Design and Development of Framework to Create Cloud-Free Reference Repository of IRS Satellite Images for Image Analysis

D S Haripriya
M.Tech (CSE)
KIET-W
NRSC (ISRO)

Sai Durga
ASST. PROF
KIET-W

Mani Kumar Vedantam
Scientist-E
NRSC (ISRO)

Manju Sarma
SCIENTIST-F
NRSC (ISRO)

ABSTRACT

Technological advancement and increasing availability of high-resolution satellite images, offer the potential for more accurate analyses, by which the detection and quantification of changes could improve greatly. Such remotely-sensed products are expensive and difficult to acquire, which not only prohibits but restricts their use to minimum. This paper aspires to create a reference repository of cloud free satellite images using data received from on board optical sensors of Indian Remote Sensing satellite. By treating the received data with radiometric and geometric correcting processes at various levels, clear images are generated to user environment for early detection of anomalies, mostly during the first level of processing itself. A reference repository will be created using radio-metric cloud free sub-sampled raw images of size 1K x1K, along with terrain classification tags. Despite, the reference repository framed can be used for developing reference based change algorithms and alerting data product generation system for scene specific corrections, radiometric imbalances, product quality degradation etc. The created reference repository in this paper contains abundant database with spatial data types, data structures and features. Also the sample is suitable for queries on geographic applications and may be utilized in clarifying knowledge-based queries, as a part of User Order Processing System (UOPS)..

General Terms

This paper can be classified under Remote Sensing and Digital Image Classification.

Keywords

Push-Broom Sensors, Remote Sensing, and Path-Row based referencing scheme, Scene, IRS- Indian Remote sensing , Resourcesat-2.

1. INTRODUCTION

Since the launch of the earth's remote sensing satellite technology, remotely sensed imagery has been increasingly used to monitor Earth's ecosystems by quantifying changes. To perform and evaluate relative change related parameters like detection of dead detectors (on predefined scenes), noise detection and estimation, detector anomalies such as streaking , saturation and radiometric change detecting algorithms users have to access the global reference image and run the classification algorithm for each and every given browse image. This is an arduous task. This paper aspires to develop a reference repository which provides access to good quality Resourcesat-2 satellite imagery thus facilitating in analysis and sharp decision making. It helps to browse the images of Indian Remote Sensing Satellites and gather its information. The inaccuracies due to cloud, seasonal variation and framing errors are suitably addressed through this repository. RAD corrected full scene Reference is used for alerting the images chain using cloud free scenes. This reference is used by raw data quality detection algorithms.

2. PROBLEM DESCRIPTION

Our problem domain contains problems related to processing of Indian Remote Sensing Satellite data. This processing involves initial processing of the Remote Sensing data before data product generation to detect gross framing errors, BBR errors, cloud estimation, detector anomaly like streaking, saturation. Therefore it is mandatory to frame a cloud-free reference repository of the satellite image. Otherwise, users may have to access the global reference image and run the classification algorithm for every given image and later do the analysis.

3. REMOTE SENSING AND DIGITAL IMAGES

3.1 Remote Sensing

Remote sensing is the science of obtaining information about remotely located objects or areas from predetermined distance, typically from aircraft or satellites. Remote sensors are required from these regions. Remote sensors collect data by detecting the energy that is reflected from Earth. These sensors can be installed on satellites or mounted on aircraft. Remote sensors can be either passive or active. Passive sensors

"Respond **to external stimuli**". They record radiation that is reflected from Earth's surface, usually from the sun. Because of this, passive sensors can only be used to collect data during daylight hours. In contrast, active sensors use **"internal stimuli"** to collect data about Earth's Surface. The output of a remote sensing system is usually an image representing the scene which requires further analysis and interpretation in order to extract clear and useful information. In Satellite remote sensing of the Earth, the sensors are seen "**through a layer of atmosphere** '' separating the sensors from the Earth's surface under observation. Whereas In Optical Remote Sensing, "optical **sensors detect**" solar radiation reflected or scattered from the earth, forming images. Now the governing factor is ' **WAVE LENGTH PLAYS VITAL ROLL FOR DIFFERENTIATING'** The wavelength region usually extends from the V**isible and Near Infra-Red** (commonly abbreviated as**VNIR**) to the **Short-Wave Infra-Red** (**SWIR**).[1]

