Overview on Different Algorithm for Cloud Detection and Removal for Satellite Images

Ms.Manisha Raut[1] Ms.Poonam Agarkar[2] Mr.Aditya Dhanvijay[3] Mr.Ketan Machhale[4]
[1,2,3,4]Department of Electronics And Telecommunication Engineering
[1,2,3,4]Rajiv Gandhi College of Engineering & Research, Nagpur, India

Abstract— Clouds is one of the significant obstacles in extracting information from tea lands using remote sensing imagery Different approaches have been attempted to solve this problem with varying levels of success in the past decade, a number of cloud removal approaches have been proposed. In this paper we review and discuss about the cloud detection & removal, need of cloud computing, its principles, and cloud removal process and various algorithm of cloud removal. This paper attempts to give a recipe for selecting one of the popular cloud removal algorithms like The Information Cloning Algorithm, Cloud Distortion Model and Filtering Procedure, Semi-Automated Cloud/Shadow, And Haze Identification and Removal etc. A cloud removal approach based on information cloning is introduced...Using generic interpolation machinery based on solving Poisson equations, a variety of novel tools are introduced for seamless editing of image regions. The patch-based information reconstruction is mathematically formulated as a Poisson equation and solved using a global optimization process. Based on the specific requirements of the project that necessitates the utilization of certain types of cloud detection algorithms is decided.

Key words: Cloud removal, information cloning, Poisson Equation, Haze Identification and Removal, cloud computing etc.

I. INTRODUCTION

Globally, the Enhanced Thematic Mapper Plus (ETM+) land scenes are, on average, about 35% cloud covered, as reported by Ju and Roy [1], indicating that cloud covers are generally present in optical satellite images. This phenomenon limits the usage of optical images and increases the difficulty of image analysis. Thus, considerable research efforts have been devoted to the topic of cloud removal to ease the difficulties caused by cloud covers [2]–[7]. If multitemporal images are acquired, the cloud-cover problem has a chance to be eased by reconstructing the information of cloud contaminated pixels under the assumption that the land covers change insignificantly over a short period of time.

In the past decade, a number of cloud removal approaches have been proposed. These approaches can be classified into three categories: *in painting-based*, *multispectral-based*, and *multitemporal-based*. In the first category, without the aid of multispectral and multitemporal data, the cloud-contaminated regions are synthesized using image synthesis and in painting

Techniques [4]–[5]. The information inside the cloud contaminated region is synthesized by propagating the geometrical flow inside that region. The synthesis approaches can yield a visually plausible result, which is suitable for cloud free visualization. However, the lack of restoring information of cloud-contaminated pixels makes them unsuitable for further applications. In multispectral-based approaches, multispectral data are utilized in cloud detection and removal [3]–[9]. Rakwatin et al. [3] proposed a reconstruction algorithm to restore missing data of Aqua Moderate Resolution Imaging Spectro radiometer (MODIS) band 6 using histogram matching and least squares fitting.
Compared with the multispectral-based approaches, the multitemporal-based approaches [2], [6], [7]–[8] which rely On both temporal coherence and spatial coherence have a better ability to cope with large clouds.

Image editing operation is related with global changes such as image correction, filtering, colorization or local changes in a selected region where the altering operations take place. One example of this is the commercial or artistic photomontages that consider the local changes. Along with the technologic improvement in this research topic, quite number of software has been developed for photo editing such as Adobe Photoshop. But, professional experience is required to be able to use these kinds of software skillfully and editing photos using the software takes a long time. Additionally, the edited image regions may include some visible corruptions.

In recent years, the image editing methods based on The Poisson equation have been frequently employed [10-11]. An image editing method was presented by Perez et al. based on the Poisson equation with Dirichlet boundary conditions. But, using this method, color inconsistencies occurred in edited image regions. An image matting approach using the Poisson equation is suggested by Sun et al. However, due to a long processing time, the method is not practically usable. Chuan et al. improved the method presented by Perez et al. to overcome the color inconsistency problem. But the experiments show that the improved method is still very complex. Leventhal et al. suggested an alpha interpolation technique to remove brightly colored artifacts caused by mixed seamless cloning in the result. Jia et al. presented an image editing method, called drag-and-drop pasting, which computes an optimized boundary condition automatically by employing a new objective function. But this study compares the developed method not with mixed seamless cloning but with only seamless cloning method proposed in. Georgiev suggested a new method that is invariant to relighting and handles seamlessly illumination change, including adaptation and perceptual correctness of the results.

