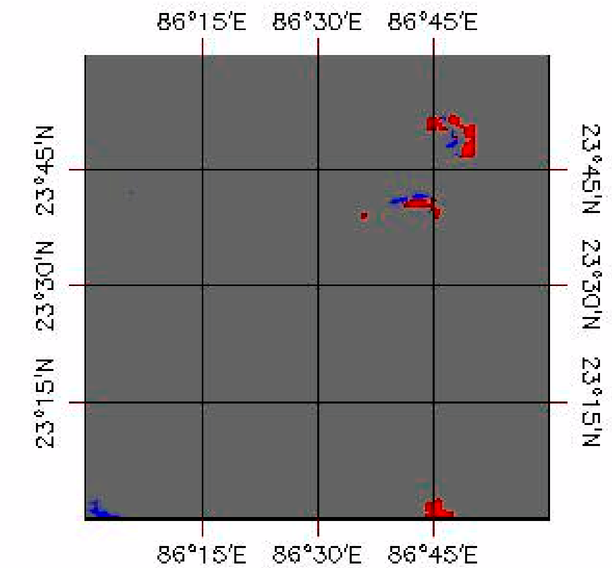

Figure 8. Change Detected image

* Corresponding Author. Phone: 09837483542.

CONCLUSIONS

With the application of Binary Division Algorithm, and change detection techniques, it is found that hotspots are clearly detected in MODIS images, which shows the efficient way to analyze the image in unsupervised way. The result and proposed algorithm are very helpful to develop a database of hotspots monitoring system for satellite data.

REFERENCES

[1] H.Hanaizumi, S.Chino, and S.Fujimura: A Binary Division Algorithm for Clustering Remotely Sensed Multispectral Images, IEEE Trans., 1M44-3, pp.759-763 (1995)

[2] H.Hanaizumi and S.Chino: binary division algorithm using a linear discriminant function for the cluster analysis of remotely sensed multispectral images, SPIE Proceedings--Volume 2579, Image and Signal Processing for Remote Sensing II, November 1995, pp. 182-187

[3] D. P. ROY, L. GIGLIO, J. D. KENDALL and C. O. JUSTICE: Multi-temporal active-fire based burn scar detection algorithm, Int. Journal remote sensing, 1999, vol. 20, no. 5, 1031- 1038

[4] C.O. Justice, L. Giglio, S. Korontzia, J. Owens, J.T. Morisette, D. Roy, J. Descloitres, S. Alleaume, F. Petitcolin and Y. Kaufman: The MODIS fire products, Remote Sensing of Environment 83 (2002) 244–262

* Corresponding Author. Phone: 09837483542.