3.2 Digital Image

Remote sensing images are normally in the form of digital **images**. An image is a two-dimensional representation of objects in a real scene. Remote sensing images are representations of parts of the earth surface as seen from space. The images may be of raster or vector type. Satellite images are digital images. A digital image is a two-dimensional array of pixels. Each pixel has an intensity value (represented by a digital number) and a location address (referenced by its row and column numbers).

3.3 Image File Formats

Remote sensing images are representations of parts of the earth surface as seen from space. Remote sensing images are normally in the form of digital **images**.An image is a two-dimensional representation of objects of a real scene.The images may be of **raster or vector** type. Satellite images are digital images. A digital image is a two-dimensional array of pixels. Each pixel has an intensity value (represented by a digital number) and a location address (referenced by its row and column numbers).

3.3.1   Raster Format

In raster data the entire area of the map is subdivided into a grid of tiny cells. A value is stored in each of these cells to represent the nature of whatever is present at the corresponding location on the ground. Raster data can be thought of as a matrix of values. There are various file formats as **TIFF (also known as TIF), file types ending in .tiff,JPEG (also known as JPG), file types ending in .jpg, GIF, file types ending in .gif, PNG, file types ending in .png , Raw image files**

GeoTIFF Image

"GeoTIFF" refers to TIFF files which have geographic (or cartographic) data embedded as tags within the TIFF file. The geographic data can then be used to position the image in the correct location and geometry on the screen of a geographic information display. GeoTIFF is a metadata format, which provides geographic information to associate with the image data. But the TIFF file structure allows both the metadata and the image data to be encoded into the same file. GeoTIFF makes use of a public tag structure which is platform interoperable between any and all GeoTIFF-savvy readers. The TIFF imagery file format can be used to store and transfer digital satellite Imagery. GeoTIFF implements the geographic metadata formally, using compliant TIFF tags and structures. [2]

3.3.2   Vector Format

In vector data the features are recorded one by one, with shape being defined by the numerical values of the pairs of **xy** coordinates. A **point** is defined by a single pair of coordinate values. A **line** is defined by a sequence of coordinate pairs defining the points through which the line is drawn. An **area** is defined in a similar way, only with the first and last points joined to make a complete enclosure. Vector data can be thought of as a list of values. Vector image allow more flexibility. They contain geometric description.

4.   THE INDIAN REMOTE SENSING SATELLITE, RESOURCESAT-2

Resourcesat-2 was launched on 20th April, 2011 along with other two satellites in PSLV-C16 launch vehicle.RESOURCESAT - 2 is a follow on mission to RESOURCESAT - 1 and the eighteenth Remote Sensing satellite built by ISRO. RESOURCESAT - 2 is intended to continue the remote sensing data services to It has three cameras of **push broom type** mounted on a single platform with a high resolution sensors LISS-IV, medium resolution LISS-III and a coarse resolution AWiFSIt provides continuity of remote sensing data for various national and international projects .[3]

4.1   Sub System

Resourcesat-2 is a three axes body stabilized spacecraft launched by PSLV into a Sun Synchronous Orbit at an altitude of 817 Km. The spacecraft is designed for a nominal mission life of five years.

Resourcesat-2 carries three optical cameras as payload.

4.1.1   Linear Imaging Self Scanning Sensor (LISS-IV) Camera

The LISS-4 multispectral high-resolution camera is the prime instrument of this sensor complement. LISS-4 is a three-band push broom camera of LISS-3 heritage (same spectral VNIR bands as LISS-3) with a spatial resolution of 5.8 m and a swath of 70 km. LISS-4 can be operated in either of two support modes:

• Multispectral (MS) mode: Data is collected in 3 bands corresponding to pre-selected 4096 contiguous pixels with a swath width of 23.9 km (selectable out of 70 km total swath). **Note:** the new swath width of LISS-4 is 70 km for band 3 - but in mono mode only.

