A MULTI-RESOLUTION HIERARCHY CLASSIFICATION STUDY COMPARED WITH CONSERVATIVE METHODS

G. B. Zhu [a, 1] , X. L. Liu [b], Z. G. Jia [c]

[a] International School of Software , Wuhan University , 129 Luoyu Road , Wuhan 430079, China - gbzhu@iss.whu.edu.cn

[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China - liuxl.j@163.com

[c] Faculty of Land Resource Engineering, Kunming University of Science and Technology, 68 Xuefu Road, Kunming 650093, China - jia.zg@163.com

KEY WORDS: Remote Sensing, Image Classification, Multi-Resolution Hierarchy

ABSTRACT

Based on the principle of remote sensing image classification, the paper roundly discusses most popular conservative methods of supervised classification and unsupervised classification of remote sensing image, and simply appraises their advantages, disadvantages, and suitable occasions. The newest methods of computerized remote sensing image classification are discussed with the attempt to bring out the trend of remote sensing image classification. In the end, this paper proposes a new classification method based on the multi-resolution hierarchy in order to have new breakthrough for the research of the classification method of remote sensing image.

1. INTRODUCTION

Remote sensing platforms have collected a considerable amount of data about the nature of the Earth's surface. How to extract high precise information from remotely sensed data, however, is one of the key issues on research and application of remote sensing. Remote sensing image classification and analysis have become the context in which special information extraction, dynamic change forecast, thematic mapping and building remote sensing dataset will be implemented. With the development of computer, digital analysis technology of remote sensing image has been improved from original image visual interpretation, and computer aided classification to automated recognition. And remote sensing classification has gradually developed from single pixel physical feature recognition, spectral and textural information extraction to image comprehension (Zhu, 2003).

For many years, various image classification methods have been developed. In most of studies, conservative methods often depended on the statistical relationship between remote sensing data and training data (Bo and Wang, 2003), which have been employed as per-pixel classifiers on the assumption image pixels are pure (i.e., pixels contain one and only one class). In fact, remotely sensed images, particularly low-resolution images, are contaminated with mixed pixels that contain more than one class on the ground. These mixed pixels reflect the composite spectral response of the classes within them (Campbell, 2002). However, pixel-classifications may lead to inaccurate classification in cases of mixed pixels and fuzzy boundaries. Simultaneously, traditional classifications are always based on a certain single strategy (e.g. grey grades), which would lead to optimal accuracy only when given case is provided. Moreover, traditional classifications are often based on not overall imaging factors including spatial and spectral information, but purely spectral data. The problem is even worse in case of hyper-spectral image resulting in the "Hughes Phenomenon" without attention to spatial relationships between corresponding pixels (Hughes, 1968). Therefore, many soft classification methods are widely discussed in recent literature, which are more flexible contrary to pixel-classifications to tackle cases of mixed pixels and fuzzy boundaries. The outputs from soft classification, when scaled from 0 to 1, provide a set of fraction images that represent proportion of classes for each pixel (Xu et al., 2005).

With the tendencies of "3-High" (high spatial resolution, high spectral resolution and high frequency resolution) and "3-Multi" (multi-sensor, multi-flat and multi-angle) (Li, 2003), on the one hand, a better classification is possible in theory because satellite images offer more information; on the other hand, it has been common recognized that the discrimination between classes becomes more difficult because the number of different classes that can be detected increases (Hsieh et al., 2001). Furthermore, it is important for the thematic mapping to pay attention to scale-related issue which is the relationship between scale, class definition, and pixel resolution. The size of pixels in relation to the expected area of homogenous cover parcels is an important factor in determining the significance of this effect (Wilkinson, 2005). Hence, a good resolution is to turn to the best appropriate spatial resolution in relation to the given spatial object in theory (Chen and Zhao, 1989). Therefore, this paper presents a new classification method based on the multi-resolution hierarchy including several following steps. The first is to build a multi-resolution hierarchy dataset of remote sensing considering natural inherent relation between pixels and multi-resolution hierarchy structure of the spatial entities because there is best appropriate spatial resolution in relation to the given spatial object in

