RESEARCH ON THE VISUALIZATION SYSTEM AND APPLICATIONS OF UNCERTAINTY IN REMOTE SENSING DATA CLASSIFICATION

YanLiu[a] , Yanchen Bo[b]

[a] National Geomatics Center of China, no1. Baishengcun,Zhizhuyuan, Haidian distric, Beijing 100044,China-keanureeves0105@163.com
[b] Beijing Normal University, Research Center for Remote Sensing and GIS, School of Geography, Beijing Normal University, State Key Laboratory of Remote Sensing Science, Beijing 100875, China-boyc@bnu.edu.cn

YF

ABSTRACT:

Thematic mapping is one of the major application fields of remote sensing techniques. Data quality of the thematic data was concerned about. At present, there are a lot of researches that focus on the uncertainty measures, modeling and visualization. But lacking an easy-to-use visualization system of classification uncertainty. The research work toward how to use uncertainty, especially how to use it to improve the classification accuracy is few. Based on the review of existing visualization method and system of the uncertainty of remote sensing data classification, we developed an uncertainty visualization system for remote sensing data classification. In this system, the static gray, color, color-saturation combined and dynamic 3D visualization method were exploited. The system was developed using the IDL language. This system can be integrated with ENVI, an popular software for remote sensing image processing, to use its powerful image processing functions. The system provides users multiple choices of uncertainty measures, such as: absolute uncertainty, relative uncertainty, relative probability entropy, relative maximum deviation. Its data format is compatible with ENVI.

1. INTRODUCTION

There are a lot of uncertainties in the thematic classification of remote sensing. The most common indexes of uncertainty are confusion matrix and Kappa analysis, which just measure the uncertainty mathematically but ignore the spatial distribution characteristics. Therefore it is very difficult to accurately and truly depict the uncertainty of remote sensing information. Uncertainty description using visual method is one of the important part of uncertainty modeling. Visualization technology allows user to explore the uncertainty of origin data and relationship among them, which would help them to realize the effect of uncertainty for project decision.

This paper characterizes the uncertainty information on pixel scale by various visualization techniques. A visualization system is developed, which could integrated operate with ENVI. What is more，besides the traditional visualization techniques (black and white, gray scale image, colour image et al.)，our system introduced the colour saturation combination techniques to characterize the uncertainty information, which can let user to choose colour to represent the class and saturation corresponding to uncertainty. The results of classification and distribution of uncertainty would be showed in the corresponding 2D image.

In this paper we will introduce some uncertainty measurements, the visualization techniques、main framework and functions of system. At last, we will show an example of this system with Landsat TM data about Lake Lanier in USA.

2. THE ACQUISITION OF REMOTE SENSING UNCERTAINTY INFORMATION OF EACH PIXEL

The pixel uncertainty information can be obtained from posterior probability vector during the process of maximum likelihood classification. Derived from the probability vector of each pixel, which gained after the classification, can develop many kinds of measurements of classification uncertainty. The maximum posterior probability vector can provide a kind of measurement for classification uncertainty. The higher the maximum posterior probability, the smaller the classification uncertainty is. In this paper we introduced four different uncertainty measurements.

2.1 Absolute Uncertainty

The definition of absolute uncertainty is:

$$U_A(x) = p(\omega_{l1}|x)/[1 - p(\omega_{l1}|x)] \tag{1}$$

The value range of UA is $(0, +\infty)$, the higher this value, the smaller the uncertainty of defining the pixel x to the class ω_{l1}.

2.2 Relative Uncertainty

The relative uncertainty reveals the misclassification between class ω_{li} and class ω_{lj}, it is defined as:

$$U_R(x,i,j) = \left| p(\omega_{li}/x) - p(\omega_{lj}/x) \right| \tag{2}$$

The value range of $U_R(x,i,j)$ is [0，1]. The higher the value of this parameter, the easier to classify between class ω_{li} and class ω_{lj}.

2.3 Relative Probability Entropy

Relative probability entropy is a concept from information theory, relating to the uncertainty of a statistical variable of different measured values. It reveals the distribution and range of uncertainty.

