Analysis of Diverse Land-Cover Classification Techniques

S.Vidhya[1], K.Anuja[2], S.Gayathri Devi[3], R.Keerthana[4], C.Neethu[5]

Department of Information Technology
K.S.Rangasamy College of Technology
Namakkal (Dt), Tamil Nadu, India.
vidhyapsubramani@gmail.com

Abstract

From the past epoch, there have been great hard work and researches are undertaken for mounting land-cover classification in data mining. Whereas the essential uses of the remotely-sensed images is for mainly land-cover classification, for the purpose of applications like monitoring the forest and agricultural areas. Through remote sensing technology an extensive diversity of digital imagery is used to cover the majority of earth's surface. Researches and studies on information extraction are still comparatively frail. Notwithstanding the significant possible established by techniques for land-cover classification, enhancements are made obligatory for attaining the acceptable performance. For getting the most out of benefits of such data, automatic and well-organized classification approaches are considered necessarily. Generally remote sensing image classification can be seen as a joint venture together for techniques like image processing and classification. In this survey diverse collection of land-cover image classification techniques are illustrated along with their prominent features.

Keywords— Classification, Land Cover Classification, Remote Sensing, Spatial Data mining, Association, Visualization, Image Mining.

I. INTRODUCTION

The learn of land use is an significant not only in agriculturally dominated where over populated developing regions then though out the world as its relationship along with dissimilar human phenomena. It has been noted that the remote sensing technique is the most effective scientific tool in conjunction with ground truth plus topo sheet for collection of spatial information and useful in classification, mapping of the land use units and identification [1]. The increasing obtain ability of remote sensing images, attained periodically through satellite sensors on the same geographical area create it extremely interesting to progress the monitoring systems capable of producing plus regularly updating image classification techniques [2]. Researches on image classification depend on remote sensing have attracted the interest of the remote sensing community as environmental plus socio-economic applications are depend on the classification outcomes [3]. Evaluation of the changes of image classification using digital analysis of remote sensing satellite data can help decision makers to enlarge efficient plans for the management of land [4]. The progression in technology of remote sensing has caused it to turn into commonly used techniques in the world.

Image classification is one of the main parts of the image analysis, remote sensing plus pattern recognition. The classification might be the object of the analysis. For instance, classification of land use as of remotely sensed data creates a map similar to image as the final product of the analysis. The image classification creates an important tool for classification of the digital images. The word classifier refers loosely to a computer program that implements a specific process for image classification. There are various image classification events used for various purposes by researchers [5][6][7][8][9][10].

The main aim of this learn is to evaluate the accuracy of every classification method to categorize the land utilize by three kinds of classification specifically, Maximum Likelihood classification, Mahalanobis, and Minimum Distance classification. The main aim of image classification will be analysed here. To continue agreement between sustainable resources as well as socio-economic needs where image and land uses studies should be deal with care.

II. DIFFERENT LAND COVER CLASSIFICATION TECHNIQUES

The following sections will present the some diverse collection of Land cover Classification Techniques in data mining. And also for each technique its salient features are elucidated.

A. An Experimental Survey on Image Mining Tools, Techniques and Application

The evolution in the image attainment and the storage technology directed to unbelievable growth for the detailed image database and the study of these images can render much useful information [11]. The image mining deals to taken out the image data relationship, the embedded knowledge, or the other pattern which are not clearly found in the images [12]. The extension of data mining to the image domain is the image mining. An efficient information-driven framework contains four levels of information [13]. There are pixel level, object

level, semantic concept level, and knowledge and pattern level. A tool language called as iArm, which extract mining association rules from images [14]. By using iArm, it is possible to dig out association rules by writing a simple, easy and understandable coding can be written and maintained by end users who require no programming knowledge iArm can be modified according to the needs of end-user and can dig out rules using both textual and signal feature in which extracted rules signifies the hidden knowledge enclosed in images and hidden relationship present in the set of images. The extracted data are used for the classification and clustering of images. A fresh system to mine is visual knowledge on the web, which establishes an image detection technique, web image, collection tasks and the SVM classifier [15].

