Volume 2, Issue 1, January 2015 ISSN: 2348-8565 (Online)
International Journal of Modern Engineering and Research Technology
Website: http://www.ijmert.org Email: editor.ijmert@gmail.com

Low Level Feature Identification of Satellite Images and Knowledge Discovery From Identified Features using Association Rule Mining

Mukta Bhatele
Head of Department
Department of Computer Science & Engineering
Jai Narain College of Technology
Bhopal (M.P.), [INDIA]
Email: mukta_bhatele@rediffmail.com

Ankita Jain
Research Scholar
Department of Computer Science & Engineering
Jai Narain College of Technology
Bhopal (M.P.), [INDIA]
Email:anikita.jain0585@gmail.com

Sourabh Jain
Assistant Professor
Department of Computer Science & Engineering
Jai Narain College of Technology
Bhopal (M.P.), [INDIA]
Email: sourabhr2s@gmail.com

B. L. Rai
Associate Professor
Department of Computer Science & Engineering
Jai Narain College of Technology
Bhopal (M.P.), [INDIA]
Email:blrai_08_76@yahoo.co.in

ABSTRACT

All weather Satellite take images of the Earth in selected spectral bands that are in both the visible and the infrared portions of the electromagnetic spectrum. In addition, many Satellites provide different types of Satellite Images which are Infrared, Visible and Water Vapor Satellite Image. These three types of imagery are important, and, in some cases, all are needed to accurately interpret atmospheric conditions. These Satellite images contain information that are important for weather forecasting and early prediction of different atmospheric condition such as typhoon, hurricanes etc. that can be extracted by Content Based Image Retrieval (CBIR). Association rule mining is applied to these features to get the knowledge about the weather. In this paper shows the frequent item set of different cloud which are exist in three different types of satellite image. This is done by the image taken from the Website of Indian Metrological Department of different day of the same time.

I. INTRODUCTION

Knowledge discovery from satellite image is extraction of image data relationship, implicit knowledge, or other pattern not explicitly stored in images and uses ideas of image processing, image retrieval, machine learning database and data mining[2]. Challenge in Knowledge discovery is to identify low level feature containing in a pixel or group of pixels in an image, or image can be effectively and efficiently processed to identify high level spatial object and relationships. Knowledge discovery process may involves the steps of image mining which are preprocessing, transformation, feature extraction, Mining, Evaluation and interpretation and obtaining the final knowledge.

Satellite and satellite images

Satellite works as a messengers and observers in the sky. They relay watch the weather, telephones calls, guide ships and aircrafts and carry out tasks that are impossible on the ground [5].

Satellite Images are pictorial representation measuring the electromagnetic energy that is recorded by a sensor but not by photography. Photograph is normally taken within a certain spectral range (visible light) whereas Satellites take images outside this limited range[5].

Different Types of Satellite Images

Based on the Image Channel satellite Images can be categorizes in of three types:

InfraRed Satellite Images: Infrared imagery can show the amount of heat emitted by the different cloud features and the surface of the earth. Infrared images can also show clouds at higher levels better because they are colder. Infrared satellite can be measured by brightness temperature. The areas that are white are ones that are colder (emit less infrared light), and areas that are dark are warmer. Since temperature in the troposphere decreases with height and high level clouds are colder than low level clouds. The higher clouds appear brighter and low clouds appear darker on an infrared image [5][6][7].

Visible Satellite Images: Visible images can show the visible light that is reflected off of clouds and the surface of the earth. The Visible imagery is only taken during the day, in the area where the satellite is located. They show can all types of clouds and are the best types of images for seeing low level systems, which is not able show up well on infrared imagery. Visible satellite images provide the information about cloud cover. White area indicates clouds while shades of gray indicate generally clear skies. Higher reflectivity (or albedo) shows thicker clouds and appear brighter than thinner clouds on a visible image. It is difficult to distinguish among low, middle, and high level clouds in a visible satellite image, they can all have a similar albedo and for this distinction, infrared satellite images are useful [5][6][7].

Water Vapor Satellite Images: Water Vapor satellite images show water vapor in the upper troposphere. This is the only area of the atmosphere which is generally important in everyday weather forecasting. This type of images are useful for pointing out regions of moist and dry air, which provides information about the swirling middle troposphere wind patterns and jet streams. Greater area of moisture in the mid can be represented by brighter area on the image and upper level of atmosphere. Drier area of moisture in the mid and upper levels shows darker area. Very cold air can also show up fairly bright [5][6][7].