The pixels' probability entropy shows the needed information for classify the pixel. The measurements for entropy is defined as the sum of weighted posterior probability of the same pixel that belong to different classes, it is defined as:

$$H(p) = -\sum_{i=1}^{M} p(\omega_i / x) \times \log_2 p(\omega_i / x)$$

(3)

Maselli (1994) put forward the Relative probability entropy:

$$H_R(p) = H(p) / H_{max}(p) \times 100$$

(4)

2.4 Relative Maximum Probability Deviation

Another uncertainty measurement which also derived from entropy vector can be defined as:

$$U = 1 - \frac{\max - \dfrac{sum}{M}}{1 - \dfrac{1}{M}}$$

(5)

Max is the maximum entropy, sum is the sum of entropy vector, M is the number of classes.

3. THE REMOTE SENSING CLASSIFICATION VISUALIZATION SYSTEM

3.1 Development Environment

This system is based on IDL language (Interactive Data Language). As the forth generation of computer language, IDL is the most ideal language for data analysis, visualization, multi-platform application and development. It can provide users of perfect tool for image analysis which can let them to analyze the scientific data freely.

The remote sensing image analysis software----ENVI is also based on IDL language, so our system can be seamlessly integrated with ENVI.

3.2 The Uncertainty Visualization Technique

3.2.1. Black-White Binary Image: Set a threshold value, then use the black dot to express the pixel whose value below this threshold and the white dot to those above the threshold. Our aim is to let all the users get a brief review of the whole classification accuracy.

3.2.2. Gray Scale Image: In the gray scale image, the deep color pixels represent the value of high uncertainty. If the uncertainty value is low, the gray scale value will be small. As the change of gray scale is hard for human eyes to recognize, we designed a gray scale scrollbar. See Figure 1, it can let users to set the gray scale which they want to display. Figure 1 shows the scrollbar set by value 0.98 and 1.0. From Figure 1 we can see that if the scrollbar value set is small then the expressed uncertainty will be cursory.
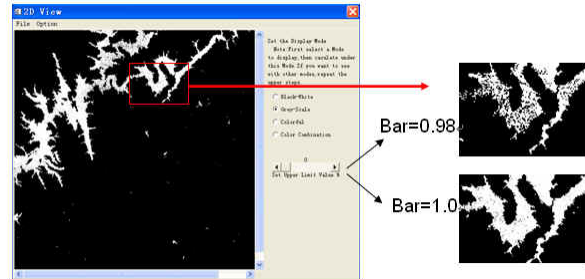

Figure 1.   The scrollbar of gray scale 2-D visualization module

3.2.3. Color Image: Human eyes could not distinguish gray-scale exceeding 64 Level, far less than the capability of distinguishing the color gradient .So the method of using colors to describe the uncertainty of different places could be much more efficient.

Color mapping is a general data visualization method in visualization system, different color corresponding to different value. Mapping between data and color would establish the relationship of them.

Pixel uncertainty value decides its color in Distribution map of uncertainty, so the size and spatial distribution of uncertainty can be reflected by different colors and distribution.

Although this method could use different color combination to presenting the distribution of uncertainty，but probably may face the problem that two adjacent regions may have the similar color that hard to recognize. The solution is we provide 40 different color combinations to the system for color choosing. According to their data，users could select corresponding color combination.

3.2.4. Color and Saturation Combination Image: The human eyes capability of distinguishing gray image is far inferior to color one，whereas uncertainty value of different land-use must have its data range. Based on this，It is suggested that we divide the whole range of value to several intervals to reduce the number of values, similarly discretization Operation.

When user's interests mostly focus on the degree of data uncertainty, not the specific numerical value, the visualization of discretization intervals is a very effective representation.

For example，after the remote sensing image classified ,entropy of pixels belonging to some classes would be normalized to the successive numerical range of [0.00,1.00].Then equidistant discretization would be operated to divide this successive range into five equivalent intervals: (0.00,0.20], (0.20, 0.40], (0.40,0.60], (0.60,0.80], (0.80,1.00], be followed by 0.00，corresponding to " 0, 1, 2, 3，4, 5" six magnitudes of uncertainty. Each land-use class have its unique color, while saturation level depicts the degree of uncertainty

see Figure 8, We named such mapping as color and saturation combination image.