World Wide Web provides access to every in the world who are ravenous for information and the World Wide Web maintains a storehouse for very large amount of data. A project that intend at generic image classification using images which are automatically gathered from the web as the training images instead of hand-made image collections [16].

A generic image classification system with an automatic knowledge is to gain mechanism from the web. It consists of two steps for processing. First step is the gathering stage, which collect images according to the keywords given from the web automatically. Second step is the learning stage in which image feature is taken out from the gathered images and link them with each classes.

A framework which provide the possibility of use 5 distance function for estimation of similarities among images and two type of quantization [17]. A plan which implements a recursive HSV-space segmentation technique is to find out perceptually major colour areas [18]. The typical colour vector of these extracted areas is used to build the images an index which requires very little storage.

An Image mining techniques which is based on the colour histogram, texture of those images [19].

Merits:

- It draws upon proficiency in computer for vision, digital image processing,
- It Requires very little storage
- It Decrease complexity
- It Decrease dimensionality

B. Comparison of Different Image Classification for Land Use or Land Cover Classification

Land Use or Land Cover organization is a data that is sensed remotely and is used in researches and in many remote sensing applications [20]. The concept of land usage is really an important phenomenon to be considered not only used in agriculture arrear but also in developing parts in the world which is becoming over populated. The technique of remote sensing is useful in categorization identification and also mapping the used land. These images that remote sensed is used to have a development in system monitoring and also automatic production and preservation of land use maps. The researchers based on this has attracted on because the socioeconomic and environmental applications are related on this result. Classification of image places a vital role in remote sensing, pattern recognition and image analysis. It is also an essential tool in digital image classification.

Super wised categorization based on multispectral remote sensed imagery is frequently worn in land use determination. It consists of ultimate logic in all the software which is processing images. To add up there is a sequence of operation that is the crucial to be executed irrespective of the classifiers used. The two operations used are defining sites for training signature extractions. This classifier is used to give a classification image of the area to be studied for the purpose of land usage classification the maximum probability algorithm which is used by the classification becomes on Bayesian property. The lowest classification of distance is related to data on training set. ERDAS software for processing image is commonly used software for all analysis related to the processing of image. Six categories of land use that are considered in the training are water bodies, habitation, open/follow scrub, forest and agriculture. It includes three common steps

Training stage – few training areas and sites are constructed with polygon

Classification stage-images are classified using Mahalanobis classifiers, maximum probability and minimum distance.

Hydra spectral image classification HSI is based on observation which relates that signature in spectral form of pixel can also be given by sparse. By giving representative working set of the classes the class label is calculated.

The joint model of sparsity is developed for the upcoming spatio notion. In this pixel signature is local spatial neighbourhood. The sparsity pattern for HSI is an extension to this representation.

ROSIS (Reflective optics system imaging spectrometer) acquires images of urban from Pavia University.

HSI gaining more popularity due to linear modernization model which implement sparsity. The result obtained is proved for its robustness.

Merits:

- It is used for Agricultural purpose
- Promotes natural resource development
- Improves future development process

C. Land cover classification in remote sensing images using structured neural networks

The Land cover classification using the Network behaviour process is determined here. Based on the related band processing and sensory image classification, the classification process is done in which the trial and the error procedure is removed [21]. First the simplified network process is done and then the sensory images are determined by the band process. Bayesian classifiers have the better approach than other. The network behaviour and the definition of the architecture are defined with the application of multi sensor images in terms of land cover classification. This classification of Land cover using sensor images are used for monitoring the forest and for agricultural activities [22]. The new approach like Artificial Neural Networks is applied to classify the multisensory images. The advantages of this approach include, fault tolerance, information about the statistical data is not to be needed in prior, and classification time is faster.

