Texture classification of vegetation cover in high altitude wetlands zone

ZOU Wentao[1, 2], Wu Bingfang[1, *], JU Hongbo[2], Liu hua[2]
[1]Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences
[2]Research Institute of Forest Resource Information Techniques, CAF Beijing, China

E-mail: zouwt@irsa.ac.cn

Abstract. The aim of this study was to investigate the utility of datasets composed of texture measures and other features for the classification of vegetation cover, specifically wetlands. QUEST decision tree classifier was applied to a SPOT-5 image sub-scene covering the typical wetlands area in Three River Sources region in Qinghai province, China. The dataset used for the classification comprised of: (1) spectral data and the components of principal component analysis; (2) texture measures derived from pixel basis; (3) DEM and other ancillary data covering the research area. Image textures is an important characteristic of remote sensing images; it can represent spatial variations with spectral brightness in digital numbers. When the spectral information is not enough to separate the different land covers, the texture information can be used to increase the classification accuracy. The texture measures used in this study were calculated from GLCM (Gray level Co-occurrence Matrix); eight frequently used measures were chosen to conduct the classification procedure. The results showed that variance, mean and entropy calculated by GLCM with a 9*9 size window were effective in distinguishing different vegetation types in wetlands zone. The overall accuracy of this method was 84.19% and the Kappa coefficient was 0.8261. The result indicated that the introduction of texture measures can improve the overall accuracy by 12.05% and the overall kappa coefficient by 0.1407 compared with the result using spectral and ancillary data.

Key words: texture measures, wetlands, vegetation cover, decision tree

1. Introduction

Remote Sensing technology has been applied in wetlands research for several decades. Along with the deep-going research on wetlands and the improvement of the spatial resolution of remote sensing data, remote sensing has become an important tool in wetland vegetation identification and classification [1].

Researches on wetlands vegetation classification has been mainly concentrated in the coastal wetlands and the commonly used remote sensing data were TM images [2,3]. TM images were also used in marsh vegetation classification [4], but the result showed that the coarse resolutions of TM images may cause uncertainty in border demarcation among different communities. Existing studies mainly put emphasis on the use of spectral information while ignoring the image texture features exhibited by the remote sensing images [5]. Therefore, it is necessary to demonstrate the effects of comprehensive utilization of texture, spectral and other characteristics represented by high-resolution remotely sensed images in wetlands vegetation classification. In this study, we chose QUEST Algorithm to establish the decision tree automatically and use the data set comprise spectral features, texture and other ancillary features to classify the vegetation in plateau where abundant wetlands exist.

2. Study Site

The study site is located in the SuoJia-QuMahe Nature Reserve, Qinghai province, China. The geographical range of this area is 34°5′-34°20′N latitude and 93°56′-94°12′E longitude. Summer in this area is short and winter is long, and there is no absolute frost-free period. The annual mean temperature is -2.5℃， annual mean precipitation is 400 mm, the altitude ranges from 4230 m to 5600 m, and this area shows a typical plateau continental climate [6].

3. Materials and method

3.1 Data and processing

The satellite data used in this study is half scene SPOT-5 image acquired on the 30th April, 2010 with orbit number 238-281.The resolution of SPOT-5 panchromatic band is 2.5 meters, and the multispectral band is 10 meters. In addition, the1:100000 topographic map, 1:100000 wetland distribution map, 1:100000 vegetation distribution map that covered this area were collected as supplementary data for training and accuracy assessment.

3.2 Classification scheme

Based on the field survey in August, 2008 and 2009 and the characteristics demonstrated by remote sensing image, the vegetation cover types were divided into eight categories. Taking into account of rivers, bottomlands and barren lands, the SPOT5 image covered this selected area was divided into a 11 classes. The name of each vegetation cover was determined by the dominance vegetation type in the community according to the field investigation. For example: *Kobresiahumilis* + *Kobresiatibetica* swampy meadows refers to the dominant vegetation species in this area are *Kobresiahumilis* and *Kobresiatibetica*, and its companion species include *Kobresiacapillifolia*, *Arenariakansuensis* and *BromusinermisLeyss* etc. Detailed divisions are shown in Table 1.

3.3 Classification Method

Decision tree is composed of a series of binary tree classifiers that based on the judge rules. It firstly divides each image into relative homogeneous subsets and then decides their proper cover types. Decision tree method is considered as a good method for remote sensing images classification. The QUEST algorithm is also accepted as a good method that has distinct advantages in operational efficiency and high classification accuracy when compared with other algorithms [7, 8]. For these reasons, the QUEST algorithm was used as decision tree building method to differentiate vegetation cover in the study area.

3.4 Classification Feature Variables

Spectrum is the most important information in remote sensing classification. Different land covers have their own unique spectral characteristics due to the differences in material composition and structure. Four bands after the fusion of panchromatic and multispectral band in SPOT5 images were chosen as test variables. In addition, the first principal component after the PCA transformation towards the fused images was also selected as one feature in the classification.