[1] Corresponding author. Tel.: +86-27-68754771; fax: +86-27-68754771.
E-mail address: gbzhu@iss.whu.edu.cn (G. Zhu)

theory (Chen and Zhao, 1989; Krawitz, 1974). The second is to employ respectively artificial neural network method and decision tree approach to determine relationships of different classes from homo-hierarchy and relationships of same classes from multi-hierarchy remotely sensed data so as to circumvent these problems by focusing on fewer classes to be identified. The third is practice decision tree approach to build the relations between a certain kinds of spatial objects from different hierarchy in order to extract semantic information. Extracted semantic information benefits the better classification of remote sensing image in return. Consequently, this method can be used for integration of remote sensing and GIS at semantic level.

2. TRADITIONAL CLASSIFICATION OF REMOTE SENSING IMAGE

Over the years, scientists have devised many image classifiers. There are two main divisions of classification: supervised classification and unsupervised classification based on whether prior information is required or not.

2.1 Supervised Classification

Supervised classification always classifies pixels of unknown identify by samples of known identity located within training areas. The analyst defines training areas by identifying regions on the image that can be clearly matched to areas of known identity on the image. Such areas should typify spectral properties of the categories they represent, and, of course, must be homogeneous in respect to the informational category to be classified. Clearly, traditional supervised algorithms are based on probability distribution models for the class of interest (Richards, 1986). The selection of these training data is a key step in supervised classification. Differences in the selection of training data were more important influences on accuracy than those among other four different classification procedures (Bo and Wang, 2003). Many methods have been devised to implement the basic strategy of supervised classification as follows: parallelepiped classification; K-nearest neighbour; minimum distance classification; maximum likelihood classification; Bayes's classification and so on.

2.2 Unsupervised Classification

Unsupervised classification can be defined as the identification of natural groups, or structures, within multi-spectral data. A typical sequence for unsupervised classification might be composed of two main stages. Firstly, the analyst specifies minimum and maximum numbers of categories by classification algorithm. The second is to find new centroids for each class, then the entire scene is classified again. Again new centroids are calculated; if the new centroids differ from those found in the preceding step, then the process repeats until there is no significant change detected in locations of class centroids and the classes meet all constraints required by the operator. Apparently, distance measures are the heart of unsupervised classification. Unsupervised classification is particularly useful when training data can not be obtained to perform supervised classification. These techniques are also used in exploratory analysis to determine the number of possible spectral classes that can be considered for a supervised classification process. Many procedures for unsupervised classification are available; despite their diversity, most are based on the above general strategy, such as K-Mean cluster method and ISODATA.

2.3 Comparison of Conventional Statistical Classifiers

The supervised method requires considerable interaction with the analyst who must guide the classification by identifying areas on the image that are known to belong to each category, but the unsupervised method proceeds with only minimal interaction with the analyst. Next, the supervised method can use these classified regions to obtain models that are used to classify the other parts of the image, the unsupervised method is based on the image itself, although some general assumptions about images can be used. In practice, unsupervised classifiers are often used in exploratory analysis to determine the number of possible spectral classes before a supervised classification process. The two strategies are not as clearly distinct as these definitions suggest, for some methods do not fit neatly into either category. These so-called hybrid classifiers share characteristics of both supervised and unsupervised methods. Table 1 shows the difference between supervised classification and unsupervised classification.

Table 1. Comparison of conventional statistical classifiers

Class	Method	(Non-)/ Parametric	Advantage	Disadvantage
Un-supervised	K-Mean Cluster	Non-parametric	Don't need prior knowledge; algorithm is easy and fast.	Accuracy is low and depends on the initial centroids.
	ISODATA	Non-parametric	Don't need prior knowledge; accuracy doesn't depend on cluster centres when iterations are enough.	Computation is complicated; accuracy is always lower than supervised methods.
Supervised	K-nearest neighbour	Non-parametric	Don't need prior knowledge and frequency distributions of spectral values; degree of the "Hughes Phenomenon" is lower.	Vast training samples needed result to large computational intensity; the output depends on data quality.
	Parallelepiped	Non-parametric	Don't need prior knowledge and frequency distributions; algorithm is easy and fast.	Define regions within a multi-dimension data space based on ranges of values within the training data.
	Minimum Distance	Non-parametric	Don't require class prior probability; don't need frequency of data distributions; algorithm is easy and fast.	Covariance matrix of classes doesn't consider; training samples is less.
	Mahalanobis Distance	Parametric	Don't require class prior probability; relative to minimum distance method, covariance matrix is considered here.	Data distributions must obey normal school; accuracy is lower than maximum likelihood method.

