Analysis of Water Bodies under Partial Cloud Conditions Using Satellite Images

Rahul Neware [1], Mansi Thakare [2]

[1] PG Scholar, Dept. of Computer Science and Engineering, G.H. Raisoni college of Engineering Nagpur, Maharashtra, India – neware_rahul.ghrcemtechcse@raisoni.net

[2] PG Scholar, Dept. of Communication Engineering, G.H. Raisoni college of Engineering Nagpur, Maharashtra, India – mansi.thakare.ghrcemtechce@raisoni.net

Keywords: water bodies; satellite images; vector data; SVM; positive and negative buffering; polygons

ABSTRACT:

The technique of obtaining information or data about any feature or object from afar, called in technical parlance as remote sensing, has proven extremely useful in diverse fields. In the ecological sphere, especially, remote sensing has enabled collection of data or information about large swaths of areas or landscapes. Even then, in remote sensing the task of identifying and monitoring of different water reservoirs has proved a tough one. This is mainly because getting correct appraisals about the spread and boundaries of the area under study and the contours of any water surfaces lodged therein becomes a factor of utmost importance. Identification of water reservoirs is rendered even tougher because of presence of cloud in satellite images, which becomes the largest source of error in identification of water surfaces. To overcome this glitch, the method of the shape matching approach for analysis of cloudy images in reference to cloud-free images of water surfaces with the help of vector data processing, is recommended. It includes the database of water bodies in vector format, which is a complex polygon structure. This analysis highlights three steps: First, the creation of vector database for the analysis; second, simplification of multi-scale vector polygon features; and third, the matching of reference and target water bodies database within defined distance tolerance. This feature matching approach provides matching of one to many and many to many features. It also gives the corrected images that are free of clouds.

1. INTRODUCTION

Satellite images are known as remotely detected images, which enables precise mapping of land cover and highlights scenes that are justifiable on local, mainland and even worldwide scales. Normally, governments and big corporate houses across the world launch and operate their own imaging satellites, which record and collect images of our planet and other different planets in the galaxy. Satellite images are extremely efficient in mapping expanded regions (Neware and Khan, 2018).

In India, the Hyderabad-based National Remote Sensing Centre (NRSC) ISRO, is at the forefront of keeping track, through satellite images, on the relative standing and position of all water surfaces in the country. It also issues a fortnightly publication giving information about water spread areas. Thus, since 2012 the Centre has been collecting time-segmented information about water surfaces. This has been organised in a geo-spatial database under the classification of water reservoirs. Listed as Water Body Information System (WBIS), this data is published on Bhuvan Geo-platform, for observation and queries relating to areas of water spread. It provides integrated view of water-spread dynamic for over 12,500 water reservoirs in the nation. But the appearance of cloud can lead to errors or underestimation of the spread in a few areas, as optical images cannot penetrate cloud cover (Bhuvan Water Body Information System, 2018; Neware and Khan, 2018).

Several methods were tried and used to restore the cloud affected areas, but most of them are used now for removing thin clouds (Chanda & Majumder, 1991). To remove cloud and shadow defects Abd-Elrahman et al. used fusion techniques. Which is also used by have Borchartt et al. and suggested a new method that works well, but the algorithm fails to compensate cloud-shadow, so the scaling factor, for each segment of the image, has be secured experimentally (Abd-Elrahman et al., 2008; Borchartt et al., 2011).

This paper advances a new approach to obtain cloud-free satellite images by comparing cloudy target images with cloud-free reference images with the help of vector data processing. The samples used in this study include the water reservoirs of Maharashtra, comprising areas greater than 3000 m. This database is in vector form represented by complex polygons. Each polygon shows a particular water body in Maharashtra state. For this analysis Quantum Geographical Information System (QGIS) is used. QGIS is open source software used for the different GIS operations (QGIS-2.8-UserGuide, 2018).