Textural characteristics are effective to improve the classification result of high resolution images. The Gray Level Co-occurrence Matrix method was commonly accepted as a well-performed method for texture extraction using remotely sensed images [9, 10]. In our calculation, the parameters were determined as follows: the moving step was set to 1, the mean value of four directions 0 °, 45 °, 90 ° and 135 ° was chosen as the final texture values for texture extraction; the size of moving window was defined as 9*9. The eight texture features, including Mean (ME), Variance (VA), Homogeneity (HO), Contrast (CO), Dissimilarity (DI), Entropy (ENT), Second Moment(SM) and Correlation (COR) were calculated using ENVI software.

Table 1. SPOT5 classification scheme

Types	Description
Rivers	Opening water, Long line shaped.
Bottom Lands	Nearby the rivers, formed by mud and sand deposition.
Barren Lands	Exposed rock or lands without cover of vegetation.
***Kobresiapygmaea*Meadows**	Dominant Species: *Kobresiapygmaea*. Companion Species: *Potentillabifurca* and *Saussurea japonica*.
***Kobresiahumilis*Meadows**	Dominant Species: *Kobresiahumilis*. Companion Species: *Kobresiacapillifolia*.
***Polygonumviviparum*Meadows**	Dominant Species: *Polygonumviviparum*. Companion Species: *Kobresiapygmaea*.
***Potentillafruticosa*shrub Meadows**	Dominant Species: *Potentillafruticosa*. Companion Species: *Gramineae*.
***Kobresiatibetica+Kobresiahumilis*Meadows**	Dominant Species: *Kobresiatibetica+Kobresiahumilis*. Companion Species: *Oxytropis.DC,Saussurea japonica*.
***Kobresiatibetica*Swampy Meadows**	Dominant Species: *Kobresiatibetica*. Companion Species: *Carex spp.*, *lysmussinocompressus*, *P.longiflora var. tubiformis.*
***Kobresiahumilis+Kobresiatibetica* Swampy Meadows**	Dominant Species: *Kobresiahumilis+Kobresiatibetica.* Companion Species: *Kobresiacapillifolia*, *Arenariakansuensis*, *BromusinermisLeyss.*
***Hippuris Vulgaris* Marshes**	Dominant Species: *Hippuris Vulgaris.* Companion Species: *Potamogetonoctandrus*, *Halerpestestricuspis.*

Note: In the following tables and figures, type1 represents Rivers; type2 represents Bottom Lands; type3 represents Barren Lands; Type4 represents*Kobresiapygmaea* Meadows; Type5 represents *Kobresiahumilis*Meadows; Type6 represents *Polygonumviviparum* Meadows; Type7 represents *Potentillafruticosa* shrub Meadows; Type8 represents *Kobresiatibetica +Kobresiahumilis* Meadows; Type9 represents *Kobresiatibetica*Swampy Meadows; Type10 represents *Kobresiahumilis +Kobresiatibetica* Swampy Meadows; and Type11 represents *Hippuris Vulgaris* Marshes.

4. Results

4.1 Decision Tree Model

The QUEST decision operation was performed under the SPSS 18.0 software environment. Using a training set obtained from variables in the data set, the depth of the Decision tree was defined as 8 to avoid over fitting in the training process.

The decision tree model obtained above by the QUEST algorithm shows that: the introduction of texture features can distinguish rivers and bottomlands from *Potentillafruticosa* meadows, *Polygonumviviparum* meadows, *Kobresiapygmaea* meadows, *Kobresiahumilis* meadows and *Kobresiahumilis* swampy meadows efficiently, and is also very effective to distinguish between *Kobresiaschoenoides+Kobresiahumilis* meadows and *Kobresiahumilis* meadows.