	Maximum Likelihood	Parametric	Introduce covariance matrices; when closing to a normal (Gaussian) distribution, accuracy is higher.	Computational intensity is larger; data must obey normal school; need enough training samples and prior probability.

3. TREND OF REMOTE SENSING IMAGE CLASSIFIACTIONAN

With the application of higher spatial resolution images, it is questionable whether traditional classification methods are still applicable for high-resolution imagery. In recent literature, many soft classification methods are widely discussed which are more flexible than per-pixel classifiers to tackle cases of mixed pixels and fuzzy boundaries. In soft models for image classification, non-probabilistic uncertainty due to vagueness and/or ambiguity should be modelled as partial belongingness to several categories at the same time, such as fuzzy set theory (Foody, 1996), artificial neural networks (Moody et al., 1996), linear mixture modelling (Settle and Drake, 1993), expert systems (Ghosh and Samanta, 2003), decision tree method (Xu et al., 2005) and texture classification (Zhang et al., 2005), and so on.

3.1 Fuzzy Set Theory

Most landscapes are so complicated that they are virtually incomprehensible without an analytical framework. The common solutions adopt a fuzzy set framework (Binaghi et al., 1999; Foody, 1996). The apparatus of fuzzy set theory serves as a natural framework for modelling the gradual transition from membership to non-membership in intrinsically vague classes. The fuzzy set framework introduces vagueness, with the aim of reducing complexity, by eliminating the sharp boundary dividing the members of a class from non-members. In some situations, these sharp boundaries may be arbitrary, or powerless, as they cannot capture the semantic flexibility inherent in complex categories. The grades of membership correspond to the degree of compatibility with the concepts represented by the class concerned: the direct evaluation of grades with adequate measures is a significant stage for subsequent decision-making processes. In practices of image classifying, the given objects are not always certain element but a fuzzy sub-set in the fuzzy set. Then practical research should focus on not the subject degree between certain element and the set but the close degree between two fuzzy sub-sets (Binaghi et al., 1999).

3.2 Artificial Neural Networks

Neural network have been proposed in recent years as alternative techniques to traditional classifiers (Moody et al., 1996). The non-parametric nature of neural networks is an advantage over traditional techniques, which assume a Gaussian distribution in the radiometric values of pixels in the image so that Neural networks classifier can define arbitrary decision boundaries. Hence, it is desirable that a feature extraction method for neural networks can preserve that characteristic. Its' another advantage is that they can better integrate non-spectral information, such as ancillary data derived from topographical maps (height and slope), which can be incorporated as an additional input to the network for a better discrimination between terrain covers with similar reflectance. Researches show that neural network methods precede maximum likelihood method, and availably integrate geographical ancillary data, especially when the model of image frequency distributions is uncertain. Unfortunately, neural networks need a long training time but a relatively short classification time for test data. However, with more high-dimensional data and multi-source data available, the resulting neural network can be very complex. Although once the networks are trained, the computational cost of neural networks is much smaller compared with other nonparametric classifiers such as the K-NN classifier (Ambroise and Govaert, 1996), inefficient calculation will be introduced inevitably into networks without efficient feature extraction methods. Thus, a good feature extraction method is necessary for more perfect neural networks.

