

PURIFICATION OF TRAINING SAMPLES BASED ON SPECTRAL FEATURE AND SUPERPIXEL SEGMENTATION

Xuehua Guan¹, Wenwen Qi^{1,2}, Jianjun He^{1,*}, Qiang Wen¹, Ting Chen¹, Zhiyong Wang³

¹ Twenty-First Century Aerospace Technology Co., Ltd., 100096, Beijing, China - (guanxh, hejj, wenqiang, chenting)@21at.com.cn

² Beijing Twenty-First Century Science and Technology Development Co., Ltd., 100096, Beijing, China - qiww@lreis.ac.cn

³ Beijing Engineering Research Centre of Small Satellite Remote Sensing Information, Beijing, 10096, China-wangzy@21at.com.cn

KEY WORDS: Beijing-2 Satellite imagery, Training Samples, Purification, Spectral Feature, Superpixel Segmentation

ABSTRACT:

Remote sensing image classification is an effective way to extract information from large volumes of high-spatial resolution remote sensing images. Generally, supervised image classification relies on abundant and high-precision training data, which is often manually interpreted by human experts to provide ground truth for training and evaluating the performance of the classifier. Remote sensing enterprises accumulated lots of manually interpreted products from early lower-spatial resolution remote sensing images by executing their routine research and business programs. However, these manually interpreted products may not match the very high resolution (VHR) image properly because of different dates or spatial resolution of both data, thus, hindering suitability of manually interpreted products in training classification models, or small coverage area of these manually interpreted products. We also face similar problems in our laboratory in 21st Century Aerospace Technology Co. Ltd (short for 21AT). In this work, we propose a method to purify the interpreted product to match newly available VHRI data and provide the best training data for supervised image classifiers in VHR image classification. And results indicate that our proposed method can efficiently purify the input data for future machine learning use.

1. INTRODUCTION

In recent decades, great process has been made in developing and launching satellites, which makes it easily to assess high spatial resolution remote sensing images. Fine spatial optical sensors with metric or sub-metric resolution, such as QuickBird, Ikonos, Worldview, as well as Chinese Beijing series, allow more detailed and accurate information extraction task. Data from these sensors enable advanced applications, such as urban mapping, precision agriculture, environmental monitoring, and military applications (Verpoorter et al. 2014 and Guan et al. 2017).

Among these applications, image classification is one of the most vital phases for remote sensing image information extraction. Generally, supervised image classification relies on abundant and high-precision training data, which is often manually interpreted by human experts to provide ground truth for training and evaluating the performance of the classifier (Huang et al. 2015). In addition, in most case, training samples labelling takes lots of time and the labelled samples do not work when images changed. In this sense, providing exhaustive ground truth for large remote sensing images is often not possible. Hence, there is an urgent demand to develop a time-saving and accurate sample labelling framework.

In this context, researchers have proposed semi-supervised classification to deal with the insufficient training samples, by taking the unlabelled samples into consideration (Camps-Valls et al. 2007). Meanwhile, active learning has received increasing attention in recent years, which aims to minimize the cost of training sample labelling process (Demir et al. 2011, Di et al. 2012, Patra et al. 2014, Persello et al. 2011 and Persello et al. 2012). Both semi-supervised classification and active learning can work with few training samples, however, it also needs to

develop specific classifier. Thus, semi-supervised classification and active learning are super solutions for scientific study when there are limited training samples. However, for actual production project in remote sensing enterprises, it seems unrealistic to develop a new classification algorithm. Taking advantage of the existing data and traditional classifiers is vital import for timely and efficient actual production project of remote sensing enterprises. In the current literature, there exist studies exploiting the crowdsourced OpenStreetMap (OSM) data as training samples for high-resolution remote sensing classification, where OSM data seems a time-saving and cost-effective way to provide labelled data for image classification (Arsanjani et al. 2013 and Johnson et al. 2016). However, due to the unprofessional production process and the absence of data quality control, OSM data could contain misleading errors. To the best of our knowledge, there have been few papers discussing insufficient labelling in actual production project.