In the tree like layered structure we have over two types as layer which is fully connected and layer which is well-separated. In fully-connected layer, all the neuron is connected successively to every other neuron in the layer. In well-separated layer, the output of the previous layer is interrupted to the next layer as their input [23]. The operation is based on the tree-like-networks (TLNs) criteria. The process takes place by the decomposition of neuron layers, global criteria is the first stage of decomposition level which are then checked on to the next level by sensor related sub-criteria followed by elementary sub-criteria

The contributions of neurons need to be simplified as, the neurons which enters and leaves the output with some definite weights. It may be positive or negative weights, those weights along with other data obtained are to be interpreted. Four steps are included 1. A generic TLN is transformed into the positive weights (TLN+) of equivalent one. 2: The (TLN+) are transformed to the unit of equivalent output network. 3: in which units of each connection layer is normalized by a multiplicative factor. 4: To simplify the process using the highest and the lowest values as linear function. The interpretation process provides the information by coding to classify the data. To validate the data and bring over the effectiveness in the land cover classification many experiments are done on the application of Agricultural and Forest Fields.

The data set of agricultural application consists of images in the multi sensor remote sensing field .From the images taken certain images of definite pixel range is selected for classification purpose. The classes selected are carrots, bare soil, stubble, sugar beet, potatoes. Winner-takes-all process determines the TLN obtained from the five classes. Bayesian classifier provides the result of the classification process. The features are extracted by the bands C, P and L with HV, HH, VV parameters. Based on the voting threshold and voting power bands that are related together are determined.

In this application along with the Fields of agricultural ATM bands, it includes the geometrical features. The process is done by k-fold cross validation process, by computing the combination of mean values for four subsets. For geometric feature extraction virtual sensors are used and for the Network behaviour PLTN is interpreted. The tabulated data are based on the significance parameters of the ATM and SAR sensor. Data fusion is the technique used to integrate over the data that are calculated all from the major multisensory images and are estimated over the land cover classification process.

Merits:

- It enhances the land-cover classification by structural neural network to great extent
- This technique eliminates over many problems and brings the classification process in the efficient way.
- It is mainly used for agricultural and forest fire applications using multi-sensor data

D. Spatial Data Mining

Data mining is the process to extract over the information from the large databases and having over the definite pattern recognition to get the statistical and visualization data [24]. Researches about on the Knowledge Discovery Data are done but it has high variations from spatial processing of data. The process of determining over the beneficial information from the previously hidden database is known as the spatial data. The characterizations, generalization, visualization, detecting non-spatial and spatial rules are the techniques involved spatial data mining.

Data of rich types, relationship with spatial variables, factors that are dependent, autocorrelation among the features of the spatial data are the certain features which are the same basic features of general data mining process. The ecosystem compositions, patterns of landscape structure are determined by the forest fires. The forest fire detection process is considered were it cause over the huge destruction to the soil, plants and even to humans. The data collected are in the form of the spatial data process. From that we determine over the process of artificial intelligence, image processing and spatial data mining process.

Drought, forest fires and other phenomena are determined by the spatial data which can be obtained by the remote sensors. Spatial data mining techniques includes segmentation, partitioning the data items, characterization and visual interpretation. The attributes which we take into account is the feature which makes the difference about relational data mining and spatial data mining. The inputs of the spatial process are the conversion of the digital images to YCBCR in which they are segmented and determine the values by the fuzzy logics.

The approach of false alarm is to false hotspot of fire system which is identified by the spatial data processing [25]. The graph based algorithm for the forest fire system by spatial outlier statistics are determined [26]. The GIS and the weather detection and many other data mining techniques, the forest fire system is determined [27]. The animals as their mobile sensor for forest fire detection [28]. The statistics from the system relay from which the data's are obtained from the satellite images and based on the coefficients the process is determined [29].

Forest ecosystem brings the influence over the economic and ecological utilities. Artificial intelligence and Image processing are the techniques used to determine over the forest fire spatial data system. The steps involved in this are conversion of colour space, specification and visualization. The transformation from one zone to another zone of modification is the colour space transformation. The Ycbcr colour format is obtained from RGB colour by the colour space transformation.