3.3 Decision Tree Method

A decision tree, having its origin in machine learning theory, is an efficient tool for the solution of classification and regression problems. Unlike other approaches that use a set of features (or bands) jointly to perform classification in a single decision step, decision tree method is based on a multistage or hierarchical decision scheme or a tree like structure. The hierarchical structure of decision tree classifier is more computationally efficient and more flexible than a conventional single-stage one in that the nodes can have different decision rules and subsets of features. And, a tree classifier may circumvent the Hughes effect due to small training sample size by focusing on fewer classes and hence using fewer features at each node. But, decision tree classifiers come with several limitations so that the design of an optimal tree classifier still remains intangible. Furthermore, without considering the optimization of all levels in the tree, errors may accumulate at each level (Safavian and Landgrebe, 1991). So, an optimal decision tree has to consider many factors such as the tree structure, feature reduction method and computational complexity at the same time. Many tree design approaches have been proposed, targeting different design aspects or applications. In section 4, we will employ a decision tree approach to determine relationships of a certain homo-class from different hierarchies for multi-resolution hierarchy dataset so as to produce soft classification and extract semantic information from remote sensing data.

3.4 Texture Analysis

The conventional multi-spectral classification methods have been successfully used for the detection of areal objects from satellite images. However, they are still problematic for the detection of object classes in urban areas. The reasons are: 1), Objects in urban areas are very complicated because they are characterized through more their structures than their spectral reflection properties; 2), The conventional multi-spectral classification methods extract the object classes only according to the spectral information, while a large amount of spatial information is neglected. (Zhang, 1999). In order to extract urban object classes accurately, it is necessary to incorporate spatial and structural information as well as the spectral information. Thus, it would be advisable for discrimination of land-cover classes to add to the radiometric bands of the sensor ancillary information related to the textural features of an image, which can be analysed from the autocorrelation spatial structure of the digital numbers. In this way, the results obtained from pixel-by-pixel classifiers simultaneously taking both radiometric and textural information into account could be improved (Chica-Olmo and Abarca-Hernández, 2000). Texture algorithms are usually divided into three major categories: structural, spectral and statistical. Structural methods consider texture as a repetition of basic primitive patterns with a certain rule of placement (He et al., 1987); spectral methods analyse the power spectrum based on the Fourier transform (Matsuyama, 1980). The third and most important group in texture analysis corresponds to that of statistical methods, which are mainly based on local statistical parameters (Sun and Qin, 1993).

3.5 DISCUSSITON

Based on the analysis of image classification methods, we draw some conclusions as follows:

- Almost all image classification methods are performed based on pixel;
- Almost all image classification methods employ a single kind of image without ancillary data;
- Different image object is always identified within same resolution hierarchy;
- None of image classifiers is a panacea for any case (Giacinto and Roli, 1997).

With the above causes, it is always available for almost all image classifiers they don't obtain satisfactory recognition accuracy and preserve the efficiency of feature extraction, understanding and analysis at certain situation. In fact, the application of image data from certain single remote sensing instrument is constrained by the limitation and difference of geometric, spectral and spatial resolution. And, different spatial objects form a set of pixels with different size and diverse lightness because of their inherent heterogeneous spectral and spatial features. Moreover, apart from the spectral information, there is also spatial information available in the image. Therefore, in the procedure of image classification, it is important and necessary for desirable accuracy to import and integrate multi-source and multi-resolution remotely sensed data describing fully land cover information from different angles. Integration of multi-source and multi-resolution remotely sensed data must contribute to accuracy and reliability of classification enormously. At the same time, with the application of higher resolution images, space-scale-related factor is considerable issue which is the relationship between scale, class definition, and pixel resolution. There are some reasons as follows: 1) Within seemingly homogeneous land cover areas at one resolution, more pixel variability emerges as we move to a higher resolution; 2) By itself, the single pixel is too small to be considered a forest even though spectrally it resembles one, because the human concept of "forest" implies to most individuals a certain minimum spatial extent; 3) In theory, there is the "best" appropriate spatial resolution in relation to the given spatial object. So, there is the optimal result of recognition to identifying the given spatial objects with the best appropriate spatial resolution (Chen and Zhao, 1989; Krawitz, 1974).

4. IMAGE CLASSIFICATION BASED ON MULTI-RESOLUTION HIERARCHY

Based on the above analysis, this section presents a new classification method based on the multi-resolution hierarchy for the purpose to fully employ and integrate multi-source and multi-resolution image data. This method is capable of exploiting the information contained in relationships between each pixel and those that neighbour it, because it gives attention to not only spectral information but also spatial information and even other ancillary information. This method has been performed following steps as:

1. Build the multi-resolution hierarchy frame;
2. Conduct hierarchy classification;
3. Evaluate classification performance.