To date, remote sensing enterprises accumulated lots of manually interpreted products from early lower-spatial resolution remote sensing images by executing their routine research and business programs. These interpreted products describe the land cover by assigning each patch a class label, while each patch may contain multiple land cover classes. Inspired by the exploitation of OSM data as training samples in classification, our work attempt to purify the interpreted products to competent of training samples for remote sensing classification (Verpoorter et al. 2016). Due to the actual production needs, the interpreted products may contain misleading errors where a patch may contain multiple classes. Moreover, these manually interpreted products may not match the VHR image properly because of different dates or spatial resolution of both data, thus, hindering suitability of manually interpreted products in training classification models, or small coverage area of these manually interpreted products. We also

* Corresponding author

face similar problems in our laboratory in 21st Century Aerospace Technology Co. Ltd (short for 21AT).

In this work, we propose a method to purify the interpreted product to match newly available VHR data and provide the best training data for supervised image classifiers in VHR classification. Specifically, interpreted products are processed based on spectral feature and superpixel segmentation to generate training samples. The remainder of this paper is constructed as follows. Section II describes the study sites and data sets. Section III presents the methodology of our work. Section IV describes the experimental results and analysis. Finally, section V concludes our work.

2. DATASETS

As shown in Figure 1, a high-resolution RGB image obtained by Beijing-2 satellite over Haidian District in Beijing, China, is utilized in the experiment. The RGB image was pre-processed by stitching and even colour in terms of the specific requirements of actual production project. The Beijing-2 is a satellite constellation, which has three satellites each carries a 1m resolution panchromatic sensor and a 4m resolution multi-spectral sensor. More details of Beijing-2 satellite parameter are

list in Table 1. This satellite constellation has been operated by the 21AT since July 10 2015 (Wen et al. 2017). The Twenty First Century Aerospace Technology Co., Ltd. is a Beijing based high-tech enterprise and is the first commercial Earth observation satellite operator and service provider in China. The satellite constellation can provide daily targeting capability anywhere on Earth. The interpreted product is manually interpreted from previously acquired coarser resolution images in 21AT. The interpreted product of Haidian district is shown in Figure 2, which contains three classes: vegetation, built-up area, water.

Parameter	Pan image	Multispectral image
Spectral range	B5:0.45~0.90 μm	B1: 0.45~0.52 μm B2: 0.52~0.59 μm B3:0.63~0.69 μm B4: 0.77~0.89 μm
Spatial resolution	1 m	4 m
Width	23 km	23 km
Revisit cycle	1 day	
Designed life	7 years	

Table 1. Parameters of Beijing-2 satellite

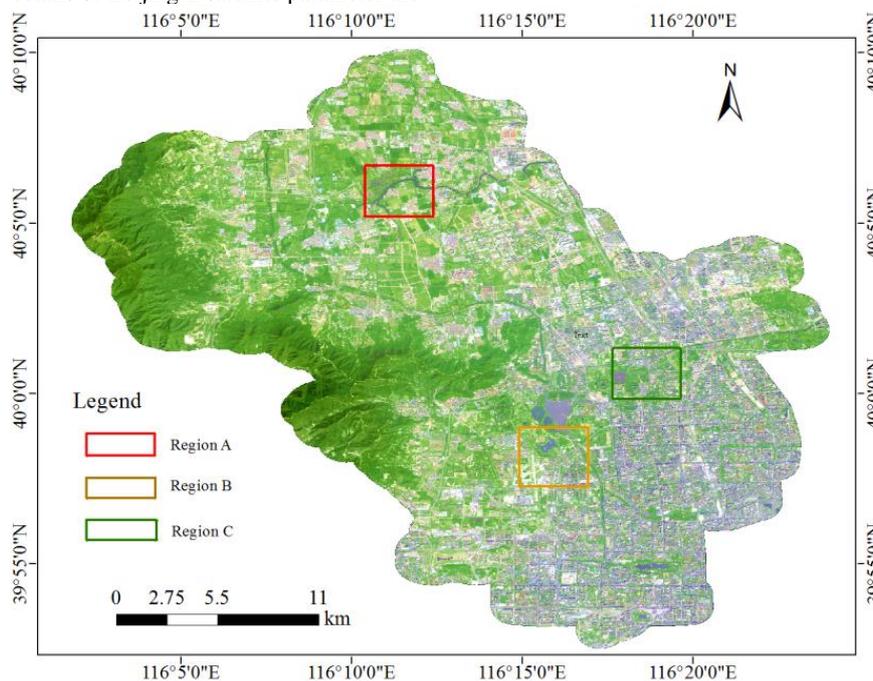


Figure 1. Pre-processed Beijing-2 image over Haidian District with test regions marked as A, B, and C.