• Mono mode: Data of the full 12 k pixels of any one single selected band, corresponding to a swath of 70 km, can be transmitted. Nominally, band-3 data (B3) are being observed and transmitted in this mode.

It does have SWIR so unable to differentiate snow and cloud.

4.1.2   Linear Imaging Self Scanning Sensor (LISS-III) Camera

The medium resolution multi-spectral sensor, LISS-III is similar to that of LISS-III of Resoucesat-1 and operates in four spectral bands - B2, B3, B4 in visible near infrared (VNIR) and B5 in Short Wave Infrared (SWIR) providing data with 23.5m resolution.

4.1.3   Advanced Wide Field Sensor (AWiFS) Camera

AWiFS operates in four spectral bands identical to LISSIII, Providing a spatial resolution of 56 m and covering a swath of 737 Km. to cover this wide swath, the AWiFS camera is split into two separate electro-optic modules, AWiFS-A and AWiFS-B. [4]

All the three sensors specifications are given in Table 1.

4.2   Satellite Parameters

Satellites have several unique characteristics which make them particularly useful for remote sensing of the Earth's surface.The parameters which are used to specify the specification of a satellite are called satellite parameters.

4.2.1   Orbit

The path followed by a satellite is referred to as its **orbit**. Satellite orbits are matched to the capability and objective of the sensor(s) they carry. Orbit selection can vary in terms of altitude (their height above the Earth's surface) and their orientation and rotation relative to the Earth.

4.2.2   Path

An orbit is the course of motion taken by the satellite, in space and the *descending* ground trace of the orbit is called a 'Path'. There are 341 paths.

4.2.3   Row

The lines joining the corresponding scene centres of different paths are parallel to the equator and are called Rows. There are 49 rows.

Table1. Sensor Specifications

SPECIFICATIONS	AWIFS	LISS-III	LISS-IV
No. Of Bands	4	4	1(mono) , 3(MX)
Spectral Bands (µ)	B2 0.52 – 0.59 B3 0.62 – 0.68 B4 0.77 – 0.86 B5 1.55 – 1.70	B2 0.52 – 0.59 B3 0.62 – 0.68 B4 0.77 – 0.86 B5 1.55 – 1.70	B2 0.52 – 0.59 B3 0.62 – 0.68 B4 0.77 – 0.86 B3-default band for mono
Resolution (m)	56	23.5	5.8
Swath (km)	740	140	70/23
Revisit (days)	5	24	5
Data Rate (Mbs per stream)	105	105	105
Quantization	12-bits	10-bits	10-bits
Gains	100% Albedo No Gain Setting	100% Albedo No Gain Setting	100% Albedo No Gain Setting

Path-Row based world image is shown in Figure 3

4.2.4   Scene Definition

The camera scans the ground track line by line continuously. The satellite motion along the track

Provides continuous imaging of the ground. This continuous stream of data is segmented to convenient sizes. These segments are called scenes.The across track length of the scan (swath) is determined by the pixel size and number of elements in a line. Each imaging sensor scans line by line during its integration time, which is fixed for each camera. Thus, each camera scans a fixed number of lines in fixed intervals of time. Therefore, the along track length of a scene is based on the number of lines used to constitute that scene.
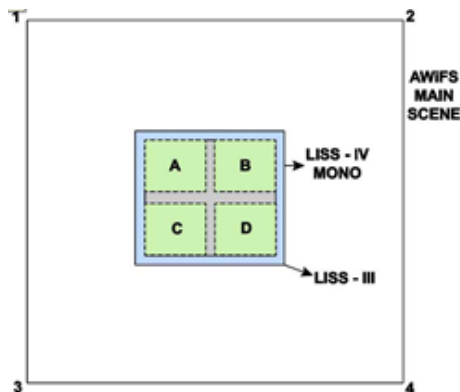
Figure 2 Scene layouts of LISS-IV and LISS-III in AWiFS main scene

4.2.5   Swath

As a satellite revolves around the Earth, the sensor "sees" a certain portion of the Earth's surface. The area imaged on the surface, is referred to as the **swath.** [5]

5.   REPOSITORY

We create a standard repository and a reference repository for the obtained cloud-free, geometrically corrected, Recourcesat images for the study region.

