# Cloud removal using multi-temporal satellite images and fuzzy logic

Soumya Sucharita, Satyajit Mohanty, Manoj Pandya, M. B. potdar, Leena Patel, Krunal patel

**Abstract**— Satellite images are often obscured by cloud cover and shadows. This poses a major challenge to processing of data regarding the surface underneath. Various methods have been proposed by different researchers to obtain cloud and cloud shadow free images from satellite images. In this paper, an approach is suggested which generates cloud free image from multi-temporal images. This paper briefly reviews the existing methods for reconstruction of cloud contaminated images and then suggests an approach based on segmentation of image using Fuzzy logic. The cloud contaminated patches are detected in an image using Fuzzy C means algorithm. Following this step, blobs are detected and those pixels are replaced with the corresponding pixels from the cloud free image under the assumption that the land cover changes insignificantly during a short period of time. Filtering technique is used to remove the visible seams in the reconstructed image. Experimental analysis is conducted on satellite images and results are obtained. Both thin and thick clouds are removed effectively. We have also compared the performance of various segmentation algorithms in cloud detection. This proves the improved performance with the proposed approach.

**Index Terms**— Segmentation, Clustering, Image reconstruction, Fuzzy C means algorithm, segmentation

## 1 Introduction

Remote sensing is an effective tool for monitoring of the Earth's surface. It helps in improving the perception of our surroundings and is useful for analyzing geo-spatial data, change in land cover. However, optical sensors present in satellite can't penetrate vegetation or clouds. It is often difficult to acquire cloud free imagery in optimal time periods. by reconstructing the information of cloud-contaminated pixels with data from the cloud free image, under the assumption that the land covers change insignificantly over a short period of time. The basic method is to identify clouds and shadows in a target image and replace them with cloud and shadow-free parts from the reference images that are acquired at the different period of time. The workflow of the proposed cloud removal method consists of the following steps: cloud detection using Fuzzy C means algorithm, blob detection and information reconstruction. In the first step, Fuzzy C means clustering is adopted to detect clouds and cloud shadows for the input images. Then blob detection is used for labeling the detected regions and also provides the statistics of labeled region. In the last step, pixel values replacement is performed to fill in the missing data after removing the cloud-contaminated pixels.

In Section 2, authors would be discussing existing literature related to cloud detection and reconstruction of images. Section 3 would comprise of implementation methodology and details of various algorithms to be used in the proposed approach. Section 4 contains the results obtained by following the approach. The paper would then end with a suitable conclusion in Section 5. Moreover, in section 6, authors discuss the possible future avenues for research.

- *Author Soumya Sucharita & Satyajit Mohanty are affiliated with BITS Pilani, Pilani campus, Rajasthan, India, E-mail: satyajitm55@gmail.com*
- *Manoj Pandya, M. B. potdar, Leena Patel and Krunal patelare affiliated with BISAG, Gandhinagar, Gujarat, India. E-mail: mjpandya@gmail.com*

## 2 Comparison of Various Algorithms for Segmentation

Image segmentation is the process of partitioning an image into non-overlapping regions each with its unique characteristics like texture, color or intensity. Segmentation transforms the image into something that is easier to analyze. This is done by assigning a label to every pixel in an image such that pixels with the same label share certain characteristics. The result of image segmentation is a set of segments that collectively cover the entire image. All the pixels in a region are similar with respect to some characteristic such as color, intensity, or texture. Pixels of the neighboring regions are significantly different with respect to the same characteristics. There are various algorithms to segment an image. The algorithms which have been tested by us are listed below.

### 2.1 Edge detection

Edges consist of a set of points at which image brightness changes sharply. This can be useful in detecting the edges of clouds as well as cloud shadows. But the disadvantage is that boundaries of other features like wastelands, reservoirs etc. may be detected. The Canny Edge detector was developed by John F. Canny in 1986. It is often referred to as the optimal edge detector. Canny algorithm aims to maintain a low error rate, good localization and minimal response. One can achieve low error rate by detecting only the existent edges. Basically, it should minimize the number of false positives. The distance between edge pixels detected and real edge pixels have to be minimized. This will give a good localization. By minimal response, the detector must return only one response for each edge.