In this study, three test regions are selected to validate the proposed framework for the purification of interpreted product according to remote sensing images, with each sub-image covering about a study area of about 4.2 km \times 3.1 km, 4.4 km \times 3.9

km and 4.4 km \times 3.5 km, respectively. And all the processes implemented in our framework are realized based on the MATLAB 2015a.

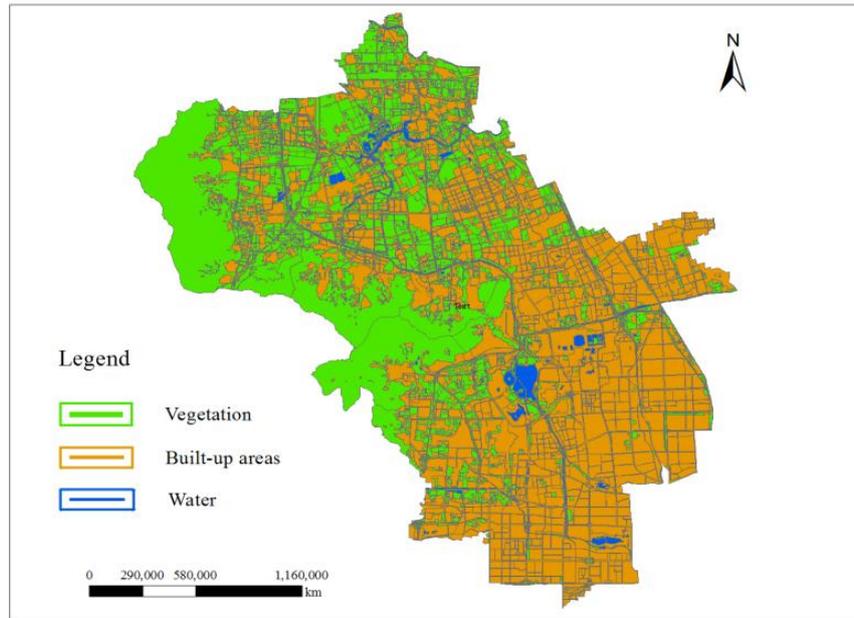


Figure 2. The corresponding interpreted product of Haidian District

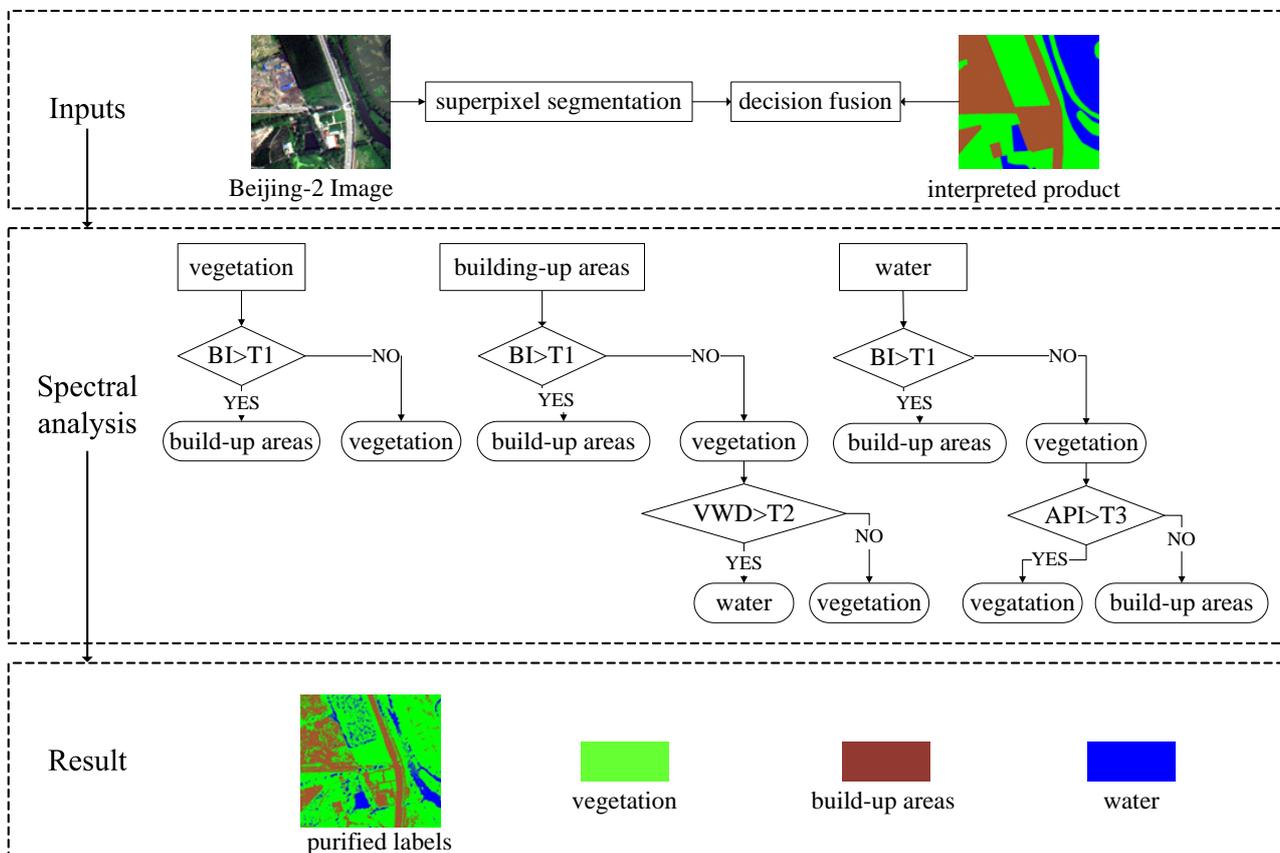


Figure 3. The Graphical Representation of the Overall Framework in This Article

3. METHODOLOGY

3.1 The overall framework

As it was described before, the interpreted product may be generated from remote sensing image of different dates or spatial

resolution. Spatial resolution difference leads to inconsistent of object boundaries in the image and the land cover changes from time to time in the high-resolution image. Consequently, it is not appropriate to utilize the interpreted product directly as training

samples for high resolution image classification. In order to reduce the misleading mistakes introduced by the coarse interpreted product, it is very crucial to match the interpreted product accurately with the high resolution image to be classified. Therefore, our work proposed a purification framework, which exploits the interpreted product to generate training samples for high resolution classification, as illustrated in Fig. 3. It includes the following two points:

