Adaptive Modeling on Satellite Image Processing and Info Extraction using Knowledge Fusion Mining

Dhayalan.D[1] Parvathi.P[2] Senthil Murugan .J[3]
[1]Assistant Professor [2]PG Scholar [3]Research Scholar
[1,2,3]Department of Master of Computer Applications
[1,2]Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala College of Engineering, Chennai, India [3]ST Peters University, Avadi Chennai, India

Abstract— Satellite Imagery is the most advent of landscape analysis. In existing system, a system has been designed to analyze the changes acquired in a particular landscape by utilizing satellite imagery of the landscape. This system formulates a Comparative analysis with historical and recent imagery of the landscape with respect to the changes in the soil, water, weather, landscape. The proposed system designs a system that formulates a comparative analysis with historical and the recent imagery by integrating the concept of image replacement. The resultant of the existing methodology does not predict or suggest anything for the future about the landscape whereas the proposed methodology does it with by means of image replacement. The satellite imagery under subject is tuned up with its part with a minute changes in the picture with a vision to the future and the comparative analysis has been made. Performance analysis has been done to the comparative analysis system with respect to the time and visualized graphically.

Key words: Knowledge Fusion Mining, Landsat and MODIS Imagery, Adaptive Modeling, Multispectral Topographical Algorithm, Apriori Algorithm

I. INTRODUCTION

The Landsat imagery was usually used to yield estimates, monitor crop condition, forest fire detection and land cover change mapping analysis alone. To compare and analyze [1] these images the spatial and temporal adaptive reflectance fusion model (STARFM) was used. Medium resolution sensors were used in the present approach that have a perfect abstraction resolution for vegetation mapping at the sector scale so as to predict the satellite detected pictures. The captured imagery within the urban areas was therefore terribly cloudy and with such a large amount of disturbances to capture, therefore in the planned system we tend to fails to spot the clarity of images. [2] Landsat scenes area unit regarding thirty fifth cloud lined on the average globally and chance of taking two cloud-free observations of Landsat imagery at southern Asia among 48 days is a smaller amount than hour. This present method has some restrictions. That are,

- [2] Landsat is proscribed by a 16-day get back cycle and this was created worse by cloud pollution in those images.
- STARFM requires radio-metrically, geometrical consistent of both Landsat and MODI which is difficult.
- STARFM also requires better pair of MODIS landscape images. MODIS at depths view does not match with Landsat therefore more precise co-registration between Landsat and MODIS is required.

To avoid these problems we introduce another system. This system invents a[1] comparative analysis with historical and also the recent imagery of the landscape with relevance the changes within the water resources, soil resources, weather condition and landscape. The planned system designs a system that make a comparative analysis with historical and the recent imagery by integrating the concept of image replacement.

The scope is to Comparative analysis with historical and also the recent imagery of the landscape with reference to the changes within the soil, water, weather, landscape. The satellite imaging beneath subject is tuned up with its give a second changes within the image with a vision to the long run and therefore the comparative analysis has been created. Performance analysis has been done to the comparative analysis system with relevance the time and unreal diagrammatically.

In planned system a potential solution for applications that need fine spatial resolution Adaptive modeling was introduced. Adaptive modeling blends Landsat and MODIS [3] knowledge to come up with artificial "daily" surface coefficient product at Landsat spatial resolution. It needs a minimum of two image pairs because the inputs into the formula. And perform three levels of changes to predict the exact solution. That are, Moisture level changes, Environmental level changes, Material level changes. It uses Apriori algorithm.

It is organized as follows. Section II describes the key concepts. The proposed system was described in Section III. In Section IV, we tend to gift experimental results, and section V concludes the research.

II. THEORETICAL BACKGROUND

A. *Multispectral Topological Algorithm (MSTA)*

A method of calculating an appropriate similarity value between two different satellite images is at the core of our methodology[4]. Existing research on map comparison be relevant to the detection of temporal changes, to comparison between different mapping methodologies, to validation of models, and to assessment of map accuracy.

In Multispectral image processing two image processing methods were used: Principal Component Analysis and Region of Interest Analysis. PCA is employed to estimate the first knowledge with lower dimensional feature vectors, since the info from multiple spectral bands usually involves a precise degree of redundancy. The basic approach is to compute the characteristic vector of a square matrix (eigenvector) of the covariance matrix, and estimated the uniquedata by a linear arrangement of the primary eigenvectors designed for a multispectral image[5] of k spectral channels each with m x n pixels, the image data is

efficient intotwo-dimensional array X of size s by k where as = (m x n). The matrix of covariance X is defined as:

$$C_x = \frac{\bar{X}^T\bar{X}}{s-1}$$

Where C_x is a matrix of size (k x k); and $\bar{X}$ is the (s x k) mean-centered matrix of X, resolute by the mean for first calculating in each column, and subtracting column mean since every value in that column. In the PCA decomposition, the p_i vectors stand eigenvectors of the covariance matrix C_x. On behalf of every p_i,

$$C_x p_i = \lambda_i p_i$$

Where k_i is the eigenvalue related with the eigenvector p_i. Each principal component PC_i is denoted by,

$$PC_i = X_1 p_{1i} + X_2 p_{2i} + X_3 p_{3i} + \ldots + X_k p_{ki}$$

Each principal component may be a weighted total of the k channel image.

B. Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM)

To better make use of Landsat and MODIS data, the Spatial and Temporal Adaptive Reflectance Fusion Model was fixed. It is charity aimed at spatial evidence since [6] fine-resolution Landsat imagery and sequential data from coarse-resolution, MODIS images to produce approximations of surface reflectance that are high resolution happening together space and time. In principle, the daily collection of MODIS imagery and seasonal Landsat imagery[9] allows the invention of synthetic "daily" Landsat-like views of Earth's surface.

The STARFM used associations of one or more sets of experimental Landsat-scale on other MODIS observation dates. STARFM was at first primarily established at NASA Goddard Space Flight Center. It has been greatly improved in computing efficiency for large area processing.

C. Apriori Algorithm

The Apriori algorithm was introduced by to discover frequent itemsets and association rules in a transactional dataset that contain support and confidence better than the user-specified minimum support and minimum confidence correspondingly. [7][8]Association rule has the form $X \rightarrow Y$, where X and Y are a subset *I*, I is a set of items, and $X \cap Y = \varnothing$.

L_k is a set of large k-itemsets. This set contains *k*-itemsets that have minimum support. C_k is a set of applicant k itemsets. Itemsets are probably huge itemsets. In the Apriori algorithm, the apriori-gen operates has the argument L_{k-1}that's the set of all giant (*k-1*) itemsets. The output of this function is a superset of this.

1) L1 = {large 1-itemsets};
2) for (k = 2; $L_{k-1} \neq \varnothing$; k++) do begin
3) C_k = apriori-gen (L_{k-1});
4) forall connections t $\in$ D do begin
5) Ct = subset(C_k, t);
6) forall candidates c $\in$ Ct do
7) c.count ++;
8) end ;
9) $L_k = \{c \in C_k \mid c.count \geq minsup\}$
10) end;
11) Result = $\cup k L_k$;

Three best extensively-used processes for choosing exciting rules, that are Support, confidence and lift. Support and confidence of the rule A →B are defined as follows:

$$Support(A \rightarrow B) = P(A \cup B) \quad (1)$$

$$Confidence(A \rightarrow B) = P(B|A) \quad (2)$$

support(A→B) is the percentage of transaction in a transactional dataset D that contain both A and b whereas Confidence(A→B) is the percentage of transaction in D is containing and that also contain B. Equation (2) is also stated as follows:

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)} \quad (3)$$

Calculate the correlation between A and B in the rule A→B, the correlation quantity High may be charity which is calculated as follows:

$$\text{lift}(A \rightarrow B) = \frac{P(A \cup B)}{P(A) \cup P(B)} \quad (4)$$

Based on the value of lift(A,B) in Equation (4), the relation of occurrence of A and B is described as follows. If lift(A,B) IS greater than 1, then A and B are absolutely correlated meaning that the occasion of A implies the occurrence of B. If(A, B) is fewer than 1, A and B are adversely correlated. A and B are independent if lift(A,B) is equal to 1. It means that there is no correlation between A and B.

III. PROPOSED SYSTEM

Adaptive modeling approach blends Landsat and MODIS data to make synthetic daily seeming reflectance produces at this spatial resolution. It require a minimum of two image pairs as the inputs into the algorithm[3]. And perform three levels of changes to predict the exact solution. That are, Moisture level changes, Environmental level changes, Material level changes. It uses Apriori algorithm.

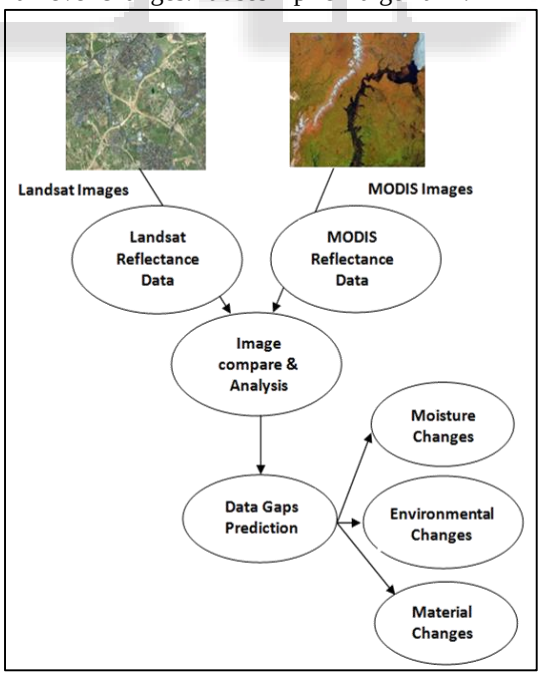

Fig. 1: The overall concept of proposed system

This Adaptive modeling focuses on cloud free images from data of Landsat and MODIS. Adaptive modeling improves finer spatial resolution of images. It provides better handle varied pixels level if unpolluted neighbor pixel occurs[4]. The Adaptive modeling with improved approaches have been successfully used in

