THE APPLICATION OF NEURAL NETWORK DATA MINING ALGORITHM INTO MIXED PIXEL CLASSIFICATION IN GEOGRAPHIC INFORMATION SYSTEM ENVIRONMENT

NANNA SURYANA HERMAN

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

The Application of Neural Network Data Mining Algorithm into Mixed Pixel Classification in Geographic Information System Environment

Norhayati Mohd Rasip
Faculty of Information Technology and Communication
Technical University of Malaysia Melaka
yati_mr@yahoo.com

Nanna Suryana Herman
Faculty of Information Technology and Communication
Technical University of Malaysia Melaka
nsuryana@utem.edu.my

Abstract

With the rapid growth of satellite technology and the increasing of spatial resolution, hyperspectral imaging sensor is frequently used for research and development as well as in some semi-operational scenarios. The hyperspectral image also offers unique applications such as terrain delimitations, object detection, material identification, and atmospheric characterization. However, hyperspectral image systems produce large data sets that are not easily interpretable by visual analysis and therefore require automated processing algorithm. The challenging of pattern recognition associated with hyperspectral images is very complex processing due to the presence of considerable number of mixed pixels. This paper discusses the development of data mining and pattern recognition algorithm to handle the complexity of hyperspectral remote sensing images in Geographical Information Systems environment. Region growing segmentation and radial basis function algorithms are considered a powerful tool to minimize the mixed pixel classification error.

Keywords

Data mining, Hyperspectral images, Mixel problem, Image Segmentation, Neural network, Image Classification.

I. INTRODUCTION

With the development of remote sensing technology, the information on earth surface has been recorded remotely and transformed into remote sensing images. Reflectance from each pixel is measured at many narrow, contiguous wavelength intervals. However the increasing of spatial resolution associated with most satellite sensor systems imaging is higher than the normal sizes of objects on the ground.

Therefore one pixel in the satellite image possibly covers more than one object. Although the detailed information is recorded, yet there have a serious problem to recognize a particular geographical object due to overlapping of two or more associated spectral properties.

Data mining technique has been applied into the process of searching for knowledge or major interest of information within a large spatial database. The process searches for pattern recognition and representation have been applied in many areas such as in marketing, telecommunication, finance and others.

However, this technique is rarely applied into extracting information from large remote sensing image as typical data inputs into spatial databases associated with Geographical Information Systems (GIS) [7]. It is also stated that data mining techniques is potentially applicable in the domain of remote sensing images from which large quantities of data are gathered. More over, the data mining technique is an interactive and iterative process involving data preprocessing, pattern recognition and interpretation. In this context classification is recognized in GIS environment as a common technique to extract information from remote sensing image data. It converts measurements of earth surface characteristics into thematic maps that suit predefined user's requirements.

This paper discusses the development of neural network data mining algorithm to handle the complexity of pattern recognition extraction associated with hyperspectral remote sensing images and applied within Geographical Information Systems environment. Therefore the incorporation of the proposed algorithm into GIS increases the usefulness of GIS.

This paper discuss also on how the hyperspectral image is segmented into several homogeneous regions to derive lass characteristics needed for mixed pixel decomposition which allow extracting the spectral and spatial features associated within each homogeneous region. The training scheme is used to classify the object using the supervised classification techniques.

II. RELATED WORKS

This part of the paper is intended to discuss the state of the art of hyperspectral image processing techniques which at the end able to provide information on the need of the proposed algorithm. In other words, there is many techniques have been studied but the decomposition of mixed pixel problem using data mining technique still have need more attention for both the ongoing and future research.

The result of data mining and hyperspectral remote sensing information processing analysis allows the development of intelligent information processing [17]. This implies there is a need to investigate multi-discipline theories and technique of hyperspectral remote sensing data mining such as image processing, spatial database organization, pattern recognition, information retrieval and data mining [12]. Related to the information retrieval, M. Datcu et al. applied semantic and visual

317

content to retrieve relevant images based on user requirements. [13] [14] On the state-of- the art data mining extract the useful pattern that are previously unknown from large collection of images. Many researchers has been applying the supervised classification using supportive vector machines to identify spectral characteristic and texture features characterizing spatial information. These characteristics and features are possible to be identified using Gabor wavelet [7]. On the image processing, generally focuses on extracting and understanding features within a single image. However the advent of high resolution satellite imagery, context-based algorithm, pixel based and per-parcel image processing is more practically applied to remote sensing images [19].

