SPATIAL DATA MINING FOR PIXEL LEVEL CHANGE DETECTION

Chinu Jose, M.Tech GIS, NIIT University
Neemrana, NH-8, Delhi-Jaipur Highway
Alwar District, Rajasthan-301705

Abstract:

Spatial data mining is a highly demanding field because huge amount of spatial data have been collected by various applications such as remote sensing, GIS, environment assessment and planning and so on. Developments of raster data capture technologies and demands from application fields call for advanced raster data analysis methods. Raster datasets usually have huge amount of pixels which challenges the efficiency of spatial data mining algorithms. In this research paper an example of land use land cover in Alwar district of Rajasthan was presented to describe the course of spatial data mining. The paper mainly focuses on pixel level identification of change that has occurred in mutlidate raster data.7 years land use land cover data of Alwar was interpreted in a single go. All the land use land cover classes were analyzed individually and the change in pixel if any was recorded. The data mining process was made much simpler by automating the task using ArcObjects. The work done in this paper can be used to enhance the ability of interpreting data to generate useful knowledge regarding the land use land cover pattern over a period of time.

About the Author:

Ms Chinu Jose,MTech GIS
Currently pursuing MTech GIS at NIIT University BTech in Computer Science and Engineering from Mahatma Gandhi University,Kerala.

E mail ID: chinu.jose@st.niituniversity.in

Contact No: +91 8829844711
+91 9400454860

Introduction

In recent years, with the development of remote sensing and data storage technique, a great number of spatial data are generated every day, much of which is remote sensing image data. However, the use efficiency of the huge quantities of remote sensing image data is still low. It is very difficult for people to process with thousands of image data and find out knowledge from them. As the researches of data mining, information retrieval, multimedia database and other correlative field have rapid progress, it is become possible to manage and analyze large amounts of remote sensing images and find out useful information in different applications.

In this work a pixel level data search has been done. Identification of change in land use patterns was the core area in this research work. The land use land cover classified image for year 2004_05 (lulc04_05) is first analyzed for a particular category of land use or land cover. Now we search for the same pixel position in all the other images. The resulting information is recorded to form the final result.

Data Mining

Data mining is the process of employing one or more computer learning techniques to automatically analyze and extract knowledge from data collected in a large database. Its purpose is to identify trends and patterns in data so that users can extract hidden predictive information from the database. It is a powerful technology with great potential to help researchers focus on the most important information in their raw data.

Spatial data mining is the process of discovering interesting and previously unknown, but potentially useful patterns from large spatial datasets. Extracting interesting and useful patterns from spatial datasets is more difficult than extracting the corresponding patterns from non-spatial data. This is due to the complexity of spatial data types, spatial relationships and spatial autocorrelation. Performing data mining on raster data is even more challenging.

Remote Sensing Data Classification

District Alwar, Rajasthan was taken as the study area. AWIFS data for 7 years was chosen for the work. The image has to be classified so that it can be used for further analysis. The image classification process converts multiband raster imagery into a single-band raster with a number of classes. The imagery was first classified into several land use land cover categories such as built up, kharif, rabi, zaid, double/triple, current fallow, plantation/orchard, deciduous forest, scrub forest, Other wasteland, gullied, scrubland and water bodies. The focus of change detection in this study was mainly on five classes namely built up, cropped area, and degraded land, forest and water bodies. For this purpose the classified data was recoded into desired number of categories. Recoding is the process used for merging similar classes. Change detection can now be carried out on the thematic data with the desired land use land cover categories.
Change detection is one of the common image processing techniques. It can be done with the many image processing software. However a pixel level change detection or analysis cannot be achieved with them. In this work a pixel level identification of change was achieved.