5.1   Standard Repository

5.1.1   GLC_2000

GLC_2000 is a worldwide accepted Geotiff file which taken as the reference for classification. It describes the land cover classes for the year 2000 based on the SPOT VEGA2000 dataset. Geotiff is essentially the TIFF format with the ability to embed georeferencing information with the image. Georeferencing means to 'define its existence in physical space.' Essentially, it establishes a relation between the image and a set of coordinates. Alternatively, it may also determine the location of other geographical features. GeoTIFF is a metadata format. It allows geographic information to be associated with the data of the image. Hence, the metadata and image data can be encoded into the same file, in lines with the TIFF format. The metadata can include map projection, coordinate systems, ellipsoids, datum's, etc. This Geotiff file contains 657881190 pixels which are classified into 23 standard classes . This Geotiff file is converted into Path – Row based reference scheme thus eliminating the use of coordinate system of georeferencing
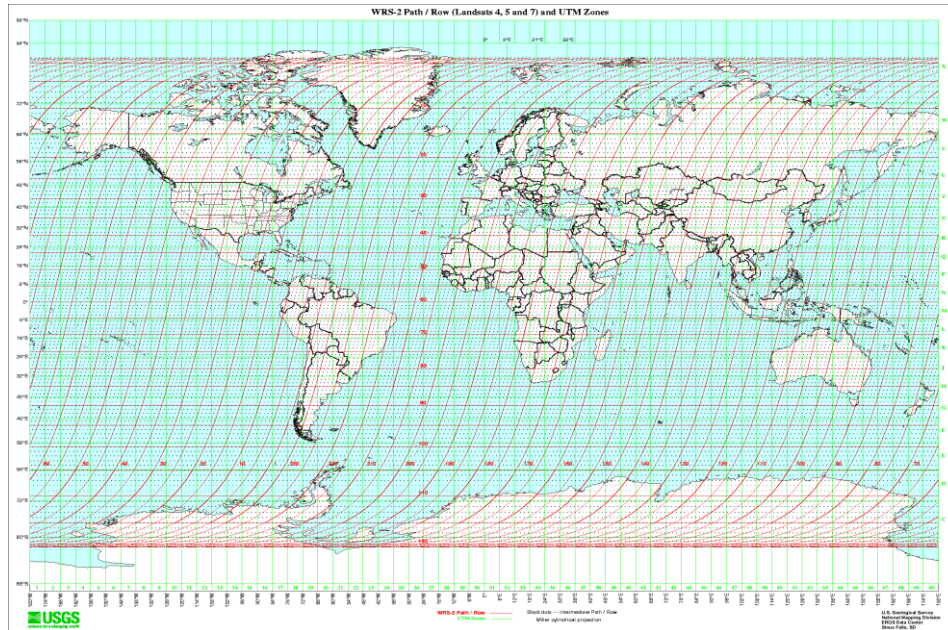
Figure 3 Path- Row based world image