different fields such as generating dense time-series synthetic. It has a monitoring forest state, appraising gross primary Productivity, mapping daily evapotranspiration, improving classification of conservation tillage and analyzing dry land forest phenology. The overall concept of proposed system was shown in Fig (1)

IV. EXPERIMENTAL RESULTS

In this proposal, we first get input as LANDSAT and MODIS imagery. Then these imagery are involved in grayscale conversion and RGB validation. Then histogram algorithm was used to get the exact information from the imagery. Histogram algorithm helps to get the pixel information of imagery without any confusion. These are performed as image validation check of satellite imagery.

To obtain reflectance data, Landsat and MODIS imagery uses this Algorithm[3].

Later the images are associated and analyzed using Multispectral Topographical Algorithm (MSTA) [7]. It produces the report of data gap prediction. Using this imagery data we perform three level of changes to obtain moisture level, environmental level and material level data[10].

In these processes, if any problems are raised during execution the image replacement and object removal methods are used to prevent obtain wrong kind of results.

Finally Performance Analysis was performed to measure the details of outcome. The overall working flow of this proposed system was shown in Fig (2).

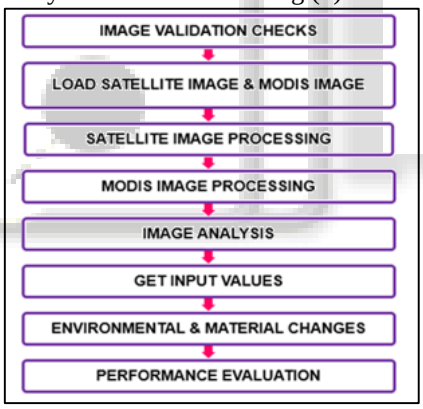

Fig. 2: Work Flow of System

A. *Landsat & Modis Imagery Scan Module*

Landsat have an ideal resolution[6] which is much suitable for land use and land cover change mapping, crop condition monitoring and yield estimation, forest fire detection and global ecosystem carbon-cycle studies.[2] Landsat is limited by a 16-day revisit cycle and it is very difficult to acquire cloud-free remotely sensed data with ideal resolution.

It helps to obtain data that are presence used to a rise products reaching from undergrowth, land surface cover, fire occurrence, snow cover and sea ice cover on the lands and oceans. It is 36-band spectro-radiometer measuring visible and infrared radiation.

In this module we check the image as valid or not by using histogram algorithm. If images are valid then reflectance data are gathered using STARFM. If image is not valid that was removed using image removal or object replacement algorithm. Fig (3) shows how MODIS imagery processed.

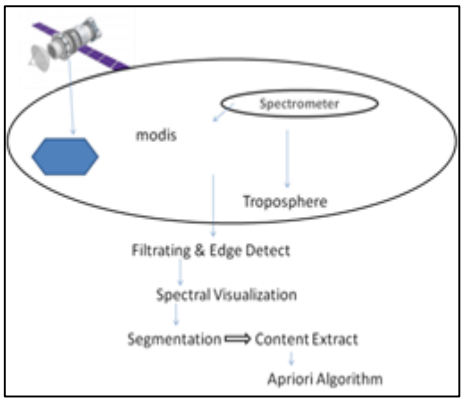

Fig. 3: MODIS imagery Processing

B. *Image Analysis & Compare Module*

Image taken by LANSAT and MODIS are blend [3]together and analyzed by adaptive modeling to get dense temporal information about the region. Adaptive modeling requires a pair of MODIS and Landsat images and these images should be radio-metrically, geometrically consistent. The MSTA (Multispectral Topographical Algorithm) was used to compare the reflectance data. Fig (4) shows the process involved in this model and the outcome.

C. *Data Gaps Prediction Module*

Due to cloud contamination in the images some pixels[5] in the MODIS BRDF/albedo algorithm show filled values and have data gaps in them.For these gaps, we used substituted Bidirectional Reflectance Distribution Function (BRDF) parameters, which were obtained from a BRDF lookup table based on the MODIS International Geosphere–Biosphere Programme (IGBP).