ISSN 2229-5518

### 2.2 Watershed algorithm

Any grayscale image can be viewed as a topographic surface where high intensity denotes peaks and hills while low intensity denotes valleys. The algorithm begins by filling every isolated valley (local minima) with different colored water (labels). As the water rises, depending on the peaks (gradients) nearby, water from different valleys, obviously with different colors will start to merge. To avoid that, one must build barriers in the locations where water merges. The process of filling water and building barriers is repeated until all the peaks are under water. Then the barriers created give the segmentation result. This is the working principle of the watershed algorithm. But this approach gives over-segmented result due to noise or any other irregularities in the image.

### 2.3 K means algorithm

Finding groups of objects such that the objects in a group will be similar (or related) to one another and different from (or unrelated to) the objects in other groups is called clustering. Intra-cluster distances are minimized and inter-cluster distances are maximized. The K-means algorithm is an iterative technique that is used to segment an image into K clusters. The basic algorithm is 1. Take any K cluster centers. 2. Assign each pixel in the image to the cluster that minimizes the difference in intensities between the pixel and the cluster center. 3. Calculate the cluster centers by averaging all the pixels in the cluster. 4. Repeat steps 2 and 3 until only a minimum number of pixels change clusters. In image processing context, distance is the squared or absolute difference between a pixel and a cluster center. K can be selected randomly or using a heuristic. The quality of the solution depends on the initial set of clusters and the value of K.

### 2.4 Thresholding

The simplest method of image segmentation is called the thresholding method. This method uses a threshold value to turn a gray-scale image into a binary image. The main issue of this method is selecting the threshold value. It can be selected by trial and error method. The range is relative to the possible signal values of an image’s class. The pixels above this particular values (level) are replaced with white pixels and all other pixels with value zero (black).

## 3. Implementation Methodology

The various methods involved are described below.

### 3.1 Histogram matching:

In histogram matching, an image is taken as input and based on its’ histogram, the output image is obtained. the required inputs for this method are the input image and the specific image from which the histogram can be obtained. the objective is to obtain a transformed image of the input image, such that the sum of absolute errors between the intensity level histogram of the transformed picture and that of a reference picture is minimized. This is done so that the histograms of the cloud free image and the original cloud contaminated image match. This will be useful because when we replace the cloudy patches in the original image with the corresponding cloud free patches of the multi-temporal image, there might a huge difference in intensity between the patch and the image. so, the reconstructed image will show the seams of the replacement, which is highly undesirable. to avoid this, histogram matching can be used.

### 3.2 Segmentation using Fuzzy C means algorithm

Clustering of data is a method by which large sets of data are grouped into clusters of smaller sets of similar data. Objects in one cluster have high similarity to each other and are dissimilar to objects in other clusters. In hard clustering, data is divided into crisp clusters, where each data point belongs to exactly one cluster. In fuzzy clustering, the data points can belong to more than one cluster. Also, each point has associated membership values which indicate the degree to which the data points belong to the different clusters.

Fuzzy segmentation algorithms are considerably effective in cloud detection, considering the level of noise and uncertainty involved in a cloud contaminated image. Also, other features like reservoir etc. would not be detected even if they have intensity levels similar to that of clouds. This is a huge advantage of the fuzzy algorithm.

Fuzzy C means clustering method is a simple and powerful method to segment an image. Fuzzy partitioning is carried out through an iterative optimization of the objective function with the update of membership matrix of each pixel and the cluster centres cj.

The aim of the fuzzy c means algorithm is to minimize the following objective function-

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2$$

Here, uij denotes the degree of the membership of the ith pixel to the jth cluster, xi denotes the location of the pixel, cj denotes the location of the cluster center and m denotes the level of fuzziness. N is the total number of pixels and C is the number of cluster centers chosen

Algorithm: 1. Choose any n cluster centers. 2. Assign a fuzzy membership to each data point. Compute the new centre of each class.

The center is calculated using this formula-

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

3. Move the centers. 4. Calculate the new fuzzy membership of

ISSN 2229-5518

each point. 5. Repeat steps 2,3,4 till convergence is attained i.e. membership values don't change much. The final membership matrix gives the membership values (or degree of belongingness) of the points to each cluster. After preprocessing, the cloud contaminated image is segmented using Fuzzy C means algorithm.