The Global vegetation Monitoring Unit carries out several activities related to Land Cover mapping and monitoring. In particular the GVM Unit is coordinating and implementing the **Global Land Cover 2000 Project (GLC 2000)** in collaboration with a network of partners around the world. The general objective is to provide for the year 2000 a harmonized land cover database over the whole globe. The year Two Thousand is considered as a reference year for environmental assessment in relation to various activities, in particular the United Nation's Ecosystem-related International Conventions. [6]
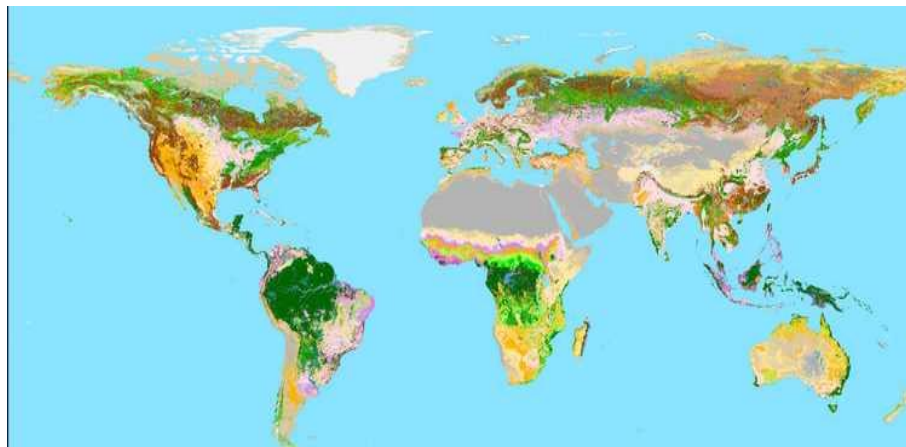
Figure 3   GLC_2000 Geotiff file

5.1.2   Reference Systems
Every position of earth can be represented using a reference system. Two such widely used reference systems are:

5.1.2.1 Coordinate Reference Systems (Cartesian Coordinates)
Cartesian coordinates are coordinates that measure the position of a point from a defined origin along axes that are perpendicular to each other.

5.1.2.2 Geodetic Reference Systems (Geographic Coordinates)
Geodetic coordinates (sometimes called geographic coordinates) are angular coordinates (latitude and longitude), closely related to spherical polar coordinates and are defined relative to a particular Earth geodetic datum.

5.1.3   Worldwide Reference System
The Worldwide Reference System (WRS) is a global notation system for satellite data. It enables a user to inquire about satellite imagery over any portion of the world by specifying a nominal scene center designated by PATH and ROW numbers. The WRS has proven valuable for the cataloging, referencing, and day-to-day use of imagery transmitted from the satellite sensors. WRS notation assigns sequential path numbers from east to west to 341 nominal satellite orbital tracks, starting with number 001 for the first track which crosses the equator . A specific orbital track can vary due to drift and other factors; thus, a path line is only approximate. The orbit is periodically adjusted after a specified amount of

drift has occurred in order to bring the satellite back to an orbit that is nearly coincident with the initial orbit.

The world reference scheme is a path-row referencing scheme used to eliminate the usage of latitude and longitude and facilitates convenient and unique identification of a geographic location. It helps the in georeferencing and mapping. It is useful in preparing accession and product catalogues and reduces the complexity of data products generation. Using reference scheme the users can arrive at number of scenes that cover his area of interest. [7]

LISS-IV senor has 4 sub scenes namely LISS4A, LISS4B, LISS4C and LISS4D.

AWiFS sensor has 4 sub scenes AWiFS-A, AWiFS-AWiFS-C and AWiFS-D along with the full scene AWiFS-F.
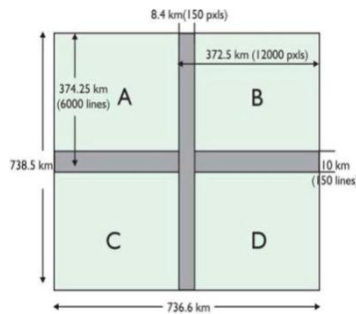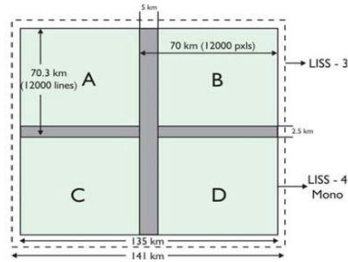

Figure 4   AWiFS Full scene and sub-scenes

Figure 5   LISS-III and LISS-IV scenes

5.1.4   GDAL

GDAL (Geospatial Data Abstraction Library) is a library for reading and writing raster geospatial data formats, and is released under the permissive X/MIT style free software license by the Open Source Geospatial Foundation. As a library, it presents a single abstract data model to the calling application for all supported formats. It may also be built with a variety of useful command-line utilities for data translation and processing. [8]

5.1.5   Metadata

Metadata is the data that describes about data. Meta is a prefix that in most information technology usages means "an underlying definition or description. "Metadata summarizes basic information about data, which can make finding and working with particular instances of data easier.