II. Methods Used in Feature Identification and Knowledge Discovery

CBIR to extract the features

Content based image retrieval (CBIR) is a method of retrieving images from a large dataset of image. CBIR is based on the low level visual features like texture, color and shape. Similarities between the images are calculated using the Euclidean distance.

$$|Q,T| = \sum |\omega_i - t_i| \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where

Q is query image and qi is low level feature of Q. T is a certain image in database and ti is low level feature of T. ω_i is the weight factor [3].

Color Feature /Grey level Extraction: Color Feature /Grey level Extraction represents the distribution of the pixels in the image over the gray-level scale. If each pixel is placed in a bin corresponding to the colour intensity of that pixel it is visualised. Then all of the pixels in each bin are added up and displayed on a graph. Histogram is a key tool in image processing which is one of the most useful techniques in gathering information of image. It is used to view the contrast of an image. If the grey-levels are well spread out, it defines a high contrast image. Likewise if they

concentrate near a certain level the image is low contrast

Texture Feature Extraction: For Texture Feature Extraction there are many techniques which are used such as co-occurrence matrix, fractals, Gabor filters, and variation of wavelet transform.

One of the most traditional techniques is Co-occurrence matrix for encoding texture information. It describes spatial relationships among grey-levels in a image. A cell defined by the position (i, j) in this matrix registers the probability at which two pixels of gray levels i and j occur in two relative positions. A set of co-occurrence probabilities (such as, energy, entropy, contrast, Correlation, Homogeneity) has been proposed to characterize textured regions.

Shape Feature Extraction: Shape is an important feature to identify and distinguish objects in pattern recognition,. Descriptors of Shape are classified into boundary-based (or contour-based) and region based methods [18].

Knowledge Discovery using Association Rule Mining

Association rule mining is one of the popular and well researched method for discovering interesting relations between variables. It is intended to identify strong rules discovered in databases using different measures of interestingness.

IV. PROPOSED WORK

All the three types of Satellite Images (all of the same time) are used in the proposed work. These Images are Visible Satellite Image, Infrared Satellite Image, and Water Vapor Satellite Image. All three types of imagery are important for different reasons, and, in some cases, all three are needed to accurately interpret atmospheric conditions [7].

The Satellite Images may look like an actual picture of the Earth, a weather satellite image is composed of thousands of points known as pixels.

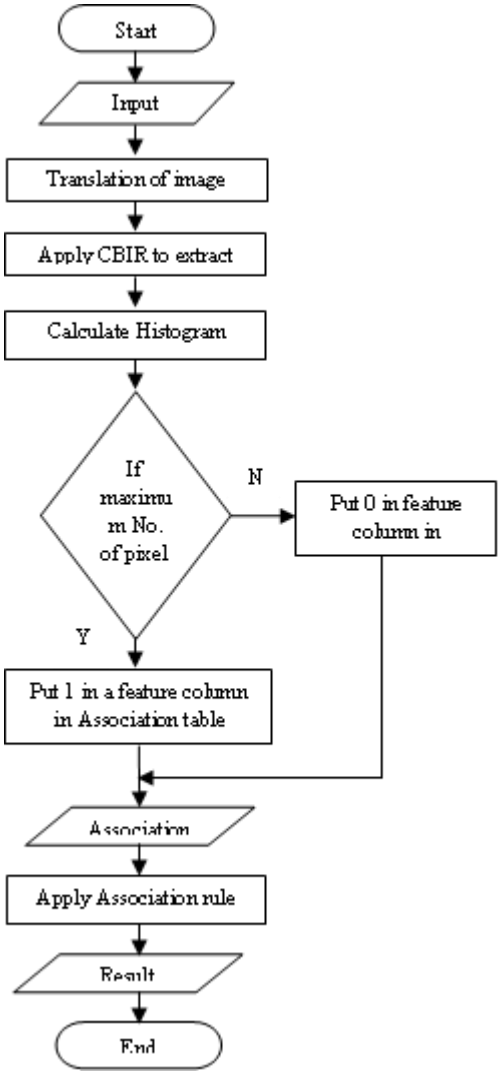

Figure 1.1 Flow Chart of Proposed work

Most weather satellite images are collected in a gray tone display. In this format, each pixel is assigned a tone that represents a level of energy (called the brightness value) sensed by the satellite. The tone is white, black, or an array of intermediate gray shades (known as gray scale). Typically, there are 256 possible brightness values or shades of gray in a satellite image. Different features on the Earth or in the atmosphere have different brightness values, therefore the relative brightness aids in

the identification of feature in a satellite image [7].

