Random forest and watershed transformations for mapping quasi-circular vegetation patch using fused multispectral CBERS-04 images

Qingsheng Liu1,*

1 State Key Lab. of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing China

*Corresponding author e-mail: liuqs@lreis.ac.cn

Abstract. To identify the quasi-circular vegetation patches (QVPs) and monitor their pattern dynamic is the most essential procedure for studying the establishment, encroachment, and maintenance of the QVPs. High resolution satellite remotely sensed technique is a cost-effective approach to distinguish the QVPs. However, the adhesion between the QVPs or between the QVPs and the vegetations of other shapes makes the detection accuracy of the QVPs low. This study used two-seasonal CBERS-04 pansharpened multispectral images to detect the QVPs by integrating random forest classification and watershed transformation image segmentation technique. The results showed that the method proposed by this study could raise the detection accuracy of QVPs, and 5 m spatial resolution CBERS-04 pansharpened imagery may be enough to obtain the number of the QVPs, but not sufficient to calculate the area and shape of QVPs. In the future, more advanced image segmentation algorithms and finer resolution images should be applied to detect the QVPs.

1. Introduction

To map the quasi-circular vegetation patches (QVPs) and monitor their pattern dynamic is the most essential step for studying the establishment, encroachment, and maintenance of the QVPs [1]. Previous researches indicate that high resolution satellite remote sensing technique is a cost-effective approach to detect the QVPs. In general, the image at a spatial resolution of better than 10 m can be used to map the QVPs [2-6].

China-Brazil Earth Resource Satellite 04 (CBERS-04) has a panchromatic and multispectral camera which image the Earth with a panchromatic band (a spatial resolution of 5 m) and three multispectral bands (a spatial resolution of 10 m). The CBERS-04 panchromatic image can result in similar detection accuracy of the QVPs with that of Gaofen 1 satellite panchromatic image (a spatial resolution of 2 m) [6]. Both the single-date and the multitemporal multispectral fused CBERS-04 multispectral images (5 m) have also been used to map the QVPs, which demonstrate that for single-date image, the image acquired in the spring season can produce a high detection accuracy, and for multi-date combined images, the combined data from the spring and winter season (late autumn) can result in higher detection accuracy [7-9]. However, whether the single-date or multi-seasonal images, or the different classification methods, the detection accuracy is low (the highest precision rate, the highest recall rate, and the highest F values were 66.3%, 77.6%, and 58.5%) [7-9]. It may be attributed
to the adhesion between the QVPs and the QVPs, and the adhesion between the QVPs and the vegetations of other shapes.

The watershed transformation is one of the usually applied image segmentation methods [10-12], which are widely used to segment the adherent cell images [13, 14], and map the individual tree [15]. So far, it has not been applied in the detection of the QVPs.

Following Liu et al. [9], the goal of this work is to segment the QVPs classification results of random forest (RF) on basis of the early spring-winter combined CBERS-04 data by the watershed transformation.

2. Materials and methods

2.1. CBERS-04 image pre-processing

In this research, the spring CBERS-04 image (acquired in 27 March 2016) and autumn image (acquired in 24 October 2015) were used to map the QVPs. The March image was first geometrically corrected using twenty GCPs, and the October image was registered to the geometrically corrected March image. The Gram-Schmidt sharpening method, one of image spectral sharpening approaches, was used to sharpen three 10 m multispectral bands with one 5 m panchromatic band [16]. Then, the quick atmospheric correction method was applied to transform the digital numbers to the spectral reflectance [17]. The study site of interest was clipped from the whole scene (see Figure 1). These image processing works were performed in ENVI V5.5 software.

![Figure 1. The left image is subset of the pansharpened CBERS-04 imagery (a false color RGB image, R-Near infrared, G-Red, and B-Green band) acquired on 37 March 2016. The right image is the classification result from random forest (the white is the vegetation).](image)

2.2. Random forest classification

The RF method has been widely used in recent years. The studies discussed in the literature demonstrate that the RF method can obtain higher accuracy for classifying land cover, identifying crop types, and detecting tree species [18-20]. The RF was performed using the ImageRF method of EnMap-Box v2.20, which can be as a plugin embedded into ENVI software [21]. In this work, three pansharpened multispectral bands, and nine spectral indices (normalized difference vegetation index (NDVI), optimized soil adjusted vegetation index (OSAVI), normalized difference water index (NDWI), two-band enhanced vegetation index (EVI2), modified triangular vegetation index (MTVI2), red-green index (RG), green chlorophyll vegetation index (GCVI), the brightness component of tasseled cap transformation (TCB), and the greenness component of tasseled cap transformation (TCG)) resulted from three multispectral bands were inputted into the RF classification as the
predictive variables. The methods for calculating nine spectral indices can be found in the literature [9]. The optimal predictive variables are determined based on variable importance analysis and classification accuracy assessment. According to our previous research [9], twenty variables was elected to classify the QVPs, excluding RG of March image, and green band, NDWI, and RG of October image. The number of trees and the number of possible splitting variables were set to the default values in RF for each tree node, which have been proved that they are generally a good selection. The training samples were 816, 1875, and 609 pixels for the QVPs, bare soil, and water bodies, respectively, and the validation samples included the number of the QVPs, bare soil, and water bodies were 516, 536, and 358 pixels for accuracy assessment, respectively. The RF classification result were removed the small scattered patches and noise using the Sieve Function (5 pixels) in ENVI v5.1 software. The RF classification result were combined into two classes: one class was the QVPs, another was the result combined the bare soils with water bodies. Finally, the processed RF classification result were resampled into 1 m resolution image using nearest neighbor method, and then converted to the JPG format image as the input image for watershed transformation.