II. CLOUD DETECTION AND REMOVAL ALGORITHM

Clouds are one of the significant obstacles in extracting information from tea lands using remote sensing imagery.
Different approaches have been attempted to solve this problem with varying levels of success.
The common algorithm or techniques in general include:

1) The information cloning algorithm

All rights reserved by *www.ijsrd.com*

2) Cloud Distortion Model And Filtering Procedure
3) Semi-automated Cloud/Shadow, and Haze Identification and Removal
4) Second Highest Value
5) Modified Maximum Averaging
6) Hybrid Method
7) Regression of visible difference (REVD)

III. THE METHODOLOGY

A. The Information Cloning Algorithm

It consists of the following steps: cloud and shadow detection, blob detection, cloud removal, information reconstruction

1) Cloud and Shadow Detection

Given a cloud-contaminated image, called target image, and its corresponding images captured at the same position but different times, called reference images. A semiautomatic window based thresholding approach is adopted to detect clouds and cloud shadows in both the target and reference images. In the thresholding-based approach the cloud boundaries in the cloud contaminated images are defined. The cloud detection based on window based thresholding approach is by considering hypothesis that, regions of the image covered by clouds present increased local luminance values to automatically detect the presence of cloud in a region. The window size can be varied depending on cloud size, if clouds are of varied sizes and occupying only 1% of larger resolution. Choose the threshold value depending on the statistical properties of the image. Once the cloud pixels are identified, their shadows are roughly predicted according to the cloud location. The dark and connected components within the neighborhood of the predicted shadows are identified as shadow components. This approach is simple and can detect most clouds and cloud shadows.[13]

2) Blob detection

The blob detection [7] must be performed after cloud and shadow detection. The blob detection block supports variable size signals at the input and output. The aim is to remove the cloud and shadows according to the statistics that are computed during blob detection. The blob analysis will return the labeled region and its statistics such as pixel location, number of pixels and blob count.

3) Information Reconstruction

The details of selected blobs are used to reconstruct the information of corresponding cloud-contaminated regions using information cloning algorithm [6]. The reconstruction problem is mathematically formulated as a Poisson equation [6] and solved using a global optimization process The cloud contaminated region in the target image is denoted as Γ, and its boundary is denoted as $\partial\Gamma$. Let f be an unknown image intensity function defined over the cloud-contaminated region Γ (i.e., the unknown that is to be calculated). Let f be the image intensity function defined over the target image minus the cloud-contaminated region Γ, and let V be a guidance vector field defined over the cloud-contaminated region Γ. The vector field V is defined as the gradient of the selected patches and is used to guide the reconstruction process [8] to optimize the pixel intensities in the cloud-contaminated regions. Thus, the information of cloud-contaminated region is reconstructed by several different blobs in a reference image. When the cloud-contaminated region Γ contains pixels on the border of the target image, these pixels are calculated to remove the boundary values.

B. Cloud Distortion Model and Filtering Procedure

Assume that an image of the earth is produced when a light cloud cover exists over t

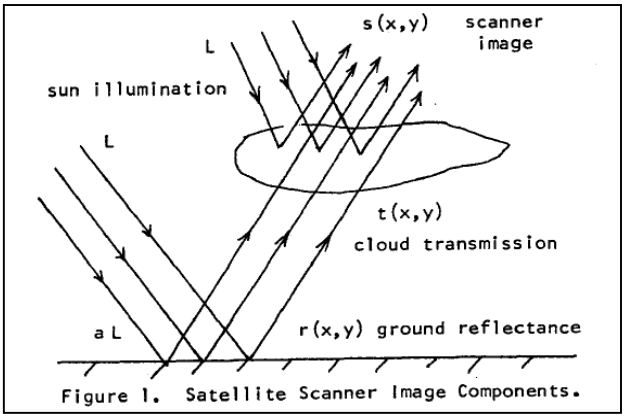

Fig. 1: the region of interest as shown in Fig.1.