In all types of imagery the degree of contrast, or gray tone difference, between an object and its background is important. The greater the contrast, the easier it is to identify features in satellite imagery. When contrast is poor, enhancement techniques can be used to make accurate interpretation easier. To enhance an image, all the pixels in a specific range of brightness values are highlighted to locate the features of interest. For example, flood forecasters use IR imagery to look for specific cloud top temperatures that indicate heavy precipitation. By highlighting all the pixels with the corresponding brightness values, one can locate the area within a storm where heavy precipitation is most likely [7].

Finally, satellite images are often described in term of their resolution. Resolution refers to the size of the smallest feature that can be seen in an image. Since one pixel is the smallest element in an image, the area represented by one pixel is equal to the image resolution. Each pixel represents the average brightness over an area. Image resolution is determined by the satellite sensor, the type of transmission used, and also type of display hardware used to view the imagery.

Feature Extraction:

To extract the entire feature discussed above, CBIR (Content Based Image Retrieval) is used. CBIR generally retrieves the Image from database based on Query Image. To retrieve the images from database in CBIR, some features are extracted based on color, texture and shape from both, queried Image and database image and then feature of both images are matched. Those images that have same feature as queried image are get retrieved.

CBIR uses color, texture and shape to extract or retrieve the feature from the image. In the proposed work all the features will be extracted using gray level color and texture. Shape features will not be used to extract the feature because cloud has no specified shape.

Feature extraction from Infrared Satellite Image

Low cloud is extracted from an infrared satellite image if maximum number of pixel in the 255 bin is less than 4300 pixel. Likewise if maximum number of pixel in 255 bin is greater than 4300, the High Cloud is extracted from the Image. This is decided by observing the following table which shows the different day's images and there maximum no. of pixel in 255 bin.

Feature *extraction from Visible Satellite Image* Visible image gives two features which are Thin Cloud and Thick Cloud. Thick cloud is identified in image if there are no. of pixel in 255 bin is greater than 4300. Otherwise thin cloud is existing in the image. The fixed threshold is decided on the basis of observing the first two images.

Feature extraction from Water Vapor Satellite Image: Water Vapor gives two features which are moisture and dry. Moisture is identified in image if there are no. of pixel in 255 bins is greater then 4350. If number of pixel in 255 bin is less than 4350, the drier feature is extracted. The fixed threshold is decided on the basis of observing the all the images.

Knowledge Discovery using Association Rule:

All the features extracted from three types of satellite image using CBIR are collected in a table to create an Association Transaction table. In this table row represents image ID and column represents extracted feature from satellite images. Any cell contains 1 if feature exists in image otherwise 0. This table helps to

get the knowledge about the weather. Association Rule mining is applied on this table which help to know that which event can occur together.

V. RESULT

A. Association Table

In table 5.1, all the feature that is extracted from satellite images are shown and the last column shows the expected result.

Table 1 Association Rule

Image ID	High Cloud	Low Cloud	Moisture	Drier	Thick cloud	Thin cloud	Rainfall
16_08	Y	N	Y	N	N	Y	No
20_08	N	Y	Y	N	Y	N	No
22_08	Y	N	Y	N	N	Y	No
23_08	N	Y	Y	N	N	Y	No
24_08	Y	N	Y	N	Y	N	Yes
25_08	Y	N	Y	N	Y	N	Yes
26_08	N	Y	Y	N	Y	N	No
29_08	Y	N	Y	N	Y	N	Yes
30_08	Y	N	Y	N	Y	N	Yes

The above table is showing the association table in which columns are associated with the extracted feature and the rows are associated with image ID. Each cell in the table contain Y or N. Y represents the Cloud Feature which is specified in the table is exist in the image and N represents not exist.