Applied to satellite image classification implies that each pixel should be labeled as belonging to a class as defined by users e.g., land use of land-cover classes. The measurements are available as feature vectors, and usually concern reflection in different spectral bands. A reflectance is a local earth surface characteristic, and be considered a class property [3]. Generally, a class cannot be associated with the single reflectance. The reflection measurements, however, do not represent an exact square or a point, but rather a weighted average in a circular in area called instantaneous field of view (IFOV). In this regards, mixed pixels pose a problem which is the errors that are caused by spectral overlap.

Recently, many researchers in the remote sensing keep improving image classifier because the mixed pixel is not influenced by the precision of object recognition and classification but to become an obstacle to the quantification analysis in remote sensing images [2]. The mixed pixels leads to errors that make the subsequent area estimation inaccurate because of mixed pixel is assigned to one class only, while in fact it belongs to two or more classes. Figure 1 showed that spectral confusion only adds to this problem as many mixed pixels composed of X and Y will be allocated to class Z , but the area of pure pixel belonging to Z will never be divided between classes X and Y Several techniques such as linear spectral unmixing, mathematical morphology theory, hypothesis testing Hough transform, neural network and maximum-likelihood have been applied to overcome this problem. A general approach to mixed pixel classification is linear spectral unmixing, which uses a linear mixture model to estimate the abundance fractions of signatures within a mixed pixel. C. Chang and H. Ren compared some popular classification algorithm which is spectral unmixing, the orthogonal subspace projection (OSP), maximum likelihood, minimum distance, and Fisher's linear discriminant analysis (LDA). They converted the mixed pixel to the homogenoeus ones by designed mixed-to-pure pixel converter.
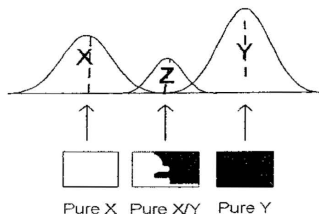

Fig 1. Spectral confusion caused by mixing of ground cover types

Y. Liu et al. 2001 have studied about classification algorithm development of understanding system for remote sensing images. They have been considered the spectral features in different regions. It was expected that the spectral features for one object are similar and the spectral features are different for different objects. However in a commonly cases, the spectral features for the same object are often difference that are caused of shadow of mountains, the difference for image acquiring time and etc., in the other hand , the spectral features for different objects are the same in sometimes because of the image resolution [3], [20]. L. Soh and C. Tsatsouli, 1998 presented a methodology that integrates traditional image processing algorithm with technique adapted from knowledge discovery in databases (KDD) and data mining to generate a system that successfully segments natural scenes image. They used segmentation technique which serves as a preprocessor to classification [10].

On the neural network, the most commonly used family of neural network for pattern classification task is the feed-forward network, which includes multilayer perception and radial basis function (RBF) network [10]. Another popular network is the self organizing map (SOM) [1] or Kohonen Network [8] for data clustering and feature mapping. M.A Ibrahim et al. and as stated also by [19] both have applied the mixed pixel in all the stages of the supervised neural network and applied the modeling of the previous knowledge in the form of more weight to the abundant classes in the region in order to improve the classification accuracy of back propagation neural network algorithm.

III. PROPOSED SEGMENTATION AND CLASSIFICATION ALGORITHM

Recently data mining technique using neural network have been applied in difference research area. A comparison of neural networks, fuzzy c-means and the analytic linear mixture modeling approach has been studied by Atkinson et al [16] who have mentioned that neural networks gave the most accurate estimates but often takes a very long time to establish an acceptable training data.

Hence, image segmentation has growing tremendous importance to extract spatial and spectral information from the remote sensing images. Huang et al [5] applied segmentation scale to generate image objects for object-oriented classification and the result is more accurate accuracy than those of based-pixel approach. D. Ming et al [3] also stated that image segmentation for information extraction or target recognition based on parcel unit extraction and the experiment results showed that the extract simple target quickly an rightly. M. P. S. Silva et al [15] built a reference data set by using segmentation algorithms to partition the image into regions which are spatially continuous, disjoint and homogenous. So it evidently that image segmentation is the ground work.