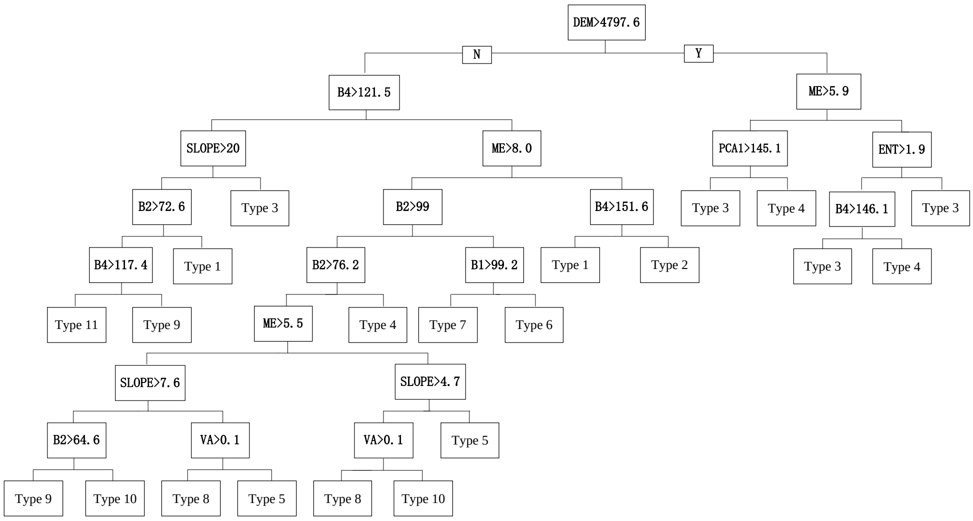

Figure2. Classification Based on data set comprised of Textures, spectral and ancillary data

The texture feature ENT can be used to differentiate barren lands from *Kobresiapygmaea* meadows. Introduction of VA can greatly improve the classification accuracy between *Kobresiaschoenoides+ Kobresiahumilis*meadowss and *Kobresiahumilis* meadows or enhance the distinction between *Kobresiaschoenoides+ Kobresiahumilis* meadows and *Kobresiahumilis+Kobresiaschoenoides* swampy meadows. From the figure above, we also found that the adding of Elevation and Slope data can play an important role to separate swampy vegetation cover and *Kobresiapygmaea* meadows , barren lands or other land types.

4.2 Classification results

The decision tree model described above was first run to obtain a preliminary classification map of selected image. Then post-processing of the preliminary result was accomplished using the following method: recoding, clustering and removing analysis. Finally, the vegetation cover map in wetlands densely distributed area was obtained (Figure 3).

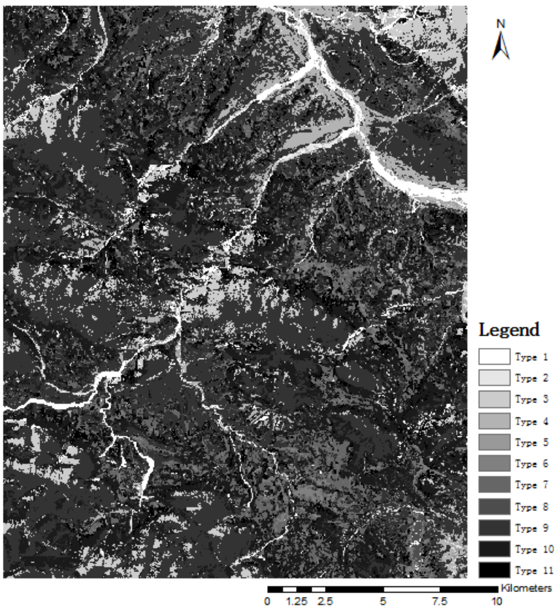

Figure3. Result of Wetlands Vegetation Cover by QUEST Decision Tree

4.3 Accuracy validation and comparison

The confusion matrices were used to verify the classification result using additional texture measures. The results with and without the additional use of texture data were both achieved. Two QUEST Decision Tree classification processes were conducted using the same training samples. Classification accuracies of the two classification methods were compared in the following Table 4.

It can be seen from table 4 that the classification based on the data set only using spectral features and ancillary data got a relative low accuracy. Many cover types, e.g. *Kobresiatibetica*swampy meadows, *Kobresiahumilis+Kobresiatibetica* swampy meadows, *Hippuris Vulgaris* marshes, *Kobresiapygmaea*meadows, etc., were confused with others. The overall accuracy was 66.80% and the corresponding kappa coefficient was 0.6349. The accuracy cannot satisfy the application requirements for the low credibility of the result. In contrast, by introducing the additional texture variables, the latter QUEST decision tree method improved the accuracy significantly for the cover types that tend to confused with others. The vegetation communities such as *Kobresiahumilis* Meadows, *Kobresiahumilis* and *Kobresiatibetica* swampy meadows, *Hippuris Vulgaris* marshes, *Kobresiatibetica* Meadows, *Kobresiatibetica* and *Kobresiahumilis* meadows have higher classification accuracy than other types. The producer's and user's accuracy of these vegetation communities reached 95.83%, 100%; 95.45%, 91.30%; 95.00%, 82.61%; 94.74%, 78.26% and 91.67%, 95.65% respectively with the usage of texture variables. This method obtained an overall accuracy of 84.19% with a corresponding kappa coefficient valued 0.8261; the improvement of overall accuracy and kappa coefficient was 17.29% and 0.1912, respectively.