2. LITERATURE REVIEW

Taking a look at the relevant literature, we find that in their study Abraham and Sasikumar analysed together with visible and infrared satellite images, remotely sensed RGB and monochrome images, using adaptive segmentation algorithm to detect the cloud in cloudy regions. First, they converted RGB images into YIQ and then measured average intensity. Variation in window size indicated the cloud present in an image. Then, applying threshold to images, they extracted the cloud (Abraham & Sasikumar, 2013). Meanwhile, to extract cloud Yuan and Hu advanced the segmentation strategy, which employs the Bag-of-Words (BoW) model to divide image into super-pixels to extract its features. Then Simple Vector Machine (SVM) is applied to separate cloud and cloud-free areas in a training data sample (Yuan & Hu, 2015). Finally, GrabCut algorithm, which delineates the foreground and background parts of an image correctly, is employed get better cloud region. But on their part Bai et al. have focused on the smart information reconstruction (SMIR) technique to restore cloud-contaminated images through use of Extreme Learning Machine-ELM (Bai et al., 2016). For instance, on an image of a lake captured by Moderate Resolution Imaging Spectroradiometer (MODIS), SMIR is applied to get the loss value of reflectance, resulting from cloud, in the image. Then, application of Back prorogation (BP) neural network algorithm gives a comparison between cloudy and non-cloudy pixels in the image. Finally, a thorough analysis creates cloud-free images in time series.

In their study, to interpret cloud coverage in satellite images Ik-Hyun & Mahmood categorized cloud as thick cloud, thin cloud and ground cloud, using local and global threshold, the latter of which determines presence of cloud in an image (Ik-Hyun & Mahmood, 2015). This approach involves feature extraction and feature matching, outlier removal, re-sampling process. It uses uniform feature scale invariant feature transform (UF-SIFT) to obtain more correct image classification. However, to create cloud-free satellite images, Chen et al. execute spatially and temporally weighted regression (STWR) model. This model involves choice of spatial and temporal pixels, interpretation of target and reference images, analysis of cloud-affected area through use of spatially and temporally weighted regression (STWR), removal of similar pixels in images through use of inverse distance weighted (IDW) and finally, creation of repetitive cloud-free images by processing cloudy images (Chen et al., 2017). Meanwhile Machiwal et al. present an extensive survey of water management techniques available in remote sensing area, the remote sensing and GIS applications for monitoring the surface water bodies and the pre-processing stages for data extraction and analysis in remote sensing area. Expounding the fundamental concept of GIS in water bodies' data monitoring, it also highlights its limitation areas (Machiwal et al., 2010).

On the other hand, the study of Hargrove et al. puts forth the Goodness-of-Fit (GoF) approach to calculate degree of sameness in two or more than two spatial datasets. This method interprets the degree of best among different datasets that are in polygon form and can be applied to the entire satellite map (Hargrove et al., 2006). For comparing two datasets, this approach uses GoF algorithm, which marks the intersection between two polygonal data that assigns the values of insideness and area weighting term. Meanwhile, the research paper of Huh focuses on the matching technique for spatial data sets' boundary edges, which uses the string-matching algorithm. In this, spatial data sets' boundary points are culled in pre-processing stage. Then, polygon objects are separated in the form of string for further interpretation (Huh, 2015). To ensure correctness of string matching, buffering is executed. Finally, the study of Martínez et al. advocates executing Boolean operations on polygons. The stated algorithm is effective for concave polygons, several contour and intersecting edges. It creates the resulting contour as common intersecting area between two polygons. Then, the polygons' edges are subdivided and those that lie inside or outside of another polygon are chosen. When the common edges of two intersecting polygons are joined, the contour is created (Martínez et al., 2009).

From the above literature review it is evident that researchers have used satellite image datasets in diverse areas of study. It also transpires that water management has also emerged as a crucial field in today's world and the fact that remote sensing can facilitate its study. Particularly, the analysis of water bodies needs to be focused on and improved, with special emphasis on proper analysis of features of surface water bodies.

3. VECTOR DATA ANALYSIS IN SATELLITE IMAGES

The structure of a vector data model is far removed from the structure of a raster data model. A vector model stores all spatial data as a line, point or polygon. On the other hand, a raster data model uses sets of grid cells to record all data. In remote sensing, vector data is referred to as coverage. In this paper, vector data is represented as coverage. The three types of spatial data - a point, line or polygon - are referred to as features,

and a vector GIS can be termed a feature-based technology (Qingxiang et al., 2012). The shape of the spatial entity is stored using two-dimensional (x, y) coordinate system.