Taking above explanation on the advantage of image segmentation techniques therefore the proposed algorithm has adopted both segmentation and classification techniques. The proposed algorithm is illustrated in Figure 1.
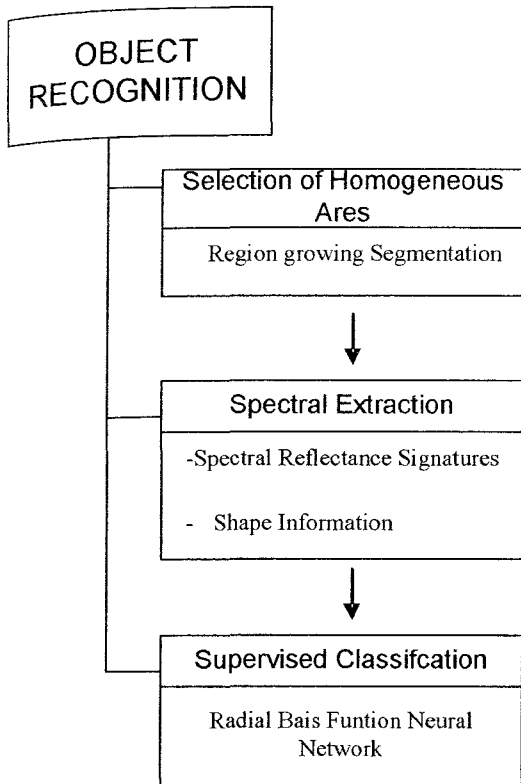
318

Fig2. Proposed hyperspectral image algorithms of pre-processing, feature extraction and classification.

The Figure 1 shows that the first step is image segmentation that will be applied to separate regions from the hyperspectral image in order to recognize and regionalized them as homogeneous objects. The second step is the obtaining spectral reflectance from each region measurement. Finally, the Radial Basis Function (RBF) Neural Network is used to classify object from overlapping class properties.

A. Selection of Homogeneous Areas

Region growing is an approach to segmentation in which neighboring pixels are examined and added to a region class if no edges are detected. This process is iterated for each boundary pixel in the region. If adjacent are found, a region merging algorithm is used in which weak edges are dissolved and strong edges are left in tact.

A region growing algorithm has been developed to sample the training and test data. The motivation for using region growing to sample test and training examples is to get a dataset covering a relatively homogeneous area and meanwhile smooth the data.
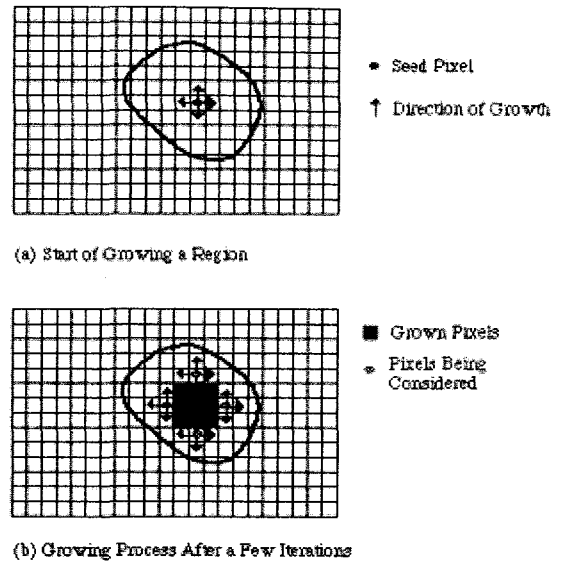

(a) Start of Growing a Region

(b) Growing Process After a Few Iterations

Fig3 . Example of region growing

We adopted the region growing segmentation algorithm developed by INPE's [9] and granted in the SPRING software system [4], which is freely available on the internet. The two parameters: a similarity an area threshold values the way SPRING's region growing works can be explained as follows [9]
:

1. The image is first segmented into atomic cells of one or few pixels
2. Each segment is compared with its neighbors to determine if they are similar or not
3. The segment continues growing by comparing it with all the neighbour until there is no remaining joinable region, at which point the segment is labeled as a region
4. The process moves to the next uncompleted cell, repeating the entire sequence until all cells are labeled.