- (1) Superpixel segmentation and decision fusion between interpreted product and high resolution remote sensing images.
- (2) Reassign the class label of each pixel in the interpreted product according to three designed spectral analysis indexes.

3.2 Superpixel segmentation

To relief the boundary offset between the interpreted product and the high spatial resolution image, our study employs superpixel segmentation to decline the boundaries according to the high spatial resolution image. Firstly, a superpixel segmentation algorithm is applied on the RGB image, dividing the image into a series of superpixels. In this work, simple linear iterative

clustering superpixels(SLIC) (Achanta et al. 2010) is employed. Supposing an image is to be divided into k superpixels. Pixels are clustered according to their colour similarity and proximity in the image plane. Besides, two parameters, number of desired superpixels and weighting factor between colour and spatial differences, need to be defined when applying the SLIC segmentation on the image. Subsequently, the class boundary of interpreted product is refined according to the superpixel segmentation image by decision fusion. In addition, the decision fusion is realized by the majority voting according to the following rule:

$$x = x(\text{argmax}_s(\text{Vote}(s))) \quad (1)$$

where x represents the class label for each pixel of the interpreted products, and s the superpixel to which the pixel belongs. This step aims to eliminate misleading pixels, which refer to the boundary offsets of the same object revealed in the interpreted product and remote sensing image.