In this study, polygon feature represents water bodies in vector data format. Polygon encloses a homogeneous area which is a set of (x, y) coordinate pairs. Figure 1 shows a water body at Nizamabad and Adilabad through a complex polygon structure.
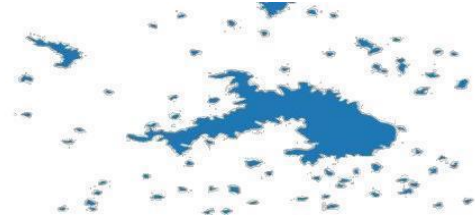

Figure 1. Polygon Representing Water Body at Nizamabad and Adilabad

4. METHODOLOGY

4.1 Creation of Database for Vector Analysis

Creation of Database for Vector Analysis Database for the vector data processing is done in QGIS framework which includes the coverage of Maharashtra water bodies whose area is greater than 3000 m. Figure 2 shows workflow of this study. Figure 3 (a) and (b) display the water bodies listed in Bhuvan Geo-platform's Water Body Information System (WBIS), representing Maharashtra state and Jayakwadi respectively.
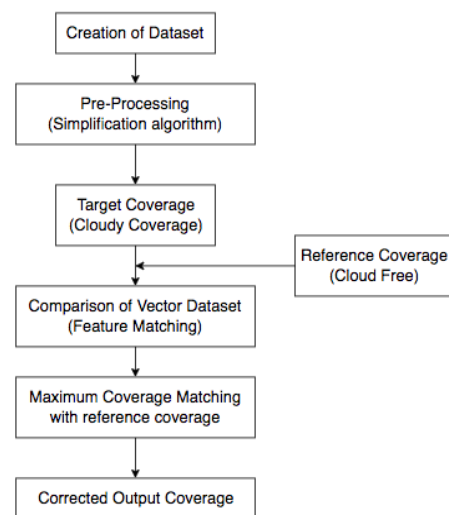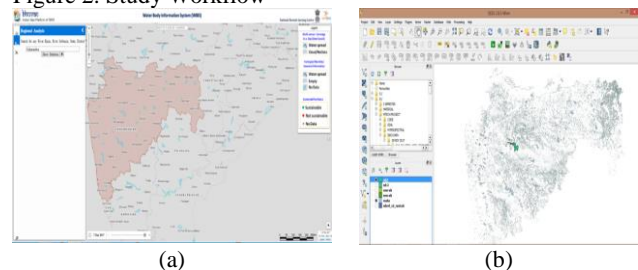

Figure 2. Study Workflow

(a)   (b)

Figure 3. (a) Showing different water bodies from Maharashtra State in Bhuvan portal, (b) Selection Specific District (Maharashtra) to create a particular Subset of Database

4.2 Simplification of Multi-Scale Vector Polygon Features

For comparing the two polygon vector data sets for the feature matching purpose, geometric characteristics of polygon play an important role. Some nodes of polygon feature also play a significant role. And among the outline feature point of the polygon, some decisive points can be extracted to recognize shapes. The specific points (shown in Figure 4) are the articulation points of the polygon (Shao and Tong, 2010). During the process of gradual change of gradient, theses points can decide the geometric shape of elements, known as the principle points. The rest of the point also characteristics point, less important than the principal points (Xie et al., 2010). In order to meet the actual requirement for GIS data producing, an adaptively improved algorithm based on Douglas-Peucker is used, which simplifies the vertices of polygon without changing its actual features. It also uses a parameter, which is distance threshold-T, to increase the matching accuracy.
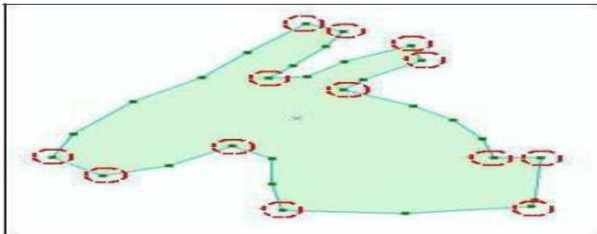

Figure 4. Principal Points in Polygon Features

The delineated method is implemented in the QGIS framework through visual inspection, quantitative measures and applications in water body contours. QGIS speeds up the data processing workflow with acceptable results as shown in Figure 5 (QGIS-2.8-UserGuide, 2018).