B. Spectral Extraction

The order of magnitude higher spectral resolution of hyperspectral imagery has been compared to multispectral imagery which creates opportunities for new feature-extraction technique such as image residuals, spectral fingerprints, absorption-band parameters, spectral derivative ration.

a. Spectral reflectance signature

The spatial resolution of satellite remote-sensing systems is too low to identify many objects by their shape or spatial detail. Therefore, the measuring the spectral signatures of surface materials, such as vegetation, soil and water, over the spectral range in Figure 3 is used to identify and delineate geographical object.
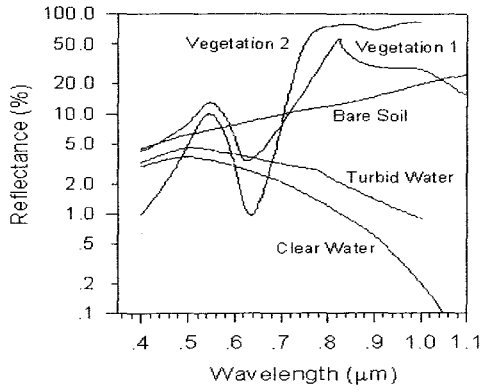
319

Fig4. Example of reflectance spectra of five materials: clear water, turbid water, bare soil and two types of vegetation.

B.Supervised Classification

B..1 Radial basis function network

Radial Basis Function (RBF) networks can be regarded as special Multi Layer Perceptron (MLP) structure in which the parametric statistical distribution model and non non-parametric linear perceptron algorithm are combined together in serial sequences. Radial Basis Function network based classifier also can generate complex decision regions using simple kernel functions, particularly, for remote sensing image classification and feature extraction [6]. Then, the learning rate of RBF network is very fast than back propagation network [18].

Radial basis Function Neural Network is a three layer feed forward network with a single layer of hidden units operating on some sort kernel functions and the output layer linearly combining the outputs of the hidden layer as illustrated in Figure 4. Rather than employing an inner product between the input vector and the weight vector, hidden neuron activation in RBF network are computed using the exponential of a distance measure (usually the Euclidean distance or a weighted norm) between the input vector and a prototype vector that characterizes the signal function at a hidden neuron.
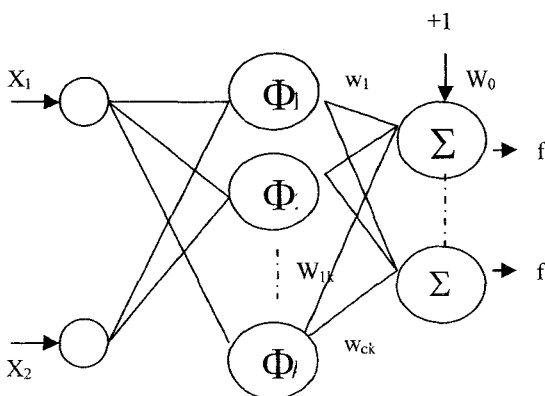
. Typical architecture of classifier based on RBF neural network

In order to implement the RBF neural network, we adopted the Least Mean Square learning algorithm to adjust parameters used in the hidden layer and weights of the output layers adaptively and simultaneously so as to archive possible best performance for the RBF neural network algorithm [18].

IV. CONCLUDING REMARKS

In this paper, the proposed segmented classification algorithm is proposed. It took the advantage of segmentation and radial basis function neural network classification to achieve better object recognition result.

Algorithm is considered potentially applicable to extract information on the geographical object recognition from hyperspectral images

V. ACKNOWLEDGEMENT

VI. REFERENCES

[1] Barsi,Neural, "Graph based Neural Self- Organization in Analyzing Remote Sensed Images", *IEEE International Geoscience and Remote Sensing*, 2003, pp. 2937- 2939.

[2] C. Chang and X. Zhao, "Least Square Subspace Projection Approach to Mixed Pixel Classification For hyperspectral Images", *IEEE Transaction on GeoScience and remote Sensing*, 1998, pp. 99-119.

[3] D. Ming et al, "Features Based Parcel Unit Extraction From High Resolution Image", *IEEE Transaction on Geoscience and Remote Sensing*, 2005, pp.1875-1878.