Table 4. Accuracy and Kappa Coefficients of Different Methods

Types	Data Set comprised of spectral and ancillary data			Data Set comprised of texture, spectral and ancillary data		
	Producer's accuracy	User's accuracy	Overall accuracy	Producer's accuracy	User's accuracy	Overall accuracy
Type1	93.75%	65.22%	66.80%	100.00%	69.57%	84.19%
Type2	62.07%	78.26%		71.43%	86.96%	
Type3	60.71%	73.91%		65.52%	82.61%	
Type4	85.71%	52.17%		85.71%	73.91%	
Type5	77.27%	73.91%		74.07%	86.96%	
Type6	69.57%	69.57%		94.74%	78.26%	
Type7	63.64%	60.87%		95.45%	91.30%	
Type8	86.67%	56.52%		95.00%	82.61%	
Type9	57.14%	69.57%		72.00%	78.26%	
Type10	56.00%	60.87%		95.83%	100.00%	
Type11	48.15%	56.52%		91.67%	95.65%	
Kappa coefficient	0.6349			0.8261		

In summary, the accuracy of the QUEST decision tree method can be significantly improved if based on a combination of texture features, spectral features and other ancillary data compared to that based only on spectral and ancillary data. This conclusion is supported by all the cases in single class, overall accuracy and kappa coefficient. The higher accuracy can fulfil the requirements in practical works.

5. Conclusion and Discussion

5.1 Conclusion

1) The QUEST Decision Tree based on texture, spectral and ancillary data is an effective method for high plateau wetlands vegetation classification. The data set used in this study is composed of texture variables calculated by GLCM, spectral characteristics of SPOT-5 fused image, the first component of

Principal Component Analysis, DEM and other ancillary data. The overall accuracy of this method is 84.19%, which can fulfil the requirements in practical works.
2) The comparison is conducted between the QUEST decision tree method with and without texture features, the results showed that the classification method based on the spectral and ancillary data achieved an accuracy of 66.80% and the corresponding kappa coefficient was 0.6349. While with the additional application of texture variables, the result reached an overall accuracy of 84.19% with a corresponding kappa coefficient of 0.8261. The improvement of overall accuracy and kappa coefficient was 17.29% and 0.1912, respectively. Our research revealed that the texture characteristics of remotely sensed images is one effective feature in classification of wetlands vegetation in plateau area.
3) Texture variables including Mean, Variance and Entropy are effectual features in classification of wetlands vegetation in plateau area. There are eight texture variables calculated by GLCM in our research. However, only three features Mean, Variance and Entropy were proved to be effective in the classification process.

5.2Discussion
1) The eight texture features selected in our study are based on the GLCM method. More studies are needed to identify whether other texture variables calculated by GLCM are also effective in the correlative studies.
2) Only the 9*9 window scale was selected in GLCM method. More researches are needed to confirm if textures obtained under other scales are effective in accuracy improvement in classification of wetlands vegetation and how to determine the optimal size of the moving window.

Acknowledgments
This research was funded by the "Strategic Priority Research Program - Climate Change: Carbon Budget and Related Issues" of the Chinese Academy of Sciences (CAS) (Grant NO. XDA05050109).

References
[1]Zou Wentao. 2011. Dynamic Analysis of Typical Wetlands in Three Rivers' Source Region under Climate Change.*Ph.D. dissertation of Chinese Academy of Forestry*, 9-16.
[2] Guan Yujuan, Zhang Liquan. 2008. Application of inter-tidal wetlands classified by image fusion technique. *Marine Environmental Science*,**27** 647-652.
[3]Huang Huamei, Zhang Liquan. 2005. The vegetation resource at the intertidal zone in Shanghai using remote sensing. *Acta Ecologica Sinica*,**25** 2886-2693.
[4]Chen Dinggui, Zhou Demin, Lv Xianguo. 2007. A Study on Classification of Wetland Communities in Honghe National Nature Reserve by Remote Sensing. *Remote Sensing Technology and Application*, **22** 485-491.
[5] DemareyD. 2005. Discrimination of Wetland Vegetation UsingClose-Range Remote Sensing. Ph.D. *dissertation of University of Nebraska*, **25**2686-2693.
[6] Zou Wentao, Zhang Huaiqing, Ju Hongbo etal.,2010. Study on Remote Sensing Classification of Land Use in the Nature Reserve of the Three Rivers Source Region. *Forest Resources Management*, **6** 90-96.
[7] Kim H, Lo h W. 2001. Classification trees with unbiased multi-way splits. *Journal of the American Statistical Association*, **96** 598-604.
[8] Pal M, Mather P. 2003. An assessment of the effectiveness of decision tree methods of land cover classification. *RemoteSensing of Environment*, **86** 554-565.
[9] Nyoungui A，Tonye E，Akono A. 2002. Evaluation of Speckle Filtering and Texture Analysis Methods for Land CoverClassification from SAR Images.*International Journal of Remote Sensing*, **23** 1895-1925.
[10] Yan Meichun, Zhang Youjing, Bao Yansong. 2004. Deriving Bamboos from IKONOS Image by Texture Information.*Remote Sensing Information*,**19** 31-35.