The following sections outline the main strategies of the method based on multi-resolution hierarchy classification.

4.1 Building A Multi-Resolution Hierarchy Frame

This method begins with building a multi-resolution hierarchy dataset of remote sensing based on natural inherent relation between pixels and multi-resolution hierarchy structure of the spatial entities. Hierarchical classifiers have been employed to frame multi-resolution image dataset offering the "best" appropriate spatial resolution in relation to the given spatial object for classification process. Hence, this method will be available to overcome some of the limitations of single-stage classifiers. There are two ways to build a multi-resolution hierarchy dataset. One is to integrate multi-source and multi-resolution remotely sensed data. Another is to scale transformation of image information based on spectral and spatial distributive feature of pixels so as to obtain anticipant different resolution image data. Scale transformation of image information is divided upscaling from downscaling. In general, remote information is always upscaled. The optimal upscaling transformation methods should preserve inherent information of high resolution data in the process of scale transformation (Hay et al., 1997). This paper employs the segmentation method of scale enlargement by controlling and regulating the segment parameters so as to rebuild multiple resolution images. The result of the segmentation algorithms should achieve two aims, on the one hand, to make the more coarse resolution image obtained close the real image with resolution; on the another hand, to gain more spectral and spatial information of image, along with the decrease of image interference which is the result of high frequency interference (Zhu, 2003).

4.2 Hierarchy Classification

From the spatial structure of view, the available remotely sensed data itself is of hierarchy in the process of the hierarchy classification. Therefore, the hierarchy classification strategies should include inter-hierarchy strategy and hetero-hierarchy strategy. In theory, artificial neural network method and decision tree approach are effective to manage multi-resolution hierarchy data contrary to conventional classifications. Although artificial neural network method and decision tree approach offer many benefits, in practice, they come with several limitations as well. Specially, with more high dimensional data and multi-source data available, the resulting classification can be very complex and consuming. Hence, this paper proposes to employ respectively artificial neural network method and decision tree approach to determine relationships of different classes from homo-hierarchy and relationships of same classes from multi-hierarchy remotely sensed data so as to circumvent these problems by focusing on fewer classes to be identified.

The hierarchy classification starts with inter-hierarchy definition, identification, labelling of natural classes. Within a certain given hierarchy, considering the inherent constraints of BP neural network, such as low convergence, tendency of local extremum, and computational complexity of the number of hidden layers and the number of hidden neurons, we employ directly a Kohonen self-organizing neural network with geographical ancillary data, to determine class proportions and boundary feature extraction so as to produce the relations between spatial objects (Kohonen, 1984). By directly applying the decision boundary feature extraction algorithm to Kohonen self-organizing neural networks, there will be no saving in training time. However, we will obtain a simpler network with better performance.

1. Initialization: select a new feature set using the decision boundary feature extraction algorithm for non-parametric classification.
2. Train Kohonen network:
   (1) Select features by the Parzen density estimator employing the decision boundary feature extraction method. By using the Parzen density estimator for feature extraction, we attempt to preserve the non-parametric characteristics of Kohonen neural networks.
   (2) Train a Kohonen neural network using the selected features. Using a reduced feature set, we attempt to reduce the training time of a neural network and obtain a simpler neural network, further reducing the classification time for test data.
3. Conduct classification: Apply directly the decision boundary feature extraction algorithm to neural networks.

Conventional classification isn't available for management hetero-hierarchy objects, decision tree classifiers belong to a type of hierarchical classifiers in which subsets of classes are processed at multiple stages. Hierarchical classifiers have been known to overcome some of the limitations of single-stage classifiers. In general, decision tree are divided into two different types of homogeneous decision tree and hybrid decision tree based on the rule which algorithms library is uniform or diverse. Homogeneous decision tree is located within homogeneous logic space, at the same time; hybrid decision tree is located within hybrid logic space which is composed of a set of homogeneous logic spaces. This is, there are a set of sub-decision trees which may respectively use a unique classification decision rules or different training algorithms in a global decision tree. In contrast with homogeneous decision tree, hybrid decision tree is available for settling complex problem (Li and Zhang, 2003). The best advantage of hybrid decision tree is the "Selective superiority" character, this is, sub-classes of hybrid logic space may employ different classification rules and feature extraction methods according to class feature in order to accurately classify some certain sub-classes or sub-data. Hybrid decision tree is self-adaptable for the given classification problem so as to make the algorithms flexible, expansible and accurate.

