A Novel Wrapper Approach for Feature Selection in Object-Based Image Classification Using Polygon-Based Cross-Validation

Lei Ma, Manchun Li, Yu Gao, Tan Chen, Xiaoxue Ma, and Lean Qu

***Abstract*— Feature selection is becoming a major component of object-based classification as numerous features of segmented object become available. Although common feature selection methods in object-based classification are acknowledged, wrapper-based methods remain an issue due to the diversity of accuracy assessment methods. This letter presents a new wrapper approach using polygon-based cross validation (CV) to overcome possible bias of object-based accuracy assessment for object-based classification. The new method is a two-step wrapper-based feature selection that involves the integration of: 1) feature importance rank using gain ratio and 2) feature subset evaluation using a polygon-based tenfold CV within a support vector machine (SVM) classifier. Several high-resolution images, including both unmanned aerial vehicle images and ISPRS (International Society for Photogrammetry and Remote Sensing) benchmark test data, were used to test the proposed method. Results show that, with the proposed polygon-based CV SVM wrapper, the mean overall accuracy is significantly higher than with an object-based CV SVM wrapper. Furthermore, the proposed method shows potential for comprehensively considering all types of features instead of only spectral features.**

***Index Terms*— Cross validation (CV), feature selection, object-based classification, support vector machine (SVM), wrapper.**

I. Introduction

GIVEN the current availability of a large amount of object features, such as spectral, geometry, and texture, feature selection is an important step for object-based classification for avoiding time-consuming of feature calculation, reducing complexity of classification models, and improving classification performance [1]–[3].

Generally, feature selection methods are divided into wrapper-based and filter-based approaches [4]. Wrapper-based approaches use the classifier's performance as an evaluation criterion for optimizing the feature subset, while filter-based approaches assess the importance of individual features to select features through ordering. Recent research applied feature selection to object-based classification [2], [3], [5], [6], but the focus of these studies was on filter-based methods. For example, Laliberte and Rango [7] applied the GINI index as the splitting rule to rank object features, and Van Coillie *et al*. [5] used genetic algorithms to implement feature selection for object-based classification. Furthermore, Ma *et al.* [3] conducted the correlation-based feature selection method to reduce dimensionality of object features prior to classification. However, wrapper-based approaches have a strong advantage in many remote sensing applications [8]–[10]. For instance, Löw *et al.* [9] conducted a feature selection per pixel for multispectral images by using support vector machines (SVMs) as a fitness function, while Poona *et al*. [10] performed a waveband selection for hyperspectral data using a random forest (RF) classifier as a fitness function. The lack of wrapper-based studies in object-based classification may be because of the large computational load involved in wrapper-based feature selection [8] and the use of polygon-based accuracy assessment, which likely leads to greater computational cost due to complex spatial analysis. Here polygon-based accuracy assessment method calculates the correctly part area for classified object by overlaying the classified layer and referenced layer.

Although Duro *et al*. [11] implemented a wrapper algorithm for object-based classification using the Boruta package in R, where the RF classifier was the fitness function, they used object-based accuracy assessment method instead of a polygon-based method when all algorithms (i.e., search method and fitness function) were directly packaged into an R environment. Here object-based accuracy assessment method refers to a process where to count of correctly classified object. However, the object-based accuracy assessment method, which considers single segmented objects as an entity, seemed unsuitable for object-based classification, since the segmented object is not necessarily represented as only one class because of the possibility of mixed objects [12]–[14]. We assume that such diversity using object-based or polygon-based methods for cross validation (CV) in a wrapper could generate unexpected results for object-based classification. Johnson [15] presented such an issue for land use/land cover classification using multiresolution data, where overestimated classification accuracy was achieved by pixel-based CV, but not using a polygon-based method.

The primary objective of this letter is to test the suitability of polygon-based CV for wrapper feature selection

Manuscript received October 26, 2016; revised December 10, 2016; accepted December 20, 2016. Date of publication January 16, 2017; date of current version February 23, 2017. This work was supported in part by the China Postdoctoral Science Foundation Project under Grant 2016M600392, in part by the Program B for Outstanding Ph.D. Candidate of Nanjing University under Grant 201502B008, and in part by the Special Research Fund of the Ministry of Land and Resources for NonProfit Sector under Grant 201411014-03.

L. Ma, M. Li, Y. Gao, T. Chen, L. Qu are with the Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing 210023, China (e-mail: maleinju@gmail.com; limanchunnju@163.com).