Based on the identical regions the image segmentation process takes by which the diffusion process of anisotropic process is used. After the (Ycbcr) converted image, the segmentation process is performed.

$$\partial/\partial t = u\,(x,t)\ \mathrm{div}\ (c\,(x,t)\ u\,(x,t))$$

It defines over the diffusion process and they are controlled by the factors of the equation defined and the spatial coordinator is denoted by the x.

Fuzzy logic was introduced by A.Zadeh in the University of California at 1965. After the segmentation process the Fuzzy logics are applied. It is the knowledgeable control achieving tool. The Fuzzy logic includes Fuzzy rules and Fuzzy sets. The membership functions are about the functioning of classifying the sets based on their belonging. The Fuzzy sets determine the objects behaviour.

The experiments are conducted by which taking over the spatial forest area without fire and with fire system and based on the Fuzzy logic sets the logics applied on the data where the fire detection system is detected separately The Forest fire system of spatial data mining is the best technique which provides the significant rule in the environmental damage and in human lives. Fuzzy logic and diffusion process involves in image segmentation and image processing.

Merits:

- Spatial Data mining brings a drastic change in the network connectivity, acquisition of data, mass data storage and in many technical assessments.
- It reduces the errors and improves the reliability process.
- Spatial data mining improves efficiently in the system of forest fire detecting system
- It deals with large amount of datasets and brings efficient solutions.

E. A comparison of the correlation structure in GPR images of deltaic and barrier-spit depositional environments

A general point in many practical problems in the earth science is to create an precise model of the spatial variation in physical and chemical properties of the sub surfaces [30].It may also entail the addition of different types of data such as surface geophysics, borehole measurements and direct sampling. A general approach to use geo statistical process is to report for the spatial variability that survives in heterogeneous, natural and geological environment.

Geo statistics supplies a mathematical representation which is mainly used to describe how some assets or stricture differs spatially. These spatial data can also be used in the progress of the subsurface model. Some of the information about spatial data can also be attained from borehole data or cores, but this typically provides a good measurement of variation in the vertical direction. To attain the characteristics of lateral variation, data can be gathered from "analog" projection. An alternate loom is the use of ground penetrating radar (GPR), which is a high resolution technique mainly used to image the trivial sub surfaces in two or three dimensions.

Several publications have detail described the usage of 2-D and 3-D GPR images mainly used to attain information about the litho logic variation and internal arrangements of sedimentary systems. We have selected a number of published radar data sets from various deltaic and barrier-spit depositional environments. The GPR sections are barrier-spit sites. All radar datasets were gathered using a pulse EKKO IV radar system with 100MHZ antennas [31],[32]. Developing of GPR data was performed using the software accessible from sensors.GPR data were first practiced to correct for signal saturation using a residual mean filtering; Common Midpoint Survey (CPM) was performed at each site to establish a subsurface velocity which was used to modify the gathered time segments to depth segments.

We use the semi vario gram, a plan which demonstrates the path in which the difference between data rate is associated with the distance between the data rates. An essential statement in semi vario gram analysis is that the data under testing conform to a theory referred to as “stationarity”. Stationarity means that any subset of data will have the similar statistical explanation as any other subset. We considered some of the Deltaic depositional environments like Box Elder Creek, William River, Athabasca River and American Fork River. Some of the Barrier spits depositional environments like Sandy Point, Willapa Bay

In radar data sets, two different stationarity problems must first be concentrated. The first issue has to do with the signal of radar itself. The radar signal in all the suggestions in the GPR profile may attenuate with the depth. This will forms a violate stationarity. To beat this issue, we use an Automatic Gain Control (AGC) with a short window length to access all amplitudes of reflections which is mainly used to approximate the same value. The second issue is to do with the facies variations found in the sedimentary series due to differing depositional processes. Semi variogram models are used to implement an analytic characteristic of the experimental semi variogram. We considered a number of more commonly used models such as Exponential, Spherical, Cubical, Penta spherical, Gaussian, Power models and also the nugget effects.