3.2.5. 3-D Dynamic Image: 3-D dynamic image is used to present the data uncertainty in 3-D scene. Compared to 2-D method，3-D dynamic method is more complex and has its obvious merits. It would represent the spatial distribution of pixels uncertainty, and could also map the uncertainty value to Z-axis in 3-D scene. X-Y plane can express the classes，and the altitude represent class uncertainty.

The system adopts 3-D dynamical representation to compare the numerical changes of uncertainty of different locations. With towing the mouse，we can analyze it from differ angles views. In addition，in order to show uncertainty distribution more directly to users, we add scatter plot、exterior view and grid chart three different functions to the system, see Figure 9.

4. FRAMEWORK, FUCTIONS AND INTERACE OF THE VISUALIZATION SYSTEM

The interfaces and menus of Uncertainty Visualization System keep consistent with the style of ENVI software. ENVI provides users with a complete remote sensing image processing functions, such as project convert、data preprocessing and calibration、image correction and mosaicking、filter、spatial transform 、classification et al. Besides supervised and unsupervised classification, it also provides users of interactive analysis and classification accuracy statistics function based on ROI feature space. Consequently, embedding the Uncertainty Visualization System as the subsystem could utilize the existing image processing functions of ENVI software. What is more, the data after processed could be saved as ENVI standard file format.It is very important to facilitate users of using ENVI software to save and view results.
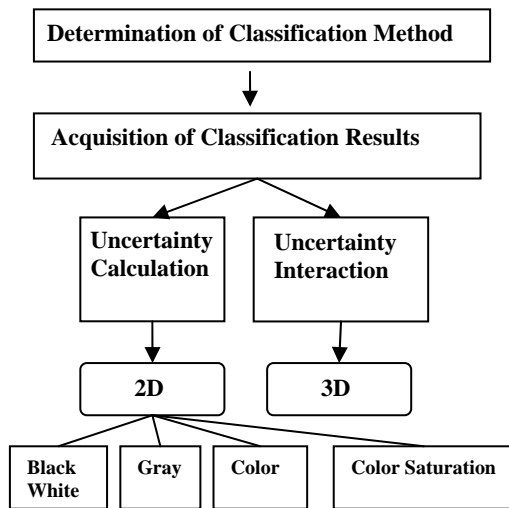

Figure 2.   The framework of Classification Visualization System

Figure 2 shows the framework of Classification Visualization System, it includes two sub module: One module is Uncertainty Calculation Module, it realizes uncertainty indexes calculate and save function based on the classification results image. The indexes are absolute uncertainty、relative uncertainty、relative probability entropy and relative maximum probability deviation;The other is Uncertainty Visualization Module, it

implements the visualization of uncertainty, involving 2-D visualization, 3-D visualization and Columnar Section. 2-D visualization introduce black and white image，gray scale image 、color image 、color and saturation combination image.3-D visualization also includes three kinds of representations. Columnar section shows the probability of one pixel belonging to different classes.

Figure 3 shows the main interface of Uncertainty visualization System.
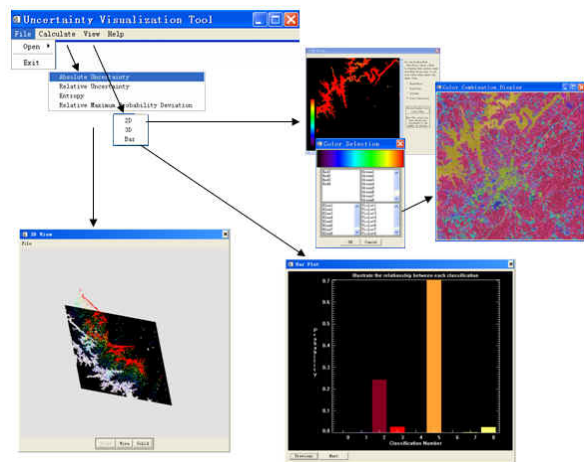