X. Ma is with the Urban and Resources Environmental College, Nanjing Second Normal University, Nanjing 210013, China.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2016.2645710

TABLE I

LIST OF OBJECT FEATURES

Feature Type	Feature Names
Spectral	Mean blue, mean green, mean red, max difference, standard deviation (std. dev.) blue, std. dev. green, std. dev. red, brightness
Texture	GLCM (Gray-Level Co-occurrence Matrix) homogeneity, GLCM contrast, GLCM dissimilarity, GLCM entropy, GLCM std. dev., GLCM correlation, GLCM ang. 2nd moment, GLCM mean, GLDV (Gray-Level Difference Vector) ang. 2nd moment, GLDV entropy, GLDV mean, GLDV contrast
Shape	Area, compactness, density, roundness, main direction, rectangular fit, elliptic fit, asymmetry, border index, shape index

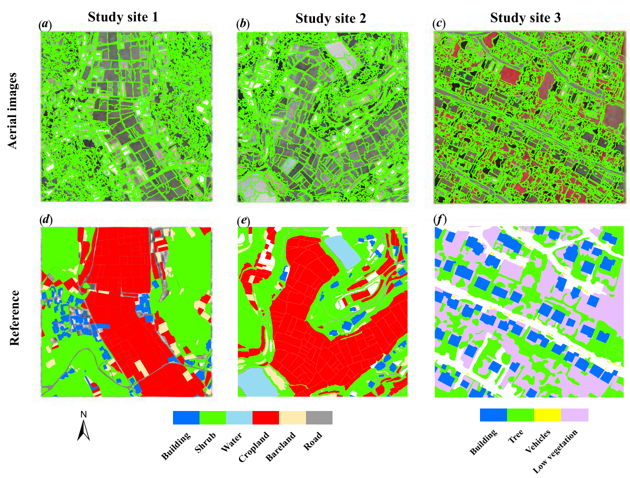

Fig. 1. Three study sites with their segmentation and reference layers. The three RGB bands of (a) and (b) correspond to the red, green, and blue bands (acquired by Canon 5D II), while (c) corresponds to the near infrared, red, and green bands (delivered by digital aerial cameras) (d), (e) and (f) are reference vector layers produced by manual interpretation for (a), (b) and (e) respectively.

in object-based classification. We, therefore, propose a new wrapper strategy for feature selection in object-based image classification to reduce the computational cost of common wrapper approaches. First, a simplified best first search was applied to ranked features using a gain ratio importance score. Second, a polygon-based tenfold CV was used to calculate the mean accuracy of the classifier as a fitness function. Finally, an SVM classifier was employed to test the proposed method.

II. METHODOLOGY

A. Study Sites and Data Sets

Two test sites, characterized by large agricultural environments, were chosen for the experiments. Both sites are located in the southwest basin of China. For this letter, two 500×500 m^2 digital orthophoto maps (DOM) were produced using 0.2-m resolution imageries, which were acquired by an unmanned aerial vehicle at a height of 750 m, and consisted of three visible spectral bands (red, green, and blue) [Fig. 1(a) and (b)]. Manual interpretation produced reference vector layers, which were primarily composed of croplands, woodlands, buildings, bare lands, roads, and water land covers [Fig. 1(d) and (e)]. In addition, an ISPRS benchmark test data set [16] with a 9-cm resolution airborne image of Vaihingen, Germany was used as the third study site (253×230 m^2) in this letter [Fig. 1(c)]. This site covers an urban scene consisting of impervious surfaces, buildings, low vegetation, trees, cars, and clutter [Fig. 1(f)]. For this letter, we ignored the impervious surfaces and clutter.

B. Segmentation and Sampling

For this letter, multiresolution segmentation [17] was applied to three DOMs, respectively. Images were segmented at scale parameter 110 for study site 1, and 130 for study sites 2 and 3, which were determined by combining precision and recall measures based on overlapping regions [18]. Color/shape weight was set to 0.9/0.1, and smoothness/compactness weight was set to 0.5/0.5. All the three bands were used to calculate object features, including spectral, texture, and shape features. The details of the selected features are given in Table I.

To obtain training sample objects, single segmented object was used as the sampling unit, and a stratified random sampling, which uses the same training set ratio (proportion of the segmented objects used for training) for each class, was applied to a list of segmented objects, which were already labeled using an overlapping ratio of 60% with the artificial interpreted reference layer (the stability of such a parameter to other variables was already demonstrated by Li *et al.* [14]).

C. Wrapper Using Polygon-Based Cross Validation

A wrapper method for feature selection uses classification performance as an evaluation criterion to find a feature set with the highest overall accuracy (OA) [19]. However, a typical problem that wrapper methods face in object-based classification is uncertainty in which accuracy assessment method is best to use. In this letter, we proposed a new wrapper methodology that uses polygon-based accuracy assessment method for CV in the process of wrapper-based feature selection (see Fig. 2). The polygon-based accuracy assessment method refers to a process where to calculate the correctly part area of classified object, but not correctly classified object count. Therefore, the new method could overcome possible bias seen in object-based accuracy assessment methods.

First, features were ranked by gain ratio to avoid global searching and reduce computational cost. This was an extension of the information gain measure to overcome the bias of selecting features with numerous values [20]. These features were then sorted into a stacked vector according to their importance score. Subsequently, a supervised classifier (SVM) was applied to training sample objects to evaluate the performance of each feature subset, when polygon-based accuracy assessment method was used to calculate the OA. If the inclusion of additional features improved classification performance, it was retained in the stacked feature vector until all other features were tested.