The decision as to which model suit best fits the data was made by considering two parameters: The sum of the squares of the weighted differences, R, and the Akaike Information Criterion (AIC).

The AIC is given by

$$AIC=n \ln (R/n)+2p$$

Where n is the number of points in the experimental semi variogram and p is the number of parameters in the model. When comparing models for any one data set, the model with the smallest AIC is the best model.

It is found that in all the cases, that the geo statistical analysis of the GPR section created semi variograms of excellent quality- both in terms of the formation of variograms and also in the terms of the agreement with the selected model. The semi variogram analysis consists of certain parameters and results such as area analysis, orientation of lag vector, range, exponential model, nugget effect, R and AIC used for concluding the best model.

Merits:

- It generates a precise model of the spatial variations both in the physical and chemical properties of the subsurface.
- GPR forms semi variograms of best quality.
- It allows us to contact the specific planes in various depositional environments.

III. CONCLUSIONS

Quite a lot of land-cover classification techniques were elucidated along with their features of several tasks in this paper. Each and every sub tasks of land-cover classification techniques tends to be extremely necessary reinforcement process for proficient information extraction. This paves the way for the development of many land-cover classification techniques. This paper provides a thought that most researchers and scientists use the above techniques for classifying the land-cover areas and also they attain substantial results. Thus the main role of this survey is to enrich knowledge about the land-cover classification techniques and its features in several fields which will be very helpful for the readers and also it meets the needs of land-cover classification researchers to innovate more advanced techniques in future.

REFERENCES

[1] C Nageshwara R., and Vaidyanadhan, R., (1981) “LandUse Capability Studies from Aerial Photo- Interpretation-A Case Study from Krishna Delta, India” Geographical Review of India, 3, pp.226-236.

[2] Bruzzone, L., Cossu, R. and Vernazza, G. (2002) “Combining parametric and non-parametric algorithms for a partially unsupervised classification of multitemporal remote sensing images” Information Fusion 3,289-297

[3] Lu, D., Weng, Q. (2007) “A survey of image classification methods and techniques for improving classification performance”, International Journal of Remote Sensing, Vol. 28, No. 5, 823-870.

[4] Gordon, S., (1980) “Utilizing Landsat imagery to monitor land use change: A case study in Ohio” Remote Sensing of Environment, 9, 189-196.

[5] Butera,M.K., (1983) ”Remote sensing of wetlands” IEEE Transactions on Geoscience and Remote Sensing GE21, pp.383-392.