We generated a metadata is for the glc2000 Geotiff file using GDAL which describes the Meta of the file. This Geotiff file is in raster formatconsisting of a matrix of cells (or pixels) organized into rowsand columns (or a grid) where each cell contains a value representing the information of land cover classes. This raster Geotiff file is read using GDAL and the corresponding Meta is generated.
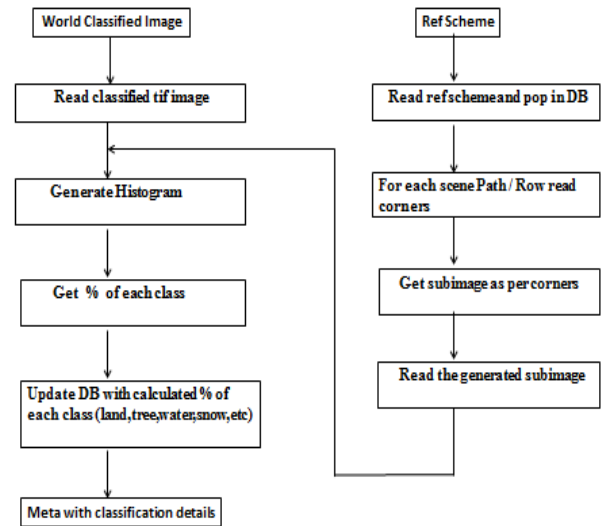

Figure 6   Algorithm for Standard repository

5.1.6   Subsetting

Subsetting is the operation that allows us to select a sub-window. Weextract the sub images from the Geotiff file by using the North-West and South-East coordinates from GRS. Extraction of a portion of image requires the conversion of the geodetic coordinates to Cartesian coordinates prior to extraction. We select a sub window from the source image with the corners given in geodetic coordinate system. Basically it helps in converting the geodetic coordinate system (GRC) to Coordinate Reference System (CRS), Path – Row based system thus eliminating the use of latitudes and longitudes longitude and facilitates convenient and unique identification of a geographic location. It helps in georeferencing and mapping. It is useful in preparing accession and product catalogues and reduces the complexity of data products generation. Using reference scheme the users can arrive at number of scenes that cover his area of interest.

5.1.7   Histogram and Database Population

We generate a histogram classification is for all the sub-images of the Geotiff file by using the metadata of the file. The classification consists of 23 bucket values corresponding to the 23 land cover classes of the tiff file. The bucket values corresponding to the classification are populated in the database along with the other parameters for all the sensors and their sub scenes.

5.2   Reference Repository

We obtained cloud-free, geometrically corrected Resourcesat - 2 imagery for the study region. We collected as many satellite images as possible.

5.2.1   Collect Cloud Free Images

We collected the cloud free image files for all the three sensors of Resourcesat-2 satellite. These files were in tarred zipped format.

5.2.2   Untar/Unzip Files

We untarred/unzipped all the files to get the browse images of the sensor's scenes and sub scenes and their metadata. They have to be unzipped two times as they are doubly zipped by proper coding or shell scripting. The meta data is in an unreadable rpt format.

5.2.3   Dump the images

Dumping the images means converting the metadata of the files which are in rpt format to a readable format using Browse Header file. Browse header file decodes the unreadable rpt format satellite metadata to a readable format dump file. The browse header specifies the path number, row number, scene, sub-scene, latitudes and longitudes, etc. of the satellite image. We then dump the images.
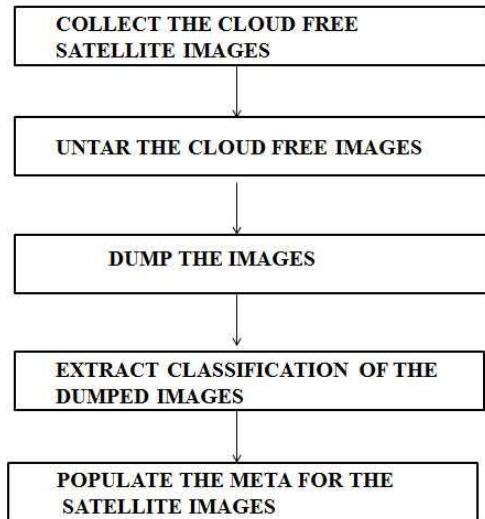