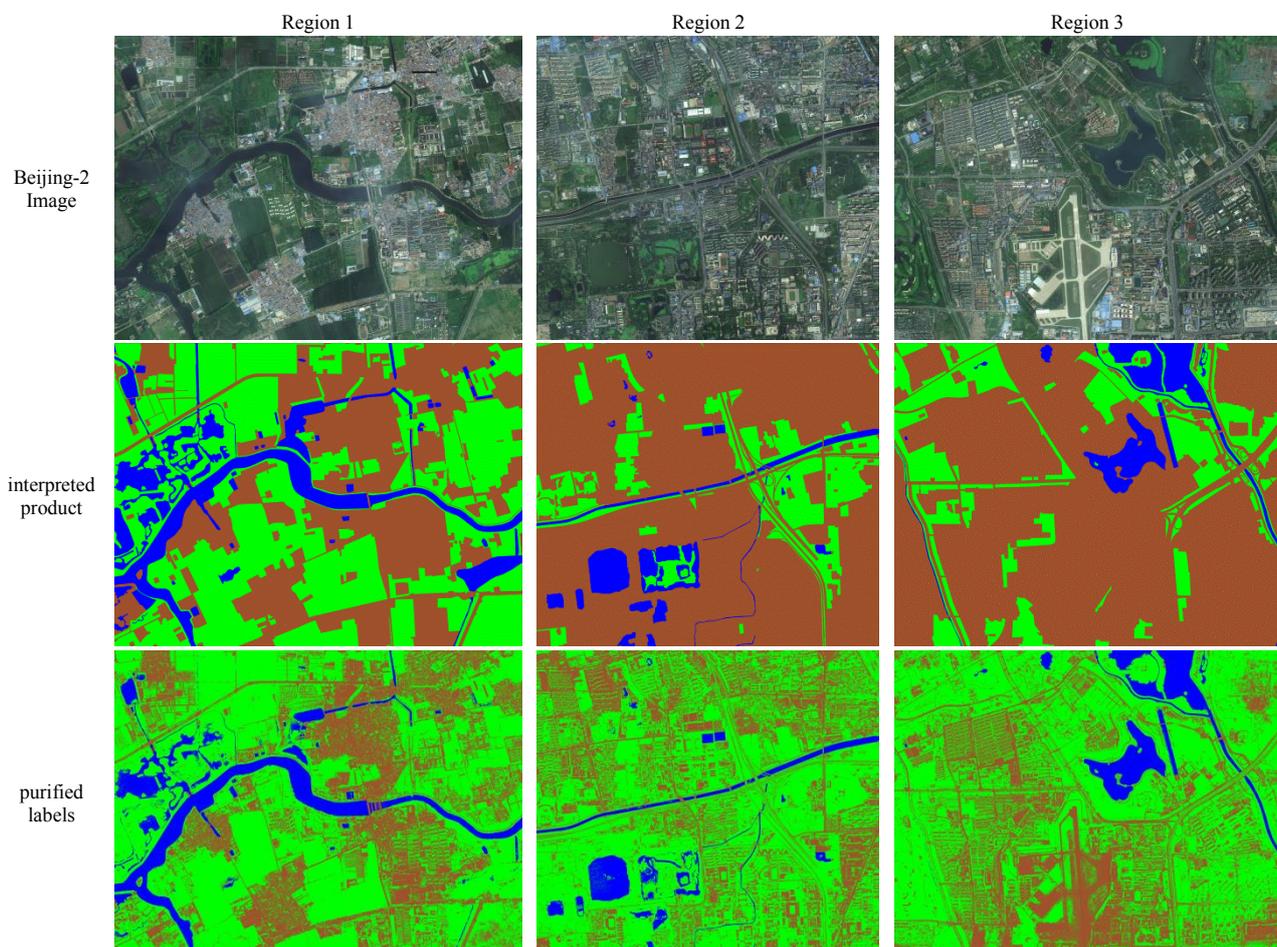


Figure 4. Test regions and the corresponding interpreted product and purified labels map

3.3 Spectral analysis

Apart from boundary offset between interpreted product and VHR image, label errors caused by image difference are misleading. In our study, a series of spectral index based on statistical analysis is proposed to complete the interpreted product according to the VHR image, namely, brightness index,

vegetation-water difference and aquatic plants index. It can be expressed as follows:

- (1) Brightness Index (BI)

$$BI = R + G + B \quad (2)$$

where R , G and B represent the pixel values of the pixel in band red, green and blue on the image, respectively. As the label of

interpreted product is patch based, a patch for one individual class may contain pixels of other class in the VHR image. Due to the significant difference of brightness, this index can effectively distinguish bright and dark class. It can be used to remove vegetation from built-up areas, or built-up areas from vegetation.

$$VWD=B-G \quad (3)$$

Dark vegetation shows similar spectral characteristics with water in Beijing-2 in our study, it is important to ensure attribute correctness of label which to be training samples for classification. To exclude misleading labels between dark vegetation and water, the VWD is presented to verify the class of interpreted product is matched with VHR image.