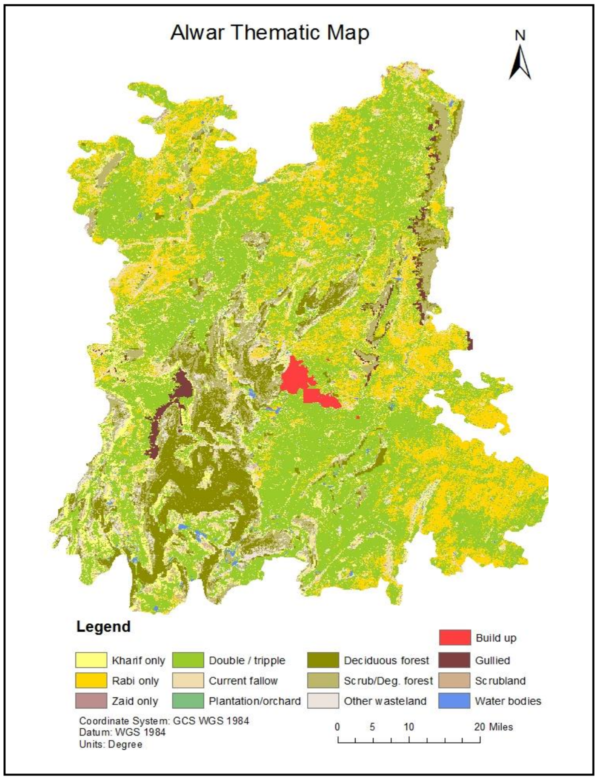

Fig 1
Alwar Thematic map with 13 classes

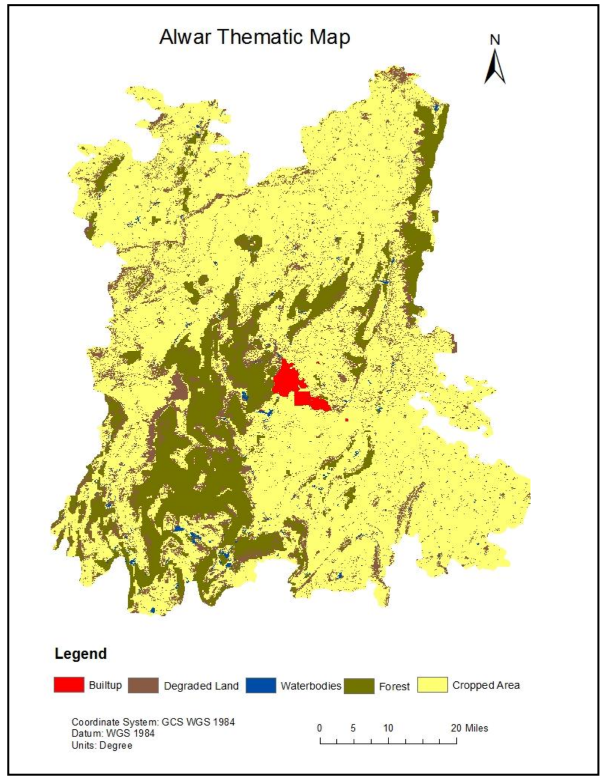

Fig 2
Alwar Thematic map with 5 classes

Flow of Remote Sensing Data Mining

i) Input

The algorithm for change detection requires certain inputs. The input has to be made ready by assigning proper naming conventions. Naming conventions are necessary so that the algorithm recognizes each image exactly while iteration. This includes source workspace, a land use class against which change detection has to be done and the result workspace. The source workspace should be a file geodatabase containing the 7 layers of classified imagery named lulc2004_05, lulc2005_06, lulc2006_07, lulc2007_08, lulc2008_09, lulc2009_10 and lulc2010_11 respectively. The input raster data should be in GRID format.

ii) Software Base

Change detection on multiple layers of raster data is performed by the addin created using ArcObjects. ArcMap 10.1 was used for prior mapping and ArcCatalog 10.1 was used for geodatabase preparation. Microsoft Visual Studio with ArcObjects SDK for .NET was used for creating ArcGIS addin in VB.Net.

iii) Process

An algorithm for accessing each pixel in the raster was used as the key basis for the change detection. Each and every pixel was identified and compared. The change detection on multiple layers will be done in comparison with the lulc2004_05 layer which is taken as the base layer. The result workspace is a file geodatabase which contains empty raster datasets. Once the necessary inputs are specified the algorithm proceeds by building a raster attribute table for the corresponding empty raster dataset which should ultimately contain the detected changes in land use category specified by the user. The attribute table will contain 7 fields which will store the pixel data for each year. It will also include the geographical location (latitude and longitude). The algorithm then performs iterations so that each pixel in the base year which corresponds to the given land use category is compared with the same pixel in all the other 6 layers and the information is recorded in the attribute table.

Result:

The analysis is performed on 7 years data as stated. The algorithm runs by comparing each of the 6 images with the image of the year 2004-05.The change is found out in terms of each class. For example, inorder to identify the change that has occurred to cropped area, the algorithm identifies the cropped area pixels in the base imagery (lulc2004_05) beforehand and the following iterations will check for the same pixels in the remaining 6 images. The results are stored into an attribute table simultaneously.

Figure 3 shows the user interface of the addin created for change detection. The user needs to specify the source and destination file geodatabase. A land use category has to be selected along with this so that the change in that particular class will be reflected.