Figure 5. Polygon Edges Simplification

4.3 Providing the Distance Tolerance

In GIS, distance tolerance is called as buffer. Buffering is the process of creating one or more zones around selected features, within a pre-specified distance from these features (smith et al., 2015). Buffering usually creates two areas, positive (outward) and negative (inward) buffers as shown in Figure 7. Here, positive and negative buffering operations are performed to create two polygons, external and internal to the original geometry as shown in Figure 6. The distance tolerance used during this study is 70 meters. This method allows the comparison between two datasets with specified distance tolerance, which can be deemed also as buffering zone.
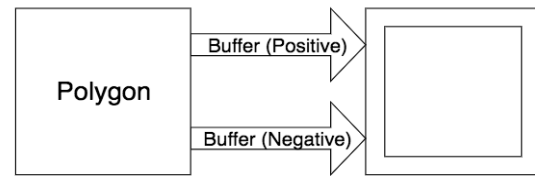

Figure 6. Demonstrates the effect of positive and negative buffering operations carried out on a simple polygon.
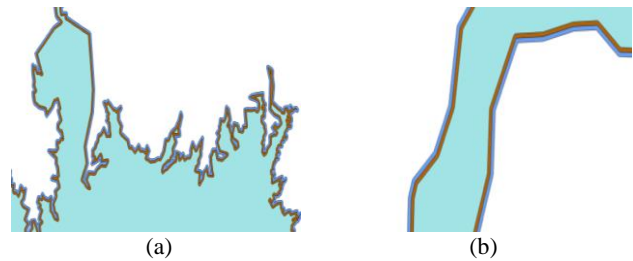

| (a) | (b) |

Figure 7. (a) Buffer Polygon, (b) Positive and Negative buffer to the original Water body (Shown in Brown Color)

4.4 Feature Matching

Feature matching means to identify identical features or entities in different sources through some similarity measures, which is a process of establishing correspondence between spatial features. The existing matching algorithms can be divided into geometric matching, topological matching, and semantic matching (Chan-Hee & Yun-Ho, 2011). During this study, the feature matching method implemented uses positive and negative buffering operations to blend with the basic topological test's setup, if the geometries match. Figure 8. represents the cloud-free reference coverage of a water body at Aurangabad district. Figure 9. represents the cloudy coverage of the same location. After the feature matching, corrected coverage is obtained of the same location as shown in Figure 10.

Figure 8. Water Body Coverage at Aurangabad-Cloud Free

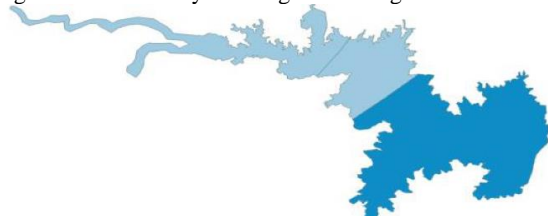Figure 9. Water Body Coverage at Aurangabad Cloud

Figure 10. Corrected Water Body Coverage at Aurangabad

5. CONCLUSION

This study paper provides the analysis to obtain cloud-free satellite images of water bodies. Feature matching of water bodies, in the form of spatial vector data, gives simplified workflow to retrieve cloud-free images. Furthermore, the algorithm that simplifies the complex polygonal structure of water bodies also removes the noise in original datasets and improves stability. This study is based on spatial vector data matching, which helps in updating more efficiently datasets of cloud-free water bodies.

The results of this study can be implemented on a large number of datasets in various remote sensing applications. This suggests that matching vector objects within a buffering zone is the proper method of relating pixel locations to geographical coordinates for the purpose of automatic recording of cloud free satellite images.

REFERENCES

Abd-Elrahman, A., Shaker, I.F., Abdel-Gawad, A.K., Abdel-Wahab, A., 2008. Enhancement of Cloud-Associated Shadow Areas in Satellite Images Using Wavelet Image Fusion. World Applied Sciences Journal 4(3), 363–370.

Abraham, L., & Sasikumar, M., 2013. Cloud Extraction and Removal in Aerial and Satellite Images. Intelligent Informatics, 341–347. doi:10.1007/978-3-642- 32063-7_36.

Bai, K., Chang, N.-B., & Chen, C.-F., 2016. Spectral Information Adaptation and Synthesis Scheme for Merging Cross-Mission Ocean Colour Reflectance Observations from MODIS and VIIRS. IEEE Transactions on Geoscience and Remote Sensing, 54(1), 311–329. doi:10.1109/tgrs.2015.2456906.