If we assume that the cloud reflection of sunlight plus the cloud transmission equals one (ignoring diffusion) and that the illumination is approximately constant on the earth's surface, then the received image at the scanner is

$$s(x,y) = a\,L\,r(x,y)t(x,y) + L[I - t(x,y)] < L \quad (I)$$

Where d(x,y) is the signal and dx,y) is the noise.The values of r(x,y), t(x,y), and a (sunlight attenuation) range between 0 and 1.A transformation is now performed by subtracting s(x,y) from L and taking the logarithm:

$$\log[L-S(x,y)] = \log[t(x,y)] + \log(l - aLr(x,y)] \quad (II)$$

If the signal is now assumed to be log[L-aLr(x,y)] and the noise is assumed to be log[t(x,y)]. Then the signal and noise are additive and uncorrelated.

Weiner linear filtering techniques can now be used to remove the noise. This method of converting a multiplicative process to an additive one and then applying linear filtering has been generalized and named homomorphic filtering. In this case both multiplied terms (reflectance and transmission) are non-negative so that the simple logarithm is an effective transform.

In order to follow the procedure outlined above, the sun illumination l must be estimated from the cloudy picture. Since a, r{x,y), and t{x,y) are all between 0 and I, the maximum value of s(x,y) cannot be greater than l (see Eq. I). If the cloud transmission at any point is zero, the value of s{x ,y) at that point will be l. Therefore a reasonable value for l in a large image is the brightest point in the image. The original data is, therefore, processed by subtracting the intensity of each point from the maximum intensity in the picture. The logarithm is then taken of the inverted data. Now the signal and noise are additive. The filter function is

$$H(\mu\, v) = \frac{S_{MP}(\mu\, v)}{Spp(\mu\, v)}$$

where SMP{V,v) is the cross power spectrum between signal and signal -plus-noise and Spp(U.v) is the power spectrum of the Signal-plus-noise. The two spatial frequency components are J and v. This is a non-causal filter which uses all the cloudy pictury points to estimate each

All rights reserved by *www.ijsrd.com*

individual signal point. In order to apply this filtering, an estimate must be made of SMP(U'v).[12]

C. Semi-Automated Cloud/Shadow, and Haze Identification and Removal

1) A Cloud/Shadow Identification

Instead of employing TRRI and CSI Index, cloud area was extracted using Unsupervised Classification for 3 visible bands (for instance, bands 1, 2, and 3 for both Land sat and AVNIR-2) and then recoding the representing classes (which is generally one or all from 28th to 30th). The Unsupervised Classification was also tried using 4 bands data; however, the result was not synchronized as that With 3 visible bands.[14]

2) B Shadow Identification

This was carried out by incorporating its direction and estimating the average distance from cloud. Shadow direction was known using the Sun Azimuth angle of the image, which is generally provided with associated metadata. Average distance of shadow from cloud was estimated by measuring two or more cases on image. Both cloud and shadow areas were extended for proper representation. Then, both were combined to one.

D. Haze Identification and Removal

1) Haze identification

For this, Haze component of Tassel Cap (TC) transformation was used. The clear and hazy area was separated using equation (III) and pixel with value greater than mean TC was designated as Haze area (Richter, 2011).

$$TC = (X1 * DN_{BLUE}) + (X2 * DN_{RED}) \quad \text{(III)}$$

Where, DN_{BLUE} : DN value in Blue band

DN_{RED} : DN value in Red band

X1 : weighing coefficients for Blue band

X2 : weighing coefficient for Red band

The index coefficient for Land sat 5 TM was used that presented by Crist et al. (1986). And, for ALOS AVINIR-2, such coefficient being un-available, the same coefficient that for Landsat5 TM was employed and result is promising one. Cloud/Shadow, and water body area was not included in the haze area. Water body was estimated using Normalized Difference Vegetation Index (NDVI) method followed by some manual editing.