2.3. Watershed transformation

The watershed transformation was used to segment the resampled 1 m resolution RF classification image in JPG format. Then, the centroid, area, and perimeter of each object in segmentation image were calculated, and the boundaries of each object was measured. On the basis, the distance of each point to the centroid was calculated for each object. After that, the statistical analysis was done to obtain the maximum distance, minimum distance, mode of the distance, the difference of the maximum and minimum distance (DMM), and a ratio between the maximum and minimum distance (RMM). Through many tests, the thresholds were determined: the maximum distance was less than 40, the minimum distance was more than 4.85, the DMM was less than 72, and the RMM was less than 3, and the area ranged between 50 and 200.

3. Results and Discussions

The overall accuracy of RF classification result was 99.4%, and the kappa value was 0.990, respectively. The final processed RF classification result was shown in the right image of Figure 1. The precision rate \( A_d \) was 58.4%, and the recall rate \( A_r \) 42.4%, and F value \( F \) 49.2%, respectively. The main reasons were from two aspects, one was because of the omission on small size QVPs (the area was less than 50), and another was attributed to the adhesions between two or more QVPs or between the QVPs and the vegetation of other shapes. Except the omission, the small size QVPs have the similar area, the DMM, and the RMM with those of vegetations of other shapes in the 5 m spatial resolution CBERS-04 pansharpened images, which can result in misclassification.

Figure 2 was the detection result using the abovementioned thresholds. The actual number of QVPs were 139 [9]. The detected number of the QVPs using the method in this study was 159, and the correctly detected number of the QVPs was 83. Overall, the \( A_d \), \( A_r \), and the F measure were 52.2%, 59.7%, and 55.7%, respectively.

Compared with the detection accuracy without image segmented by watershed transformation, the detection accuracy was raised by integrating the random forest classification and watershed transformation segmentation. The boundaries of the QVPs were rough, which indicated that the CBERS-04 sharpened image may be suitable for mapping the number of the QVPs, but not sufficient to measure the area and appearance of QVPs. In the future, finer resolution remote sensing imagery should be applied to detect the decametric-scale QVPs and measure their area and shapes. In addition, more advanced image segmentation algorithms should be applied to distinguish the QVPs in the future.

4. Conclusion

The adhesion between two or more QVPs or between the QVPs and the vegetation of other shapes is a key limitation for accurately detecting QVPs using remote sensing images. In this work, two-seasonal CBERS-04 pansharpened multispectral images were used to detect the QVPs by integrating random forest classification and watershed transformation image segmentation. The results showed that compared with the detection accuracy without image segmentation, the detection accuracy
of the QVPs was improved by the method proposed by this study. It indicated the method in this work was effective for detecting the QVPs in the YRD, China. However, there was a problem of over-segmentation, moreover, the resolution of CBERS-04 image was not sufficient to measure the area and appearance of QVPs. In the future, more advanced image segmentation algorithms and finer resolution images should be applied to detect the QVPs.

Figure 2. The QVPs detected by integrating random forest classification with watershed transformation using two-seasonal pansharpened CBERS-04 multispectral images.

5. Acknowledgments
This research was jointly financially supported by the National Natural Science Foundation of China (Project No. 41561144012, 41661144030, 41671422), the Strategic Priority Research Program of Chinese Academy of Sciences (Project No. XDA 20030302), the National Key Research and Development Program of China (2016YFC1402701), the Innovation Project of LREIS (Project No. 088RA20CYA, 08R8A010YA), and the National Mountain Flood Disaster Investigation Project (Project No. SHZH-IWHR-57). Thanks to China Center for Resources Satellite Data and Application for providing the CBERS-04 data products.

References
[1] Q. S. Liu, G. H. Liu, C. Huang, H. Li, Soil physicochemical properties associated with quasi-circular vegetation patches in the Yellow River Delta, China, Geoderma 337 (2019) 202-214.
[2] Q. S. Liu, G. H. Liu, C. Huang, C. J. Xie, L. Shi, “Using ALOS high spatial resolution image to detect vegetation patches,” in Proc. 2011 3rd International Conference on Environmental Science
and Information Application Technology (ESIAT 2011), Procedia Environmental Sciences 10 (2011), pp. 896-901.