(3) Aquatic Plants Index(API)

$$API=2*G-(R+B) \quad (4)$$

In the interpreted product, the water surface which is covered with aquatic plants may be marked as water. In fact, aquatic plants belong to vegetation according to spectral characteristics. Therefore, it is necessary to filter out the aquatic plants from water to make sure the best separability of vegetation and water. It is found in our study that the API can discriminate the aquatic plants from water, effectively. By applying the above index-based label refinement on the interpreted product, labels with uncertainty are corrected to the more appropriate class.

	Region1		Region2		Region3	
	interpreted product	purified labels	interpreted product	purified labels	interpreted product	purified labels
vegetation	27.79	19.06	22.22	18.55	25.86	19.25
Build-up	43.23	33.71	40.81	33.54	43.94	37.17
water	20.88	13.07	23.02	18.63	19.24	15.64

Table 2. The intra-class purity for the interpreted product and purified labels of three test images

	interpreted product						purified labels					
	JM distance			Transformed Divergence			JM distance			Transformed Divergence		
	vegetation	build-up	water	vegetation	build-up	water	vegetation	build-up	water	vegetation	build-up	water
vegetation	—	0.83	0.52	—	1.03	0.60	—	1.26	1.49	—	1.55	1.91
build-up	0.83	—	1.26	1.03	—	1.77	1.26	—	1.87	1.55	—	1.99
water	0.53	1.26	—	0.60	1.77	—	1.49	1.87	—	1.91	1.99	—

Table 3. The inter-class separability for Region 1

	interpreted product						purified labels					
	JM distance			Transformed Divergence			JM distance			Transformed Divergence		
	vegetation	build-up	water	vegetation	build-up	water	vegetation	build-up	water	vegetation	build-up	water
vegetation	—	0.75	0.33	—	0.98	0.40	—	1.15	1.39	—	1.41	1.79
build-up	0.75	—	1.04	0.98	—	1.45	1.16	—	1.87	1.41	—	1.99
water	0.33	1.04	—	0.40	1.45	—	1.39	1.87	—	1.79	1.99	—

Table 4. The inter-class separability for Region 2

	interpreted product						purified labels					
	JM distance			Transformed Divergence			JM distance			Transformed Divergence		
	vegetation	build-up	water	vegetation	build-up	water	vegetation	build-up	water	vegetation	build-up	water
vegetation	—	0.59	0.85	—	0.75	0.99	—	1.41	1.07	—	1.78	1.33
build-up	0.59	—	1.21	0.75	—	1.82	1.41	—	1.75	1.78	—	1.99
water	0.85	1.21	—	0.99	1.82	—	1.07	1.75	—	1.33	1.99	—

Table 5. The inter-class separability for Region 3

4. EXPERIMENTS RESULTS AND ANALYSIS

Three test images are utilized to validate the proposed method for the automatic training samples labelling framework. In the experiments, the labels purified by the proposed method are compared with the manually labelled interpreted product according to intra-class purity and inter-class separability. The test regions with the corresponding interpreted product and purified labels are displayed in Figure 4. As it can be seen from Figure 4, class labels are more accurate after purification, which are more qualified to be training samples for pixel-based classification. Water surface which are covered with aquatic plants are all fixed to vegetation. In addition, vegetation around buildings is excluded from build-up areas, where in the interpreted product they are confusing. Besides, bare soils that are caused by image acquisition time are adjusted to the Beijing-2 image.