B. Association Rule

Minimum support: 0.75 (7 instances)

Thick cloud=Y 7 ==> Moistured=Y 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Moistured=Y 7 ==> Thick cloud=Y 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Thin cloud=N 7 ==> Drier=N 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Drier=N 7 ==> Thin cloud=N 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Drier=N Thick cloud=Y 7 ==> Moistured=Y 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Thick cloud=Y 7 ==> Moistured=Y Drier=N 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Drier=N Thin cloud=N 7 ==> Moistured=Y 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Drier=N 7 ==> Moistured=Y Thin cloud=N 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Moistured=Y 7 ==> Drier=N Thin cloud=N 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Moistured=Y Thick cloud=Y 7 ==> Thin cloud=N 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Minimum support: 0.35 (3 instances)

Thick cloud=Y 7 ==> Moistured=Y 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

Moistured=Y 7 ==> Thick cloud=Y 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)

High Cloud=N 4 ==> Rainfall=No 4 <conf:(1)> lift:(1.67) lev:(0.16) [1] conv:(1.6)

High Cloud=Y Thick cloud=Y 4 ==> Moistured=Y 4 <conf:(1)> lift:(1.43) lev:(0.12) [1] conv:(1.2)

High Cloud=Y Moistured=Y 4 ==> Thick cloud=Y---------- 4 <conf:(1)> lift:(1.43) lev:(0.12) [1] conv:(1.2)

High Cloud=Y Moistured=Y 4 ==> Rainfall=Yes 4 <conf:(1)> lift:(2.5) lev:(0.24) [2] conv:(2.4)

High Cloud=Y Thick cloud=Y 4 ==> Rainfall=Yes 4 <conf:(1)> lift:(2.5) lev:(0.24) [2] conv:(2.4)

High Cloud=Y Moistured=Y Thick cloud=Y 4 ==> Rainfall=Yes 4 <conf:(1)> lift:(2.5) lev:(0.24) [2] conv:(2.4)

High Cloud=Y Thick cloud=Y 4 ==>

Moistured=Y Rainfall=Yes 4 <conf:(1)> lift:(2.5) lev:(0.24) [2] conv:(2.4)
High Cloud=Y Moistured=Y 4 ==> Thick cloud=Y Rainfall=Yes 4 <conf:(1)> lift:(2.5) lev:(0.24) [2] conv:(2.4)
High Cloud=N Moistured=Y 3 ==> Rainfall=No 3 <conf:(1)> lift:(1.67) lev:(0.12) [1] conv:(1.2)
High Cloud=N Thick cloud=Y 3 ==> Rainfall=No 3 <conf:(1)> lift:(1.67) lev:(0.12) [1] conv:(1.2)
Moistured=N Thick cloud=N 3 ==> Rainfall=No 3 <conf:(1)> lift:(1.67) lev:(0.12) [1] conv:(1.2)
Thin cloud=N 7 ==> Drier=N 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)
Drier=N 7 ==> Thin cloud=N 7 <conf:(1)> lift:(1.43) lev:(0.21) [2] conv:(2.1)
Low Cloud=N Thin cloud=N 4 ==> Drier=N 4 <conf:(1)> lift:(1.43) lev:(0.12) [1] conv:(1.2)
Thin cloud=Y 3 ==> Drier=Y 3 <conf:(1)> lift:(3.33) lev:(0.21) [2] conv:(2.1)
Drier=Y 3 ==> Thin cloud=Y 3 <conf:(1)> lift:(3.33) lev:(0.21) [2] conv:(2.1)
Drier=Y 3 ==> Rainfall=No 3 <conf:(1)> lift:(1.67) lev:(0.12) [1] conv:(1.2)
Low Cloud=Y Thin cloud=N 3 ==> Drier=N 3 <conf:(1)> lift:(1.43) lev:(0.09) [0] conv:(0.9)
Drier=Y Thin cloud=Y 3 ==> Rainfall=No 3 <conf:(1)> lift:(1.67) lev:(0.12) [1] conv:(1.2).

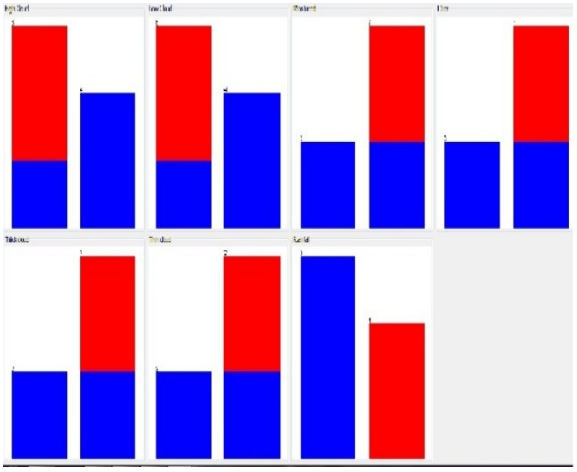

Figure: 1.2 Visualization of Input Data

C. Interpretation of association *rules*

An important conclusion of these rules is that if High cloud, Moisture and thick cloud are high then Rainfall is high.