Bhuvan | Water Body Information System 2018. Bhuvan.nrsc.gov.in, http://bhuvan .nrsc.gov.in/gis/thematic/wbis/#!/map. (20 March 2018).

Borchartt, T.B., de Melo, R.H.C., Gazolla, J.G.F.M., Resmini, R., Conci, A., Sanchez, A., de A. Vieira, E. (2011). On the reduction of cloud influence in natural and aerial images. In: 18th International Conference on Systems, Signals and Image Processing, pp. 1–4.

Chan-Hee H, Yun-Ho K, 2011. Fast Shape Matching Using Statistical Features of Shape Contexts. IEICE Transactions on Information and Systems, pp. 2056-2058.

Chanda, B., & Majumder, D. D., 1991. An iterative algorithm for removing the effect of thin cloud cover from LANDSAT imagery. Mathematical Geology, 23(6), 853– 860. doi:10.1007/bf02068780.

Chen, B., Huang, B., Chen, L., & Xu, B., 2017. Spatially and Temporally Weighted Regression: A Novel Method to Produce Continuous Cloud-Free Landsat Imagery. IEEE Transactions on Geoscience and Remote Sensing, 55(1), 27–37. doi:10.1109/tgrs.2016.2580576.

Hargrove, W. W., Hoffman, F. M., & Hessburg, P. F., 2006. Mapcurves: a quantitative method for comparing categorical maps. Journal of Geographical Systems, 8(2), 187–208. doi:10.1007/s10109-006-0025-x Huh, Y., 2015. Local edge matching for seamless adjacent datasets with sequence alignment. ISPRS International Journal of Geo-Information 4: 2061–2077. Doi:10.3390/ijgi4042061.

Ik-Hyun Lee, & Mahmood, M. T., 2015. Robust Registration of Cloudy Satellite Images Using Two-Step Segmentation. IEEE Geoscience and Remote Sensing Letters, 12(5), 1121–1125. doi:10.1109/lgrs.2014.2385691.

Machiwal, D., Jha, M. K., & Mal, B. C., 2010. Assessment of Groundwater Potential in a Semi-Arid Region of India Using Remote

Sensing, GIS and MCDM Techniques. Water Resources Management, 25(5), 1359–1386. doi:10.1007/s11269-010-9749-y.

Martínez, F., Rueda, A. J., & Feito, F. R., 2009. A new algorithm for computing Boolean operations on polygons. Computers & Geosciences, 35(6), 1177–1185. doi:10.1016/j.cageo.2008.08.009.

Meng Qingxiang, Fu a Zhongliang, Huang b Yan and Shen Meicen, 2012. A FAST MATCHING APPROACH OF POLYGON FEATURES, *ISPRS Journal of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume I-2, pp.45-49.

Neware R., and Khan A., 2018. Survey on Classification techniques used in remote sensing for Satellite Images', in 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA 2018), pp.1860-1863.

Neware, R. and Khan, A., 2018. "IDENTIFICATION OF AGRICULTURE AREAS IN SATELLITE IMAGES USING SUPERVISED CLASSIFICATION TECHNIQUE," INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT), 6(2), pp.682-688, 2018.

QGIS-2.8-UserGuide-en.pdf | Geographic Information System | Command Line Interface. Scribd. https://www.scribd.com/document/353653928/QGIS-2-8-UserGuide-en-pdf. (10 February 12, 2018).

QGIS-2.8-UserGuide-en.pdf | Geographic Information System | Command Line Interface. Scribd. https://www.scribd.com/document/353653928/QGIS-2-8-UserGuide-en-pdf. (February 12, 2018).

Shao S, Tong C, 2010. A matching method for multi-characteristic vector elements of complex polygon. 2010 International Conference on Multimedia Technology (ICMT), pp. 1-4.

Smith, Goodchild, Longley, 2015. Geospatial Analysis – 5th Edition, 2015. http://www.spatialanalysisonline.com/HTML/index.html?simple_regression_and_trend_su.html

Xie J, Xu G, Zhang F, et al., 2010. A local tangent space-based approach for single-trial representation of event-related potentials. 5th Cairo International Biomedical Engineering Conference, CIBEC2010, pp.184-187.

Yuan, Y., & Hu, X., 2015. Bag-of-Words and Object-Based Classification for Cloud Extraction from Satellite Imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(8), 4197–4205. doi:10.1109/jstars.2015.2431676.