2) Haze removal

First, using the Pixel values of Blue and Red bands for clear area, slope angle (α) was calculated, which was used to estimate haze levels map or Haze Optimized Transform (HOT) using equation (IV).

$$HOT = DN_{BLUE} * \sin\alpha - DN_{RED} * \cos\alpha \quad \text{(IV)}$$

both ALOS AVNIR-2 and Landsat TM, the haze signal to be subtracted, Δ for each HOT level and for each AVNIR-2 band was calculated. Dehazing was done by subtracting the Δ from original DN for haze area (Richter, 2011).

E. Second Highest Value

Second highest value (SH) algorithm is another method used for cloud analysis. This method works based on calculating second highest values in each row of an image matrix.

Each pixel of an image will be taken in a vector form and these vectors are sorted to find out second highest value. This value will be treated as threshold which will be used to compare with pixel resolution for discriminating cloud and non cloud regions. Pixel value greater than threshold is assumed as a cloud cover region and remaining pixel values are considered as cloud free region [2].

F. Modified Maximum Averaging

Modified Maximum Averaging (MMA) will detect and remove cloud regions better than mean and SH methods. These algorithms will support only for low-level brightness clouds. MMA is one of the enhanced algorithms used to remove high brightness clouds. Procedure used for MMA algorithm is as follows;

i) First step is to assume image as a vector.

ii) Find mean for whole vector.

iii) The subset of the pixel is extracted by comparing pixel values to the mean value of an image [2].

iv) If the pixel value is higher than mean then eliminate the pixel values and remaining values of pixel is consider as cloud free region.

v) Repeat the process until entire cloud region has been detected

G. Hybrid Method

Mean, SH and MMA have number of drawbacks and implementation constraints. These methods are not accurate for removal of multi objects from an image. Development of hybrid method is needed to overcome constraints and drawbacks of existing algorithms. Hybrid method is a combination of mean and MMA. The process of hybrid method is to apply mean algorithm for finding the threshold This threshold values will be used by MMA algorithm for cloud removal.

H. Regression of Visible Difference (Revd)

Observations for the regression technique were drawn using systematic samples of every 20th pixel from areas visually identified as cloud-free land, cloud-free water and cloudy lands. The colour composite used in the visual interpretation of the images was the Red-Green-Blue (4-2-1, default setting for Erdas Imagine).

As can be seen from Fig. 2, the colour of thick cloud usually varies between white and white cyan, partial clouds over ocean appears in dark blue while that over land appear in turquoise and aquamarine. Land cloud free pixels vary in colour depending on the cover type. They vary between golden, orange, light yellow or light pink. The reflectance values for the visible channels 1 and 2 and the brightness temperature values of channel 4 were recorded for each image by surface type.[15]

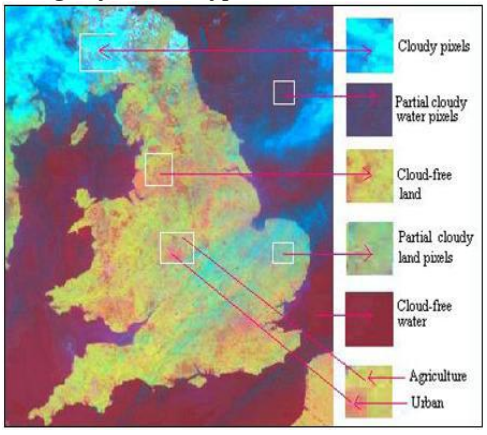

Fig. 2: Visual appearance of a four-layer (Ch1, Ch2, NDVI, Ch4) image

All rights reserved by *www.ijsrd.com*

Scatter plots were generated showing the relation between "Ch2 – Ch1" and ch2 and simple linear regressions were fitted to the data of the three surfaces. The resultant regression equations were then used to estimate the lower confidence interval of the relation so as to be further utilized in the screening of partial-cloud pixels. For this task, "Ch2 – Ch1" values were estimated from the regression equation and subtracted from the observed values to calculate the deviations. Standard errors were then calculated and used to form the lower confidence interval. The equations for cloud masking were finally calculated such that only pixels falling above and to the left of the lower confidence interval would pass the test for cloud-free land. Cloudy pixel brought in by the calculated confidence interval, were further subjected to the second test using Ch4 brightness temperature. This was done on the assumption that cloudy pixels are normally cooler than land pixels with similar reflectance properties.