[4] G. Camara, "SPRING: Integration Remote Sensing and GIS with Object-Oriented Data Modelling. Computers and Graphics ",1996, pp. 13-22.

[5] H. Huiping et. al, "Analysis to the Relationship of Classification Accuracy, Segmentation Scale and Image Resolution", *IEEE International Geoscience and Remote Sensing* 2003, pp.3671- 3673.

[6] J. A. Benediktsson et.al, "Classification and Feature Extraction of AVIRIS data " *IEEE Trans. On Geosciemce and Remote Sensing*, 1995, pp. 1194-1205.

[7] J. Li and M. Narayanan, "Integrated Spectral and Spatial Information Mining in Remote Sensing Imagery", *IEEE Transaction on Geosciences and Remote Sensing*, vol. 42, 2004, pp.673-685.

[8] J. Razmara, "A Model for Classifying Multisource Remote Sensing Images by Kohonen Neural Networks", *IEEE International Geoscience and Remote Sensing Symposium, IGARSS'04*, 2004, pp. 3849 – 3852

[9] L. Bins et. al. "Satellite Imagery Segmentation : A Region Growing Approach", *In VIII brazilian Symposium on Remote Sensing.* 1996.

[10] L. Soh and C. Tsatsoulis, "Data Mining in Remotely Sensed Images: A General Model and an Application", *IEEE International Geoscience and Remote Sensing Symposium Proceedings*, 1998, pp.798 – 800.

[11] L.C. Bastos et al,. "Radial Basis Function for Clasification of Remote Sensing Images", *International Joint conference on Neural Network,* 1999, pp. 1959-1962,

[12] M. C. Burl et. al, "Mining for Image Content,"In Proc. Conf. Systemics,, Cybernatics, Informatics/Information Systems: Analysis Systhesis, Orlando. FL, 1999.

[13] M. Datcu et. al, "Information Mining in Remote Sensing Image Archieves" : System Concepts, , *IEEE Trans. Geoscience Remote Sensing*, 2003, pp.2923-2936..

[14] M.Datcu and K. Seidel, "Image Information Mining:Exploration of Image Content in Large Archives", *IEEE Aerospace Conference Proceeding,* 2004, pp. 2420 - 2423 ,

[15] P. S. Silva et. al, "Mining Patterns of Change in Remote Sensing Image Database" *Proceedings of Fifth IEEE International Conference on Data Mining* , 2005.

[16] P.M. Atkinson et al,. "Mapping Sub-pixel Proportional Land Cover with AVHRR" *Imager International Journal of Remote Sensing*, 1997, pp. 917-935.

[17] Peijun, C. Yunhao, "On the Framework, Algorithms and Applications of Hyperspectral Remote Sensing Data Mining", Proceedings. *2005 IEEE International Geoscience and Remote Sensing Symposium, IGARSS '05,* 22005,pp.752-755.,

[18] Q. Du and C.-I Chang, An Interference Rejection-Based Radial Basis Function Neural Network for Hyperspectral Image Classification, International Joint Conference on Neural Network, 1999, pp. 2698-2703.

[19] Suryana, N. The Use of Multiplying Effect of Ancillary Data in Erosion Severity Classification: A Method for Improving Data Quality of Multispectral Images

Classification. Proceeding European GIS 91, 2 [nd] European GIS Conferences ,1991, pp. 1046 – 1057.

[20] Y. Liu et al, "Study on Image-Segmented Classification", *IEEE International Conference on Info-tech and Info-net,* 2001, pp.296-301

VI. BIBLIOGRAPHY

NURHAYATI MOHD RASIP is a research assistant and currently pursuing her Master study by research at the Faculty of Information Technology and Communication (FTMK), UTeM under the supervision of the second author. Her research interest is in spatial temporal data mining.

NANNA SURYANA HERMAN

He is a Deputy Dean for Research and Postgraduate Programme and lecturer at Faculty of Information Technology and Communication (FTMK), Technical University of Malaysia Melaka (UTeM). He got his PhD in year 1996 from the Department of GIS and Remote Sensing, Wageningen Research University, Holland. He published more than 60 papers in journals and conference proceeding and three books. His research interest is in the field of GIS interoperability infrastructure, Location Based Service, Spatial Modelling, Spatial and Temporal Databases and Information Security.