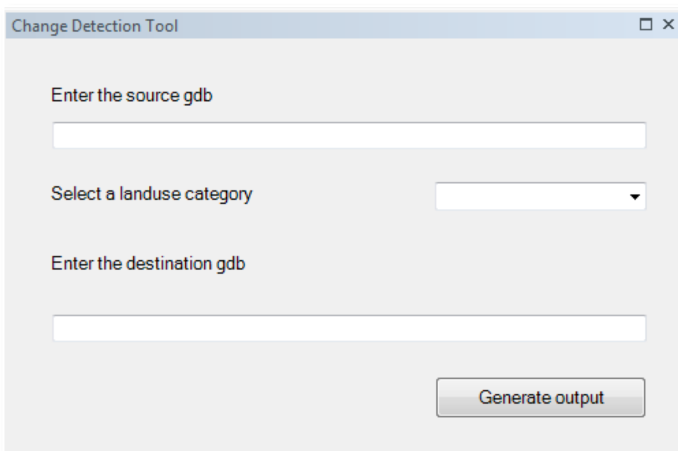

Fig 3
Change Detection Addin

The result obtained after change detection is shown in Table 1. This is in fact an attribute table generated in ArcMap as the result of executing the addin. This table shows the output obtained for the land use land cover category "degraded land". The table has been populated with the change information of a single pixel for 7 years. In figure , each record represents one pixel corresponding to 'degraded land' in the lulc2004_05 image and the change happened to the same if any in the following 6 years. Similarly , the change detection information of all the other classes can be derived.

degraded_land

OBJECTID *	lulc_2004_05	lulc_2005_06	lulc_2006_07	lulc_2007_08	lulc_2008_09	lulc_2009_10	lulc_2010_11	X	Y
1	Degraded Land	Degraded land	Degraded land	Degraded land	Forest	Forest	Forest	76.550915	27.697832
2	Degraded Land	Degraded land	Degraded land	Forest	Degraded land	Degraded land	Degraded land	76.552161	27.697832
3	Degraded Land	Forest	Forest	Forest	Forest	Forest	Forest	76.555896	27.697832
4	Degraded Land	Forest	Forest	Degraded land	Forest	Forest	Forest	76.556519	27.697832
5	Degraded Land	Forest	Forest	Degraded land	Forest	Forest	Forest	76.557142	27.697832
6	Degraded Land	Forest	Forest	Degraded land	Forest	Forest	Forest	76.587028	27.697832
7	Degraded Land	Forest	Forest	Degraded land	Forest	Forest	Forest	76.588896	27.697832
8	Degraded Land	Forest	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	76.590764	27.697832
9	Degraded Land	Forest	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	76.591386	27.697832
10	Degraded Land	Forest	Degraded land	Cropped area	Forest	Forest	Forest	76.593254	27.697832
11	Degraded Land	Degraded land	Cropped area	Degraded land	Cropped area	Cropped area	Cropped area	76.594499	27.697832
12	Degraded Land	Degraded land	Degraded land	Cropped area	Cropped area	Cropped area	Degraded land	76.595122	27.697832
13	Degraded Land	Cropped area	Degraded land	Cropped area	Cropped area	Cropped area	Cropped area	76.597612	27.697832
14	Degraded Land	Cropped area	Cropped area	Cropped area	Cropped area	Cropped area	Cropped area	76.598858	27.697832
15	Degraded Land	Cropped area	Cropped area	Cropped area	Cropped area	Cropped area	Cropped area	76.601348	27.697832
16	Degraded Land	Forest	Degraded land	Forest	Degraded land	Degraded land	Degraded land	76.60882	27.697832
17	Degraded Land	Degraded land	Forest	Degraded land	Degraded land	Degraded land	Degraded land	76.609442	27.697832
18	Degraded Land	Forest	Forest	Forest	Forest	Forest	Forest	76.613801	27.697832
19	Degraded Land	Forest	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	76.614423	27.697832
20	Degraded Land	Forest	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	76.615046	27.697832
21	Degraded Land	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	76.615669	27.697832
22	Degraded Land	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	76.616291	27.697832
23	Degraded Land	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	Degraded land	76.616914	27.697832
24	Degraded Land	Degraded land	Forest	Degraded land	Forest	Forest	Forest	76.617536	27.697832

Fig 4
Output Attribute Table

The accuracy in result depends entirely on the input data. While the classification is done care should be taken that the data is not getting distorted. In case distorted images are provided as input the result will have great percentage of inaccuracy since the whole process is analyzing each pixel individually. Raster images are usually heavy in terms of the enormous amount of pixel data contained in them. From this point of view when large datasets are provided as inputs to the algorithm used in this work, it will in turn increase the processing time taken to complete the process and give the results of change detection. One possible solution for this is to clip the image with respect to the desired area of interest and then use the algorithm, providing the same as input.

Conclusion and Future Scope

Spatial data mining is a promising field of research with wide applications in GIS. The work done in this paper is mainly useful for analyzing the land use land cover changes over a period of 7 years in Alwar District. The same concept and algorithm can be used for any other thematic data. The pixel level change detection done in this work can be used for further study and analysis in a variety of fields like Natural Resource Management, Monitoring and Planning, Climate change analysis, Suitability mapping, Environment, Forestry, Hydrology, Agriculture and so on.

References:

1. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part II Commission VI, WG VI/4
2. Spatial Data Miining: Progress and Challenges,Krzysztof,Junas. Adhikary,Jiawei Han
   School of Computing Science, Simon Fraser University
3. Data analysis Using GIS and Data Mining, Fang-Yie Leu, Tai-Shiang WANG