Follow-up is about the quantitative evaluation of the experimental results. In our study, intra-class variance is adopted to describe the intra-class purity, as variance is an important indicator of data dispersion. The larger value of the variance, the more significant of the volatility it is. In other words, if the variance decreased after purification, the intra-class purity is increased. The intra-class purity for the interpreted product and purified labels are listed in Table 2. It is evident that intra-class purity increased significantly after purification for all the three test regions. Besides, JM distance and transformed divergence are calculated to measure the intra-class separability. These values range from 0 to 2.0 and indicate how well the selected training samples are statistically separate. A larger value for JM distance and transformed divergence represent higher intra-class separability. Table 3, Table 4 and Table 5 exhibit the inter-class separability of interpreted product and purified labels for the three test regions. As can be learned from these tables, inter-class

separability for the interpreted product is rather poor, with JM distance and transformed divergence for vegetation-build-up areas and vegetation-water is lower than 1.0. However, after purification, the inter-class separability is improved significantly, as all the values are raised to above 1. Besides, the transformed

divergence for water-building is even increased to 1.99, which implying these two class are statistically separate.

5. CONCLUSIONS

In this paper, an innovative method that exploits interpreted product and remote sensing image for the generation of training samples is proposed. To make the interpreted product pure enough to be training samples of high resolution image classification, a series of approaches is used successively. To verify the effectiveness of our proposed purification framework, three test images and their corresponding interpreted product are utilized to generate training samples for high resolution image classification. Meanwhile, intra-class purity and inter-class separability are employed to evaluate the quality of the purified training samples. The experimental results illustrate the superiority of the proposed method in terms of quantitative accuracy and visual interpretation. Further research lies in the utilization of the purified training samples for high resolution image classification.

ACKNOWLEDGEMENTS (OPTIONAL)

This research was supported by the National Key R&D Program of China (NO. 2017YFC0505701).

REFERENCES

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S., 2010. Slic superpixels (No. EPFL-REPORT-149300).

Arsanjani, J. J., Helbich, M., & Bakillah, M., 2013. Exploiting volunteered geographic information to ease land use mapping of an urban landscape. *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, London, UK, pp. 29-31.

Camps-Valls, G., Marsheva, T. V. B., & Zhou, D., 2007. Semi-supervised graph-based hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10), pp. 3044-3054.

Demir, B., Persello, C., & Bruzzone, L., 2011. Batch-mode active-learning methods for the interactive classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(3), pp. 1014-1031.

Di, W., & Crawford, M. M., 2012. View generation for multiview maximum disagreement based active learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 50(5), pp. 1942-1954.

Guan, X., Liao, S., Bai, J., Wang, F., Li, Z., Wen, Q., ... & Chen, T., 2017. Urban land-use classification by combining high-resolution optical and long-wave infrared images. *Geo-spatial Information Science*, 20(4), pp. 299-308.

Huang, X., Weng, C., Lu, Q., Feng, T., & Zhang, L., 2015. Automatic labelling and selection of training samples for high-resolution remote sensing image classification over urban areas. *Remote Sensing*, 7(12), pp. 16024-16044.

Johnson, B. A., & Iizuka, K., 2016. Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping: Case study of the Laguna de Bay area of the Philippines. *Applied Geography*, 67, pp. 140-149.

Mitra, P., Shankar, B. U., & Pal, S. K., 2004. Segmentation of multispectral remote sensing images using active support vector machines. *Pattern recognition letters*, 25(9), pp. 1067-1074.

Patra, S., & Bruzzone, L., 2014. A novel SOM-SVM-based active learning technique for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 52(11), pp. 6899-6910.

Persello, C., & Bruzzone, L., 2012. Active learning for domain adaptation in the supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11), pp. 4468-4483.

Persello, C., Boularias, A., Dalponte, M., Gobakken, T., Naeset, E., & Schoelkopf, B., 2014. Cost-sensitive active learning with lookahead: Optimizing field surveys for remote sensing data classification. *IEEE Transactions on Geoscience and Remote Sensing*, 52(10), pp. 6652-6664.

Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J., 2014. A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41(18), pp. 6396-6402.

Wen, Q., He, J., Guan, S., Chen, T., Hu, Y., Wu, W., ... & Yeong, S., 2017. The TripleSat constellation: a new geospatial data service model. *Geo-spatial Information Science*, 20(2), pp. 163-173.