- These rules represents that if low cloud, thin cloud and drier are low then rainfall is low.
- There is another possibility is that if moisture and thick cloud is high then rainfall is high.
- These rules show the strong relationship between High cloud, thick cloud and moisture.
- These rules also show the relation between low cloud, thin cloud and drier attributes.
- Rainfall and moisture is associated with thick cloud.
- If thick cloud is exist in the visible satellite image then there is possible that moisture is available in the water vapor image.
- This is possible that if drier exists in the satellite image then thin cloud is available in the visible image.
- High cloud is associated with the moisture. This is possible that if high cloud is high then moisture will also high.
- Thin cloud and low cloud are associated with drier.

VI. CONCLUSION

Images are taken from the metrological department of India which shows the atmospheric conditions. Low level features are identified for knowledge discovery using Association rule mining algorithm of data

mining. This works will show which events can happen together as a result.

REFERENCES:

[1] Yo-Ping Huang, Tsun-Wei Chang and Li-Jen Kao "Using Fuzzy SOM Strategy for Satellite Image Retrieval and Information Mining" CITSA 2007, Florida, USA. ISBN 1-934272-10-8.

[2] Y. Li, T. Bretschneider "Semantics-Based Satellite Image Retrieval Using Low-Level Features" IEEE 2004.

[3] Senduru Srinivasulu, P. Sakthivel "Extracting Spatial Semantics in Association Rules For Weather Forecasting Image" IEEE 2010.

[4] Image Mining: Trends and Developments http://www.comp.nus.edu.sg/~whsu/publication/2002/JIIS.pd.

[5] Web Reference: www.imd.gov.in

[6] Web Reference: http://ww2010.atmos.uiuc.edu/(Gh)/guides/rs/sat/img/vis.rxm

[7] An Introduction to Satellite Imagery Interpretation by Eric.

[8] Mahendra Kumar Gurve, Jyoti Sarup "Satellite Clout Image Processing and Information Retrieval System" IEEE 2012.

[9] R.M. Haralick, K. Shanmugam, I. Dinstein. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, Vol.3, No. 6, 1973, 610–621.

[10] Mahendra Kumar Gurve, Jyoti Sarup. "Satellite Cloud Image Processing And Information Retrieval System" 978-1-4673-4805-8 IEEE 2012.

[11] Data Mining Techniques by Arun K Pujari.

[12] Getachew Berhan, Tsegaye Tadesse, Solomon Atnafu, Shawndra Hill, "Drought Information Mining from Satellite Images for Improved Climate Mitigation" 987-1-4673-4805-8/12 IEEE 2012.

[13] Yikun Li and Timo R. Bretschneider, "Semantic-Sensitive Satellite Image Retrieval" 0196-2892 IEEE 2007.

[14] Seema H.Jadhav1, Dr.Shah Aqueel Ahmed "A Content Based Image Retrieval System using homogeneity Feature extraction from Recency-based Retrieved Image Library" *IOSRJCE, ISSN: 2278-0661, ISBN: 2278-8727 Volume 7, Issue 6 (Nov-Dec. 2012), PP 13-24.*

[15] YU Changhui, Yuan Yuan, Miao Minjing, Zhu Menglu "Cloud detection method based on feature extraction in remote Sensing images" International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-2/W1, 2013 8th International Symposium on Spatial Data Quality, 30 May - 1 June 2013, Hong Kong.

[16] Noureldin Laban, Ayman Nasr, Motaz ElSaban, Hoda Onsi "Spatial Cloud Detection and Retrieval System for Satellite Images" (IJACSA), Vol. 3, No. 12, 2012

[17] Y N Mamatha and A.G Ananth," Content Based Image Retrieval of Satellite Imageries Using Soft Query Based Color Composite Techniques",

IJCA (0975 – 8887) Volume 7– No.5, September 2010.

[18] P.S. Hiremath, S. Shivashankar, and Jagadeesh Pujari, "wavelet based features for color texture Classification with application to CBIR", IJCSNS International Journal of Computer Science and Network Security, VOL.6 No.9A, September – 2006

* * * * *