### 3.3 Blob Analysis

In histogram matching, an image is taken as input and based on its Blob detection methods are aimed at detecting regions in a digital image that differ in properties, such as brightness or color, compared to surrounding regions. Basically, a blob is a region of an image in which some properties are constant or approximately constant. In a segmented image, it identifies the blobs and returns the pixel locations of the detected segments. Blob analysis provides the detailed properties of the detected blobs like mean, centroid, intensity, area, perimeter etc. After segmentation, the blob detected image is converted to a binary image using a proper value of threshold. The value of threshold is chosen at random. Sometimes, the threshold value could be set at a level such that even the cloud shadows got detected.

### 3.4 Image reconstruction:

Cloud contaminated pixels detected during blob analysis were replaced by pixel values of corresponding locations in the cloud free image. This was implemented by manipulating the matrix form in which the image is stored.

### 3.5 Median Filter:

Median filtering is widely used in digital image processing because it preserves edges while removing noise. The main idea of the median filter is to run through the image pixel by pixel, replacing each pixel with the median of neighboring pixels. The pattern of neighbors is called the "window", which slides, pixel by pixel, over the entire image. It is very effective in removing speckle noise and "salt and pepper" noise (impulsive noise).

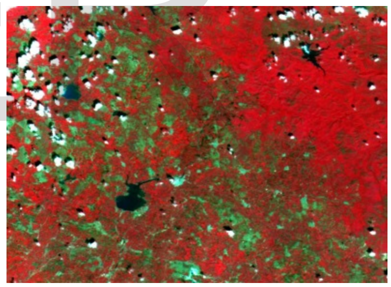
### 3.6 Image sharpening:

Enhancing the high-frequency components of an image leads to an improvement in the visual quality. Image sharpening refers to any enhancement technique that highlights edges and fine details in an image. One example of such an algorithm is an unsharp mask. It is a flexible and powerful way to increase sharpness.

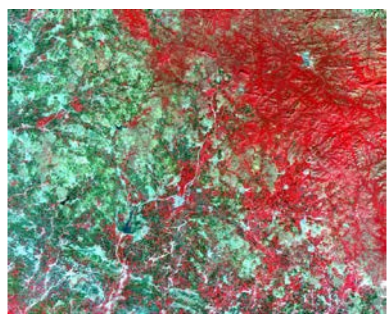
i. The procedure of removal of clouds from images is described below-
ii. First, a cloud free image of the same geographical region as that of the cloud contaminated image at a different time is collected. These are basically multi-temporal satellite images. Both images must have same spatial correlation.
iii. Then fuzzy c means algorithm was applied to the image. The resultant image was a segmented image with clouds detected.
iv. The cloud detected image was transformed into a binary image using a threshold value. The threshold value can be selected heuristically. Sometimes, the threshold value could be set at a level such that even the cloud shadows got detected.
v. Blob detection and analysis is performed on the binary image. This returns the detected cloud pixels and their properties like intensity.
vi. Then, image reconstruction was done. Cloud contaminated pixels detected during blob analysis were replaced by pixel values of corresponding locations in the cloud free image.
vii. Because of visible seams after pixel by pixel replacement, median filter was applied on the resultant image. Bigger mask size (5 by 5) gave better results but resulted in a blurry image. An image sharpening algorithm could be used in such a case.
viii. Sometimes, after pixel by pixel replacement of pixels, distinct patches are visible in the resultant image. So, one can use histogram matching to somewhat make the histograms of two images similar to each other.

## 4 Experiments and Results

The procedure described in section 3 is applied to the multi-temporal images shown in Fig.1. The stepwise implementation of the procedure is demonstrated in Fig. 2.



(a)



(b)

Fig 1: Multi-temporal satellite images (a) Cloud contaminated image (b) Cloud free image of the same region at a different point of time