Figure 3.   The main interface of Uncertainty visualization System

5. DEMONSTRATION

We take a Landsat TM image of Lake Lanier in USA as an example. The image would be classified into 9 classes, including cropland, bare soil, grass, conifer, Broadleaf, water, town and developed lands. There is litter cloud and shadow shade at the right upper of image, so we treat them as two new classes. The selected classification method is the maximum likelihood classification. We do not set any threshold during the classification process so there is no "unclassified pixels". The posterior probability of pixel belong to each class would be saved and form the probability vector during the classification. When setting the parameters of classification algorithm, we could save this vector data named Rule Image in ENVI software.
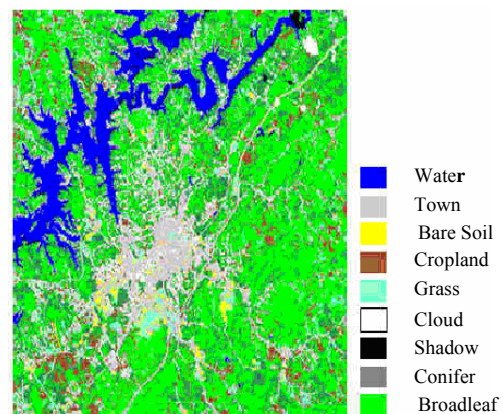
Figure 4. Maximum Likelihood Classification results of Lake Lanier
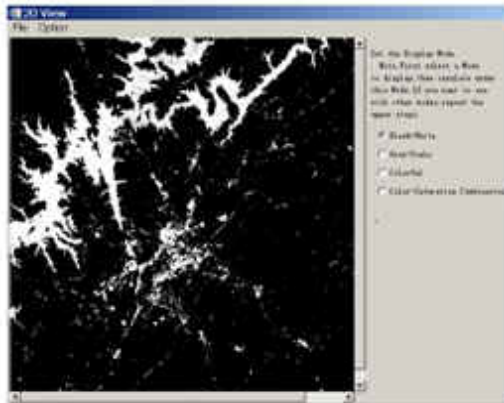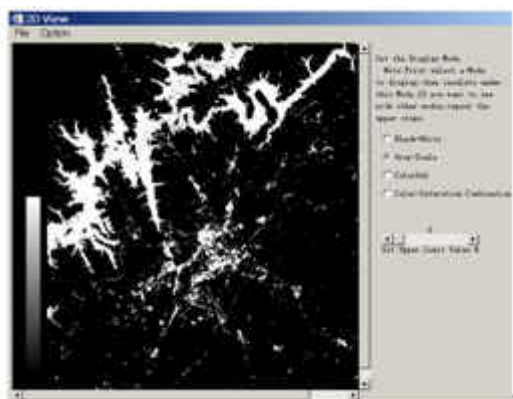