In Figure 1, i =1, 2, …, m indicates the object class. DT_j is the decision tree j (j=1, 2, …, n). P_{ij} is probability measurement level of the decision tree j relative to the object class i, when the data sample k (S_k) is inputted. O_i presents the combined weight probability measurement level of the decision tree j relative to the object class i, when S_k is inputted. W_{ij} indicates the connection weight of the decision tree j relative to the object class i, and it describes the weightiness of the j decision tree relative to the i object class.

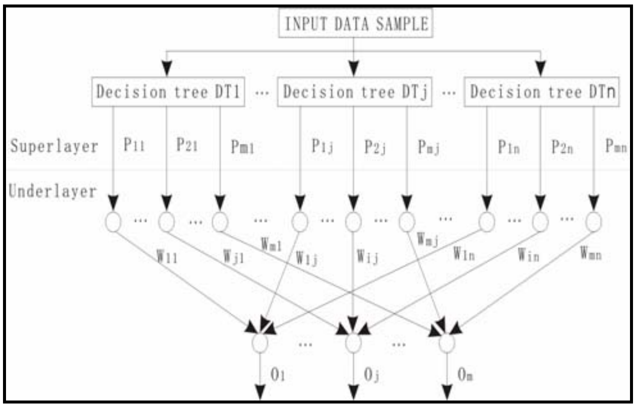

Figure 1. Hierarchy decision tree model

This paper employs hybrid decision tree to determine relationships of a certain homo-class from different hierarchies so as to produce soft classification from remote sensing data, and extract semantic information within multi-resolution hierarchy dataset. Extracted semantic information benefits the better classification of remote sensing image in return. Consequently, this method can be used for integration of remote sensing and GIS at semantic level.

The sequence for hierarchical decision tree classification is described step-by-step as follows:

1. Compute the separability between each class pair;
2. Generate a baseset of sub-classes;
3. Combine sub-class to generate a new combinative framework;
4. Conduct hybrid decision tree;
5. Assess accuracy.

4.3 Classification Validation

Classification accuracy achieved by hierarchy classification algorithms is compared with those achieved by the most widely used maximum likelihood classifier, implemented in the soft mode, and a supervised version of the fuzzy c-means classifier. Root Mean Square Error (RMSE) and fuzzy error matrix based measures have been used for accuracy assessment of soft classification.

5. CONCLUSIONS

Multi-resolution hierarchy classification is a promising approach for integration of remote sensing and GIS at semantic level. In this paper, we have applied this approach for soft classification of remotely sensed data. In particular, we implement inter-hierarchy strategies and hetero-hierarchy strategies in the process of classification. Then, this paper employs respectively Kohonen neural network method and hybrid decision tree approach to determine relationships of different classes from homo-hierarchy and relationships of same classes from multi-hierarchy remotely sensed data so as to circumvent these problems by focusing on fewer classes to be identified. The results show that multi-resolution hierarchy classification is flexible to exploiting the spatial structural information contained in relationships between a pixel and those corresponding to it, because it is implemented to extract the object classes according to not only the spectral information of the individual pixels, but also the spatial or textural information. Consequently, the method is available for the transformation of the pixel-based classification to the object-based classification. And it is a potentially useful approach to extract and analyse semantic information so as to generate semantic-based integrative system of remote sensing and geographic information systems.

REFERENCES

Ambroise, C. and Govaert, G., 1996. Constrained clustering and kohonen self-organizing maps. J. Classification, 13(2): 299-313.

Binaghi, E., Brivio, P.A., Ghezzi, P. and Rampini, A., 1999. A fuzzy set-based accuracy assessment of soft classification. Pattern Recognition Letters(20): 935-948.