Figure 7   Algorithm for Reference repository

5.2.4   Extract Classification

We generate the classification for the all satellite images using the standard repository which in turn was created from the Worldwide accepted Geotiff file and the Worldwide Reference System(WRS). We then map the classification acquired from the standard repository across its   path and row numbers tothe satellite repository and populate the data into the satellite reference repository. This is the reference repository generated for user's reference contains all the satellite images and their corresponding classifications. This is the reference repository used as visual and statistical aiding tool for anomaly detections.

The reference repository is created for satellite Browse Images and for satellite Data Products. Satellite Data Products are high resolution images and Browse images are low resolution images.

5.3   Reference Repository – An Aiding Tool

We assert that reference repository would benefit from a better understanding of satellite imagery performance. The ability to detect with satellite imagery varies greatly, and missed detection can lead to false conclusions.

5.3.1   Auto-Cloud Generation

Several remote sensing satellites are currently available, providing imagery suitable for various types of applications. Each of these satellite-sensor platforms is characterized by the **wavelength bands** employed in image acquisition, LISS-IV sensor has only 4 bands and do not have SWIR band which is used for snow and cloud discrimination. Generally snow has less reflectance and cloud has higher reflectance. But in case of LISS –IV both have high values and it is difficult to discriminate them. In such cases the reference repository is used to find the amount of cloud.
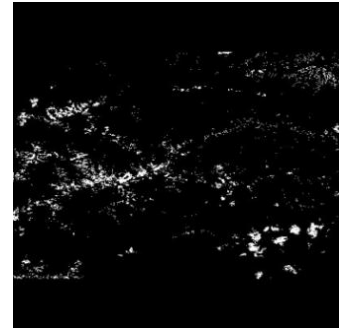

Figure 8   Image containing cloud and snow

We take a present satellite image having snow and cloud and a cloud free image from the reference repository to get the amount of cloud by subtracting the snow image from cloudfree image.

Figure 9 Clouds Free Image

AMOUNT OF CLOUD = SNOW IMAGE – CLOUD FREE IMAGE

When we subtract a snow and cloud image from the cloud free image, the snow portion gets eliminated and we are left with only cloud portion. It helps in estimation of the cloud by calculation the amount of cloud in the image.
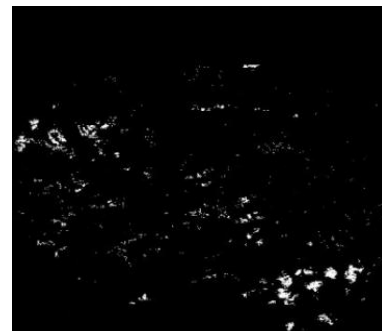

Figure 10 Amount of Cloud

5.3.2   Visual Aiding tool

Visual aiding tool is a tool that helps in detection of defects or changes by comparing the two images. The reference repository is a collection of good quality cloud free images.

5.3.2.1 Lens Defect Detection

In some cases we get lines or bands in an image. In such cases if we have the good quality of the same image in reference repository, it is easy to compare both the satellite images by just seeing them and detect the fault. We come to know the defect in lens. This helps in halting further processing of the faulty satellite image.

5.3.2.2 Radiometric and Geometric correction

The images received from the satellite may sometimes be blurred. In such cases, if we have the good quality of the same image in the repository, we can easily compare them and detect the need of radiometric and geometric corrections.
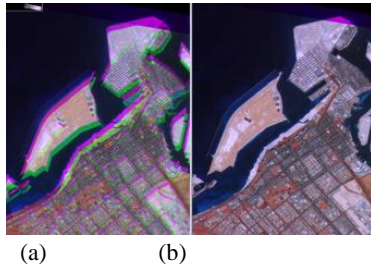

(a)   (b)

Figure 11   a) defective image b) repository image

5.3.2.3 Change Detection

Raw data received from the satellite undergoes various levels of processing. Thus there is chance to lose the actual data. The loss of actual data can be interpreted by comparing the classifications of both the given satellite image and the repository image.