Figure 5. Black and White Image
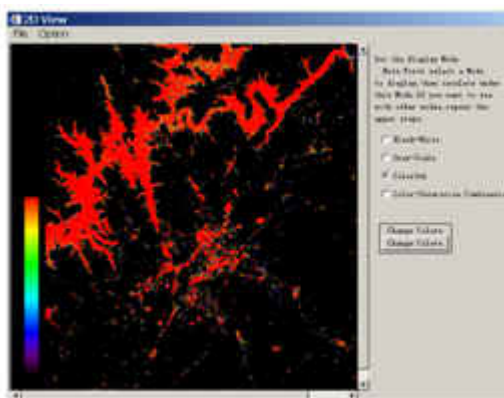

Figure 6. Gray scale Image

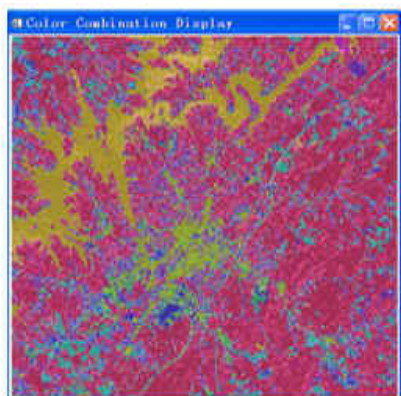Figure 7. Color Image

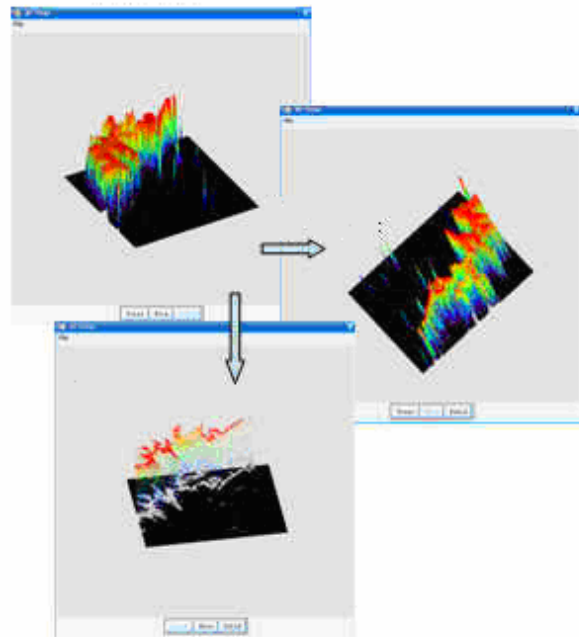Figure 8. Color and Saturation Combination Image

Figure 9. 3-D Visualization

6. DISCUSSION AND CONCLUSION

It is a very important issue of how to measure the uncertainty and which visualization method to be chosen to characterize the spatial distribution of uncertainty more exactly and comprehensively.

We first obtain the spatial distribution of uncertainty on pixel scale then visualize it in our system. A series of uncertainty parameters could be derived from the posterior probability vector, which would be generated from classification based on Bayesian theory. With the extended probability vector formula, we could calculate the uncertainty on pixel scale.

The Remote Sensing Classification Uncertainty Visualization system could visualize and hold the classification uncertainty information. Furthermore, it realizes the seamless integration with ENVI. Besides using the traditional visualization representations, the system developed the color and saturation visualization technique, characterizing the classes and their uncertainty on the same image. This new representation would improve visual effect.

However, there are some shortcomings needs to be optimized on user interactive performance in the system. For instance，establish the dynamic links between spatial distribution image of uncertainty and origin classified image, export pixel uncertainty quantitative value and location with mouse click.

REFERENCES

Alex Pang. Visualizing Uncertainty in Geo-spatial Data. http://www7.nationalacademies.org/cstb/wp_geo_pang.pdf(accessed 2007)

B.Klinkenberg,M.Joy,1994.Visualizing Uncertainty:Succession or Misclassification,*GIS/LIS* , pp.494

Congalton, R.G,Green K., 1999.Assessing the Accuracy of Remotely Sensed Data : Principle and Practices. USA:Lewis Publishers, pp.49

Davis, T.J, Keller, C.P., 1997. Modelling and visualizing multiple spatial uncertainties. *Computers & Geosciences,*23(4) , pp.397

Dykes. J.,1997. Exploring spatial data representation with dynamic graphics.*Computers & Geosciences*, 23(4) , pp. 345

Ehlschlaeger,C.R.,Shortridge,A.M.,M.FGoodchild,1997.Visualizing spatial data uncertainty using animation.*Computers & Geosciences,*23(4) , pp.387

Evans, B.J. 1997.Dynamic display of spatial data-reliability: does it benefit the map user?.*Computers&Geosciences,* 23(4), pp.409

Fisher, P.F.,1994. Visualization of the reliability in classified remotely sensed images. *Photogrammetric Engineering & Remote Sensing*, (60), pp.905

Goodchild, M.F.,Guoqing.S,Shiren,Y.,1992. Development and test of an error model of categorical dat.*International journal of Geographic Information Systems*, 6(2) , pp.87

Hunter, G.J., Goodchild, M.F. , 1995. Dealing with error in spatial databases: a simple case study.*Photogrammetric Engineering&Remote Sensing,* , 61(5) , pp.529

Lucieer, A. Uncertainties in Segmentation and their Visualisation.http://www.parbat.net/ (accessed 2004)

Maceachren,A.M.,Menno-JanKraak,1997.Exploratory Cartographic Visulization:Advancing the Agenda.*Computers and Geosciences*, 23(4) , pp.335

Van der Wel, F.J.M., Gaag, L.C., van der Gorte, B., 1998. Visual exploration of uncertainty in remote sensing classification. *Computer and Geosciences*, 24(4) , pp.335