[6] Ernst, C.L. and Hoffer, R.M., (1979) "Digital processing of remotely sensed data for mapping wetland communities. LARS Technical Report 122079" Laboratory for Applications of Remote Sensing. Purdue University, West Lafayette, 119.
[7] Lo, C.P. and Watson, L.J., (1998) "The influence of geographic sampling methods on vegetation map accuracy evaluation in a swampy environment" Photogrammetric Engineering and Remote Sensing, 64,11891200.
[8] Ozesmi, S.L and Bauer, M., (2002) "Satellite Remote Sensing of Wetlands" Wetlands Ecology and Management, 10, 381-402.
[9] Dean, A.M and Smith, G.M. (2003) "An evaluation of per – parcel land cover mapping using maximum likelihood class probabilities" International Journal of Remote Sensing, 24 (14), 2905-2920
[10] Liu, X.H., Skidmore,A.K. and Oosten, V.H. (2002) "Integration of classification methods for improvement of land-cover map accuracy" ISPRS Journal of Photogrammetry & Remote Sensing, 56, 257-268.
[11] C. Lakshmi Devasena et al. / International Journal on Computer Science and Engineering "An Experiential Survey on Image Mining Tools,Techniques and Applications" Vol. 3 No. 3 Mar 2011
[12] Ji Zhang Wynne Hsu Mong Li Lee " Image Mining: Issues, Frameworks and Techniques"
[13] Ji Zhang, Wynne Hsu and Mong Li Lee "An Information-Driven Framework for Image Mining" Database and Expert Systems Applications in Computer Science, 2001, Volume 2113/2001, 232-242, DOI: 10.1007/3-540-44759-8_24
[14] Hassan H. Malik "iARM: Image Association Rule Mining Language" 2005.
[15] Bingbing Ni, Zheng Song, Shuicheng Yan "Web Image Mining Towards Universal Age Estimator"
[16] Keiji Yanai "Web Image Mining: Can We Gather Visual Knowledge for Image Recognition from the Web?"
[17] Aura Conci., Everest Mathias M. M. Castro "Image Mining By Color Content"
[18] D. Androutsos, K. N. Plataniotis and A. N. Venetsanopoulos "A Novel Vector-Based Approach to Color Image Retrieval Using a Vector Angular-Based Distance Measure" Computer Vision and Image Understanding Vol. 75, Nos. 1/2, July/August, pp. 46–58, 1999
[19] Rajshree S. Dubey "Image Mining using Content Based Image Retrieval System" (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 07, 2010, 2353-2356
[20] Subhash Thakur, Akhilesh Singh, Seema Suraiya " Comparison of Different Image Classification Techniques for Land Use Land Cover Classification: An Application in Jabalpur District of Central India" International Journal of Remote Sensing and GIS, Volume 1, Issue 1, 2012, 26-31
[21] F.Roli,S.P.Serpico "Land Cover Classification in Remote-Sensing Images using Structured Neural Networks"Vol.4,No.3-XII,1995
[22] Bischof H.,Schneider W.,& Pinz A.J.,"Multispectral classification of Landset-Images using neural networks", IEEE Trans. On Geoscience and remote Sensing, Vol. 30, No. 3, May 1992, pp.482-490
[23] Hertz J.,Krogh A.& Palmer R.G., Introduction to the theory of Neural Computation, Addison Wesley Pub. Co., The Advance Book Program, 1991
[24] Seng Chuan TAY, Wynne HSU, Kim Hwa LIM, "Spatial Data Mining: Clustering of Hot Spots and Pattern Recognition"Proceedings. 2003 IEEE International Geoscience and Remote Sensing Symposium, Volume: 6,pp: 3685- 3687, 21-25 July 2003.
[25] Young Gi Byun, Yong Huh, Kiyun Yu, Yong Il Kim, "Evaluation of Graph-based Analysis for Forest Fire Detections", Proceedings of world academy of science, engineering and technology, volume. 10, December 2005, ISSN 1307-6884
[26] Daniela Stojanova, Panče Panov, Andrej Kobler, Sašo Džeroski, Katerina Taškova, "Learning to predict forest fires with different Data Mining techniques", Conference on Data Mining and Data Warehouses (SiKDD 2006),Ljubljana, Slovenia, pp. 255-258, October 9, 2006.
[27] Yasar Guneri Sahin, "Animals as Mobile Biological Sensors for Forest Fire Detection", Sensors, vol. 7, pp. 3084-3099,2007.
[28] González, J.R., Palahí, M., Trasobares, A., Pukkala, T., "A fire probability model for forest stands in Catalonia (northeast Spain)," Annals of Forest Science, 63: 169-176, 2006.
[29] F.Roli,S.P.Serpico "Land Cover Classification in Remote-Sensing Images using Structured Neural Networks"Vol.4,No.3-XII,1995
[30] Jol, H. M., Meyers, P. A., Lawton, D. and Smith, D., 1994, A detailedGPR investigation of a coastal barrier spit, Long Beach, Washington,U. S. A.: SAGEEP, Ann. Mtg., Env. Eng. Geophys. Soc., 107–127.
[31] Jol, H.M., and Smith, D.G.,1992, Geometry and structure of deltas in large lakes:Agroundpenetrating radar overview: Geol. Surv. Finland, Spec. Pap. 16, 159–168.
[32] Smith , D.G., and Jol, H.M., 1992b, GPR results used to infer depositional processes ofcoastal spits in large lakes: Geol. Surv. Finland, Spec. Pap. 16, 169–177.