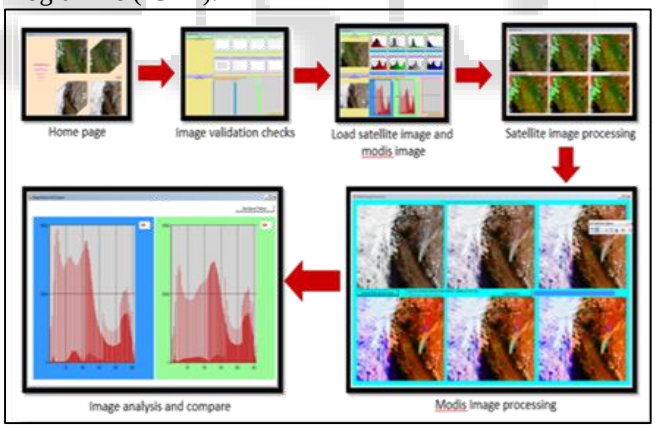

Fig. 4: Report of image analysis and Comparison

D. *Moisture Changes Module*

Agricultural remote sensing modeling[2] is also an important factor in landscape prediction that determines the effects of climate (moisture) change on agriculture. Here, we focus on the wheat growth season, from October to the following May[9].The data fusion tests in these sites focus on monitoring crop growth and vegetation based on seasonal variability.

E. *Environmental Changes Module*

To tests environmental changes happened in particular landscape, MODIS and Landsat data pairs from the same period (season or climate) and year, the same period of two different years, and different period from adjacent years[9].

The accuracy of the expected results depends on the information consistency between the MODIS depth view bidirectional-reflectance distribution perform and Landsat surface coefficient on each the paired dates and therefore the prediction dates. Based on the above made prediction the environmental changes can be found certainly.

F. Material changes module

The adaptive modeling approach was modified into Spatial and Temporal Adaptive Algorithm for mapping Reflectance Change (STAARCH) [9]. This algorithm helps to detect reflectance changes associated with land cover change and disturbance. Land cover change will simultaneously lead to material change in that region [2].

G. Performance Evaluations

Here we will evaluate the performance of the Satellite Imagery based comparative analyzer for landscape [8] prediction the data system. Two models that estimate monthly evapotranspiration area unit relatively evaluated so as illustrate however the counseled strategies are often applied. The performance evaluation module will eradicate the overall performance of the Data gaps prediction [8] and level of changes in the earth.

V. CONCLUSION

This system predicts a[1] Comparative analysis with historical and the current imagery of the landscape with deference to the changes in the soil, water, weather, landscape by integrating the notion of image standby with Apriori algorithm and Multispectral Topographical Algorithm.

The satellite imagery under subject is tuned up with its part with a minute changes in the picture with a vision to the future and the comparative analysis has been made.

REFERENCES

[1] "Comparative Analysis of Image Quality Assessment Using HVS Model" ,International Journal of Innovative Research in Computer and Communication Engineering(An ISO 3297: 2007 Certified Organization) Vol. 2, Issue 7, July 2014
[2] M. Datcu, H. Daschiel, A.Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P. G. Marchetti, and S. D'Elia, "Information miningin remote sensing image archives: System concepts, "IEEE Trans. Geosci. Remote Sens., vol. 42, no. 12, pp. 2923–2936,Dec. 2003.
[3] Gao, F., Masek, J., Schwaller M. and Hall,F. "On the blending of the Landsat and MODIS surface reflectance".
[4] M. Lew, N. Sebe, C. Lifi, and R. Jain, "Content-based multimedia information retrieval: State of the art and challenges," ACM Trans. Multimedia Comput., Commun., Appl., vol. 2, no. 1, pp. 1–19, 2006.
[5] T. Gevers and A. W. Smeulders, "Content-based image retrieval: An overview," in Emerging Topics in Computer Vision, G. M. S. B. Kang,Ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2004, ch. 8, pp.333–384.
[6] "Image Enhancement and Feature Extraction Basedon Low-Resolution Satellite Data"- Vasileios Syrris, Stefano Ferri, Daniele Ehrlich, and Martino Pesaresi- "IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. 8, No. 5, May 2015"
[7] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules," Proceedings of 20th International Conference on Very Large Data Bases, VLDB, 1994, pp. 487-499.
[8] J. Han and M. Kamber, "Data Mining: Concepts and Techniques," 2nd Edition, Morgan Kaufmann, 2006.
[9] Safa Réjichi, Ferdaous Chaabane and Florence Tupin,– "Expert Knowledge-Based Method for Satellite Image Time Series Analysis and Interpretation"-" IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. 8, No. 5, May 2015"
[10] "LandEx—A GeoWeb Tool for Query and Retrieval ofSpatial Patterns in Land CoverDatasets", Tomasz F. Stepinski, PawelNetzel, and Jaroslaw Jasiewicz.