[3] Q. S. Liu, G. H. Liu, C. Huang, C. J. Xie Using SPOT 5 fusion-ready imagery to detect Chinese tamarisk (Slatcedar) with mathematical morphological method, Int. J. Digit. Earth 7 (2014), 217-228.

[4] Q. S. Liu, G. H. Liu, C. Huang, L. Shi, J. Zhao, “Monitoring vegetation recovery at abandoned land,” in the 8th International congress on Image and Signal Processing (CISP 2015), Institute of Electrical and Electronic Engineers, Inc, 2015, pp. 88-92.

[5] Y. J. Zhang, Q. S. Liu, G. H. Liu, S. J. Tang, “Mapping of circular or elliptical vegetation community patches: a comparative use of SPOT-5, ALOS and ZY-3 imagery,” in the 8th International congress on Image and Signal Processing (CISP 2015), Institute of Electrical and Electronic Engineers, Inc, 2015, pp. 666-671.

[6] Q. S. Liu, G. H. Liu, C. Huang, B. W. Yu, Comparison of CBERS-04, GF-1, and GF-2 satellite panchromatic images for mapping quasi-circular vegetation patches in the Yellow River Delta, China, Sensors 18 (2018), 2733.

[7] Q. S. Liu, “Using the CBERS-04 multispectral data tasselled cap transformation to detect the quasi-circular vegetation patches,” in IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019), Institute of Electrical and Electronic Engineers, Inc, pp. 3708-3711.

[8] Q. S. Liu, “Comparing the Different Seasonal CBERS 04 Images to Map the Quasi-circular Vegetation Patches in the Yellow River Delta, China,” in 2018 International Conference on Sensor Networks and Signal Processing (SNSP 2018), Institute of Electrical and Electronic Engineers, Inc, pp. 328-333.

[9] Q. S. Liu, H. W. Song, G. H. Liu, C. Huang, H. Li, Evaluating the potential of multi-seasonal CBERS-04 imagery for mapping the quasi-circular vegetation patches in the Yellow River Delta using random forest, Remote Sens. 11 (2019), 1216.

[10] L. Vincent and P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE PSPB 13 (1991), 583-598.

[11] J. B. T. M. Roerdink and A. Meijster, The watershed transform: definitions, algorithms and parallelization strategies, Fundamenta Informaticae 41 (2001), 187-228.

[12] N. Salman, Image segmentation based on watershed and edge detection techniques, The International Arab Journal of Information Technology 3 (2006), 104-110.

[13] R. Tomari, W. N. W. Zakaria, R. Ngadengon, M. H. A. Wahab, Red blood cell counting analysis by considering an overlapping constraint, ARPN Journal of Engineering and Applied Sciences 10 (2015), 1413-1420.

[14] Y. Wang, Adaptive marked watershed segmentation algorithm for red blood cell images, Journal of Image and Graphics 22 (2017), 1779-1787. (In Chinese with English abstract)

[15] Y. S. Guo, Q. S. Liu, G. H. Liu, C. Huang, Individual tree crown extraction of high resolution image base on marker-controlled watershed segmentation method, Journal of Geo-Information Science 8 (2016), 1259-1266. (In Chinese with English abstract)

[16] K. Nikolakopoulos and D. Oikonomidis, Quality assessment of ten fusion techniques applied on Worldview-2, Euro. J. of Remote Sens. 48 (2015), 141-167.

[17] L. S. Berbstein, X. M. Jin, B. Gregor, and S. M. Adler-Golden, Quick atmospheric correction code: algorithm description and recent upgrades, Optical Engineering 51 (2912), 111719.

[18] Y. Y. Zhao, D. L. Feng, L. Yu, X. Y. Wang, Y. L. Chen, Y. Q. Bai, H. J. Hernandez, M. Galleguillos, C. Estades, G. S. Biging et al., Detailed dynamic land cover mapping of Chile: Accuracy improvement by integrating multi-temporal data, Remote Sens. Environ. 183 (2016), 170-185.

[19] P. Y. Hao, M. Q. Wu, Z. Niu, L. Wang, Y. L. Zhang, Estimation of different data compositions for early-season crop type classification, PeerJ 6 (2018), e4834.

[20] R. L. Pu, S. Landry, Q. Y. Yu, Assessing the potential of multi-seasonal high-resolution Pleiades satellite imagery for mapping urban tree species, Int. J. Appl. Earth Obs. 71 (2018), 144–158.
[21] S. van der Linden, A. Rabe, C. Oldenburg, M. Held, B. Jakimow, P. J. Leitao, A. Okujeni, M. Schwieder, S. Suess, P. Hostert, The EnMap-Box-A toolbox and application programming interface for EnMAP data processing, Remote Sens. 7 (2015), 9020-9044.