5.3.3   Generating good quality Satellite Images

We are having a cloud satellite image and there is a need to generation of a good quality image, in that case we can replace the portion of the cloud from the same portion of the repository image to the cloud portion of the satellite image.

5.3.4   Change in Classification

The difference in the classifications of the repository image and the classifications generated by processing the latest satellite image concludes the changes in the classification of that particular area. For e.g. as per the repository classification there is no vegetation in the year 2012 for an AWiFS scene with path and row numbers 90 and 77 respectively. After 5 years if we again generate the classification of the same image and it shows some vegetation in that scene, then this depicts the changes occurring in the area having path and row numbers 90 and 77 respectively. Reference repository acts as an aiding tool to depict changes in environment.

5.3.5   Disaster Management

The reference repository contains the classifications of all the satellite images. Hence reference repository acts as a tool for alarming the disaster in disaster management. For e.g. in

5.3.5.1 Global warming

The % of ice in the year 2011 for a LISS-III scene in polar region having path number 55 and row number 15 was 90% according to reference repository classification. The % of ice in the year 2013 for the same LISS-III scene on processing gives 60 %. It depictsthe amount of ice melted hence showing the effect of Global warming on our environment.

5.3.5.2 Flood

The % of water in the year 2011 for an AWiFS scene in polar region having path number 87 and row number 90 was 60% according to reference repository classification. The % of water in the year 2014 for the same AWiFS scene on processing gives 80 % thus depicts the increase of water in

sea level. In such cases this reference repository can be used to alarm the disaster.

5.3.6   Statistical tool for correction

If a user processes the satellite image and prepares an algorithm to generate the classification of a satellite image, in this case the reference repository is used as a statistical tool to check the correctness of the algorithm by comparing the standard classification values and the values derived by him after generating the algorithm.

6. CONCLUSION

The reference repository created acts as a referencing scheme for the retrieving the classification information of the Indian Remote Sensing Satellite. Greater access to reference repository would therefore greatly facilitate in analysis and decision-making.

7. SCOPE

The reference repository can used for analysis, cloud cover-estimation, quality detection, to perform change related parameters like detecting dead detectors noise detection and estimation, detector anomalies like streaking, saturation, radiometric change detection, inaccuracies due to cloud, seasonal variation. Etc.

This is an evolving project which can be further extended for other satellites and sensors. Presently, we have created the reference repository for Resourcesat-2 satellite and LISSIV, LISS-3 and AWIFS sensor's Browse and Products. It can further be evolved for other satellites and sensors. Creation of reference repositories for other satellites and sensors is very useful and time saving.

8. ACKNOWLEDGEMENT

9. REFERENCES

[1] What is Remote Sensing? CRISP: http://www.crisp.nus.edu.sg/~research/tutorial/intro.htm.

[2] GeoTIFF FAQ Version 2.4, Author: Mike Ruth. www.remotesensing.org/geotiff/faq.html.

[3] RESOURCESAT-2 Data Users' Handbook: www.nrsc.gov.in/pdf/resourcesat2.pdf

[4] Resourcesat – 2, NRSC: bhuvan.nrsc.gov.in/bhuvan/PDF/Resourcesat-2_Handbook.pdf.

[5] Fundamentals of Remote Sensing, 2.4 Spectral Resolution: http://sar.kangwon.ac.kr/etc/fundam/chapter2/ch apter2_2_e.html

[6] GLC2000 Legend: bioval.jrc.ec.europa.eu/products/glc2000/legend.php

[7] The Worldwide Reference System « Landsat Science – NASA: landsat.gsfc.nasa.gov/?p=3231.

[8] GDAL: GDAL - Geospatial Data Abstraction Library: www.gdal.org/