IV. CONCLUSION

Based on the specific requirements of the main project that necessitates the utilization of certain types of cloud detection algorithms, the following points could be drawn from this investigation:

1) No cloud masking technique is absolutely perfect and without errors. Variation between algorithm types depends on the nature of spectral channels used and on the surfaces upon which they operate, while the accuracy of a given algorithm is both spatial and temporal dependent.
2) In the presence of thin and lower cloud types, the optimization of output from algorithms based on simple thresholds is very difficult, and the situation is further aggravated if the surface is inhomogeneous. In order to get rid of cloudy pixels in these situations, some cloud-free pixels have to be sacrificed.
3) REVD technique is found to have excellent potential of being suitable for application in various conditions over both spatial and temporal domains. For more accurate results, the hybrid approach is preferable than the REVD alone, and to further reduce the total cost, the entire-season regression approach is quite sufficient to yield the desired accuracy.
4) In this paper, an information cloning algorithm for cloud removal has been introduced. The cloud-contaminated portions of a satellite images are detected based on window based thresholding approach. The detected clouds are removed and then the information of missing data is reconstructed with a single reference image. This approach is based on the patch-based information reconstruction strategy with the global optimization process. This approach results in cloud removed images and is tested for various input images.
5) The haze removal algorithm also worked well and it removed haze from all 3 bands; Blue, Green Red (that is, < 800nm).Haze appears in these bands. Mean at row vector and column vector have been considered for fixing threshold value for better segmentation of cloud cover regions. Enhanced algorithm also works better for removing tiny cloud regions which will be considered noise.

REFERENCES

[1] J. Ju and D. P. Roy, "The availability of cloud-free Landsat ETM Plus data over the conterminous United States and globally," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1196–1211, 2008.
[2] P.J. Hardin, R.R. Jensen, D.G. Long and Q.P. Remund , "Testing Two Cloud Removal Algorithms for SSM/I ", 1999
[3] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Restoration of Aqua MODIS band 6 using histogram matching and local least squares fitting," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 613–627, Feb. 2009.
[4] F. Chen, Z. Zhao, L. Peng, and D. Yan, "Clouds and cloud shadows removal from high-resolution remote sensing images," in *Proc. IEEE*
[5] L. Lorenzi, F. Melgani, and G. Mercier, "Inpainting strategies for reconstruction of missing data in VHR images," *IEEE Geosci. Remote Sens.*
[6] P. Perez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 313–318, Jul. 2003
[7] F. Chun, M. Jian-wen, D. Qin, and C. Xue, "An improved method for cloud removal in ASTER data change detection," in *Proc. IEEE IGRASS*, 2004, pp. 3387–3389.
[8] D.-C. Tseng, H.-T. Tseng, and C.-L. Chien, "Automatic cloud removal from multi-temporal spot images," *Appl. Math. Comput.*, vol. 205, no. 2, pp. 584–600, Nov. 2008.
[9] D. P. Roy, J. Ju, P. Lewis, C. Schaaf, F. Gao, M. Hansen, and E. Lindquist, "Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization,
[10] Perez P., Gangnet M. and Blake A., "Poisson Image Editing, ACM Transactions on Graphics", 22(3), 313-318, 2003.
[11] Islam A. and William A., "Efficient, Low-Complexity Image Georgiev T., Covariant Derivatives and Vision, ECCV" 2006, 9th European Conference on Computer Vision, Graz,
[12] Mitchell, O. R. and Chen, P. L., "Filtering to Remove Cloud Cover in Satellite Imagery" (1976). LARS Symposia. Paper 152. http://docs."
[13] Saranya M, International Journal of Computer Science and Mobile Computing, Vol.3 Issue.2, February- 2014, pg. 681-688
[14] A. K. Sah , B. P. Sah , K. HonjiA, N. Kubo , S. Senthil " semi-automated cloud/shadow removal and land cover change detection using satellite imagery"
[15] El Mamoun Haroun. Osman "Demonstrating an Efficient Algorithm for Cloud Detection and Removal for Satellite Images" Journal of Applied and Industrial Sciences, April, 2013, 1 (1): 7-15

All rights reserved by *www.ijsrd.com*