Bo, Y.C. and Wang, J.F., 2003. Uncertainly in Remote Sensing: Classification and Scale Effect Modeling. The Geological Publishing House, Beijing.

Campbell, J.B., 2002. Introduction to Remote Sensing. Guilford Press, New York.

Chen, S. and Zhao, Y., 1989. Geo-Science Analysis of Remote Sensing. The Publishing House of Surveying and Mapping, Beijing.

Chica-Olmo, M. and Abarca-Hernández, F., 2000. Computing geostatistical image texture for remotely sensed data classification. Computers & Geosciences(26): 373-383.

Foody, G.M., 1996. Approaches for the production and evaluation of fuzzy land cover classifications from remotely sensed data. Int. J. Remote Sens. (17): 1317-1340.

Ghosh, I. and Samanta, R.K., 2003. TEAPEST: An expert system for insect pest management in tea. Applied Engineering in Agriculture, 19(5): 619-626.

Giacinto, G. and Roli, F., 1997. Ensembles of neural networks for soft classification of remote sensing images, Proceeding of the European Symposium on Intelligent Techniques, Bari Italy, pp. 166-170.

Hay, G.J., Niemman, K.O. and Goodenough, D.G., 1997. Spatial thresholds, image-objects and upscaling: a multi-scale evaluation. Remote Sensing of Environment(62): 1-19.

He, D.C., Wang, L. and Guibert, J., 1987. Texture discrimination based on an optimal utilization of texture features. Pattern Recognition Letters(2): 141-146.

Hsieh, P.-F., Lee, L.C. and Chen, N.-Y., 2001. Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing. IEEE Trans. Geosci. Remote Sens., 39(12): 2657-2663.

Hughes, G.F., 1968. On the mean accuracy of statistical pattern recognizers. IEEE Transactions on Information Theory, 14(1): 55-63.

Kohonen, T., 1984. Self Organization and Associative Memory (2nd ed.). Heidelberg: Springer-Verlag, Berlin.

Krawitz, L., 1974. Earth Resources Program Scope and Information Needs. NASA-CR\141767. General Electric Co., Philadelphia.

Li, D.R., 2003. Towards the Development of Remote Sensing and GIS in the 21st Century. Geomatics and Information Science of Wuhan University, 98(2): 127-131.

Li, S. and Zhang, E.X., 2003. The decision tree classification and its application in land cover. Areal Research and Development, 22(1): 17-21.

Matsuyama, T., 1980. Structural analysis of natural textures by Fourier transformation. Computer Vision. Graphics and Image Processing(12): 286-308.

Moody, A., Gopal, S. and Strahler, A.H., 1996. Sensitivity of neural networks to subpixel land-cover mixtures in coarse-resolution satellite data. Remote Sensing of Environment(58): 329-343.

Richards, J., 1986. Remote Sensing Digital Image Analysis. NY: Spring-Verlag, New York.

Safavian, S.P. and Landgrebe, D., 1991. A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics(21): 660-674.

Settle, J. and Drake, N.A., 1993. Linear mixing and the estimation of ground cover proportions. Int. J. Remote Sens. (14): 1159-1177.

Sun, X.H. and Qin, P., 1993. Texture analysis for remotely sensed imagery, the Ninth Thematic Conference on Geological Remote Sensing, Pasadena, California, pp. 311-322.

Wilkinson, G.G., 2005. Results and implications of a study of fifteen years of satellite image classification experiments. IEEE Trans. Geosci. Remote Sens., 43(3): 443-440.

Xu, M., Watanachaturaporn, P., Varshney, P.K. and Arora, M.K., 2005. Decision tree regression for soft classification of remote sensing data. Remote Sensing of Environment(97): 322-336.

Zhang, X.D., Younan, N.H. and O' Hara, C.G., 2005. Wavelet domain statistical hyperspectral soil texture classification. IEEE Trans. Geosci. Remote Sens., 43(3): 615-618.

Zhang, Y., 1999. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering. ISPRS Journal of Photogrammetry & Remote Sensing(54): 50-60.

Zhu, G.B., 2003. Remote sensing image analysis based on hierarchical multi-resolution structures. Geomatics and Information Science of Wuhan University, 28(3): 315-320.