ISSN 2229-5518

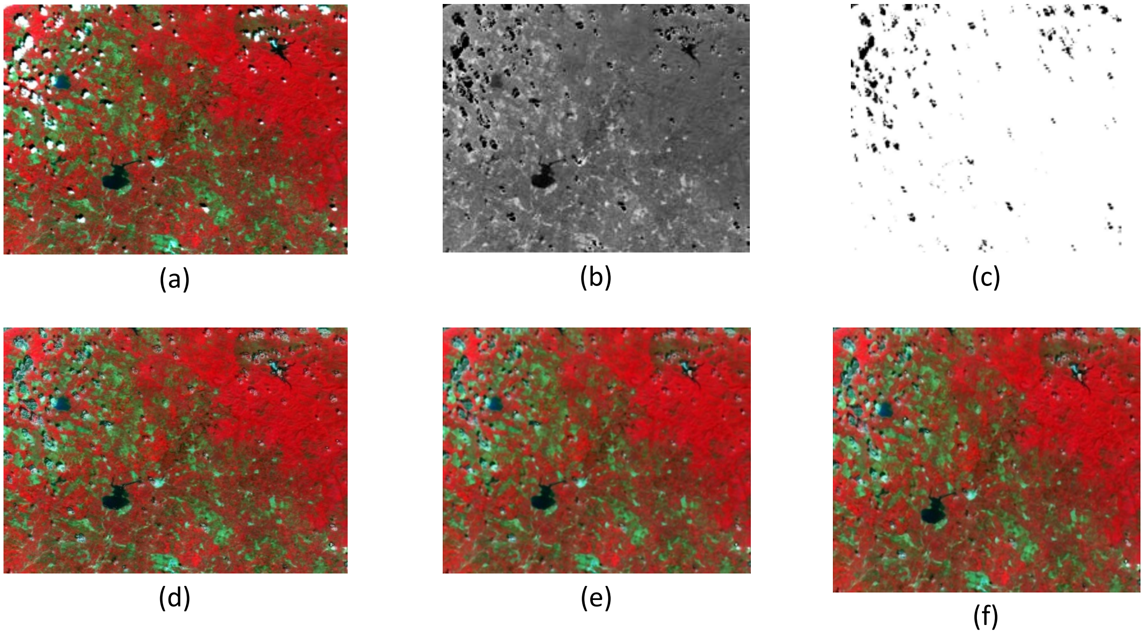



Fig 2: (a) Cloud Contaminated Image (b) Detected clouds using Fuzzy C Means Clustering Method (c) Blob Analysis using thresholding (d) Information Reconstruction (e) Median Filter with 5x5 window (f) Image Sharpening algorithm



## 5. Conclusion

In this paper, we have analyzed the validity of various segmentation algorithms in cloud and shadow detection. The findings suggest that the fuzzy c means clustering method is the best algorithm for cloud detection. Shadow detection can be done using thresholding method. Then pixel by pixel intensity value replacement, followed by median filtering can be done to reconstruct cloud contaminated images. Thus, we have suggested a simple yet effective method to accomplish the image reconstruction.

## 6. Future Scope

The image reconstruction methodology can be improved. Instead of using a filter and pixel value replacement, one could use seamless cloning method. This method will blend the replaced patch into the original image so that the edges are not visible. But since the method is image specific, the cloud removal and image reconstruction can't be made automatic.

## Acknowledgment

The authors wish to thank Mr. T.P. Singh, Director, Bhaskaracharya Institute of Space Applications and Geo-informatics, Shri M. H. kaluberme and BITS, Pilani for their valuable support and inputs in the research work.

## References

[1] Chao-Hung Lin, Po-Hung Tsai, Kang-Hua Lai, and Jyun-Yuan Chen, "Cloud Removal From Multitemporal Satellite Images Using Information Cloning," IEEE Transactions On Geoscience And Remote Sensing, vol. 51, no. 1, Jan 2013.

[2] S R Surya, Philomina Simon, "Automatic Cloud Removal from Multitemporal Satellite Images" , J Indian Soc. Remote Sens (March 2015) 43(1):57–68, DOI 10.1007/s12524-014-0396-2.

[3] Din-Chang Tseng ,Chun Liang Chien, "A Cloud Removal Approach for Aerial Image Visualisation" , International Journal of Innovative Computing, Information and Control Volume 9,Number 6, June 2013.

[4] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 442–455, 2006.

[5] C. Huang, N. Thomas, S. N. Goward, J. G. Masek, Z. Zhu, J. R. G.Townshend, and J.E. Vogelmann, "Automated masking of cloud and cloud shadow for forest change analysis using Landsat images," *Int. J.Remote Sens.*, vol. 31, no. 20, pp. 5449–5464, Jun. 2010.

[6] M. Li, S. Liew, and L. Kwoh, "Automated production of cloudfree and cloud shadow-free image mosaics from cloudy satellite imagery," in *Proc.20th Congr. Int. Soc. Photogramm. Remote Sens.*, 2004, pp. 15–23.

[7] Rafael C. Gonzalez and Richard E. Woods, "Digital Image Processing", 2nd edition, 2009.