In order to achieve a stable result, the fitness function (SVM classifier) was evaluated on a tenfold CV using a polygon-based accuracy assessment method. For comparison

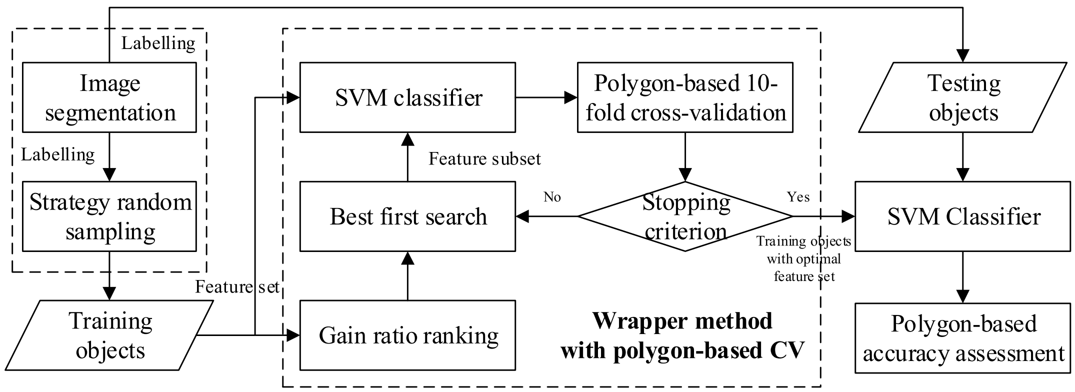

Fig. 2. Framework of proposed methodology.

with our proposed method, the same process was repeated using an object-based accuracy assessment method, which calculated the number of correct classification objects instead of calculating the area of correct classification.

D. Classification and Evaluation

In this letter, it was necessary that a classifier, which could benefit from feature selection on object-based classification, was used to compare both the feature selection methods. According to a previous systematic comparison of several classifiers [14], an SVM classifier was applied to each test site with an identical workflow, in which classification was repeated ten times with independently stratified random sampling (using a stable training set ratio of 30% sampling objects). Subsequently, the SVM classifier with radial basis function kernel was implemented using the R package e1071, and the optimal penalty parameter C and kernel parameter γ were obtained when the best CV accuracy was observed. The same parameters were applied to the SVM in the wrapper feature selection procedure.

In order to perform accuracy evaluation, all polygons in the reference layer were used as validation data. Moreover, a polygon-based accuracy assessment method overlaying the classified object and reference layers was used to calculate the area-based error matrix, which enabled OA to be obtained by calculating the proportion of correct classified area and total area [21]. The function of OA is written as follows:

$$\mathrm{OA} = \sum_{i=1}^{m}\sum_{j=1}^{n}|C_i \cap R_j| / \sum_{j=1}^{n}|R_j| \tag{1}$$

where m is the number of objects in the classified layer, and n is the number of polygons in the reference layer. C_i is the ith classified object, and R_j is the jth reference polygon. $\bigcap$ represents overlapping the classified layer and reference layer. $|*|$ denotes the area in a region.

Object-based classification was repeated ten times to obtain a list of samples of OAs for further evaluating both feature selection methods. Therefore, a two-tailed t-test was employed to check for statistical difference between the OA values of the two feature selection approaches.

III. Results and Discussion

A. Classification Output

First, the proposed polygon-based CV SVM wrapper (Polygon-SVMCV) was evaluated against the object-based CV SVM wrapper (Object-SVMCV) by summarizing visual box plots (Fig. 3), which illustrate median, minimum, and maximum of classification OA samples of both the methods. For each study site, SVM classification with Polygon-SVMCV clearly outperformed that with Object-SVMCV. This is attributed to the inherent uncertainty in object-based CV in object-based classification, since it strongly depends on the number of the homogeneous objects under accuracy assessment.

TABLE II
Statistical Evaluation of Three Study Sites Using SVM-Wrapper and SVM Classification

	Mean OA with Object-SVMCV	Mean OA with Polygon-SVMCV	p
Study site 1	0.841	0.865	<0.01
Study site 2	0.801	0.833	<0.01
Study site 3	0.668	0.729	<0.01

Furthermore, two-tailed t-test results (Table II) shows that the proposed feature selection method performed significantly well for all the three study sites. Table II shows significantly higher mean OA with the polygon-driven Polygon-SVMCV feature selection, than with Object-SVMCV feature selection ($p < 0.01$ at confidence level 95%). These results demonstrate that Polygon-SVMCV significantly improves object-based classification performance, while an opposite effect was observed in Object-SVMCV.

B. Feature Selection Output

Selection frequency of each feature for ten classifications was summarized for site 1 in Fig. 4 to provide insight into both feature selection methods. Fig. 4 shows that Object-SVMCV preferred to retain an object's spectral features (selection frequency of only three spectral features was more than six times, and all texture and shape features were never selected more than four times), while Polygon-SVMCV had no bias for any of the three feature types tested in this letter (each feature type contained at least one feature, which was selected more than six times), even though there were low importance scores for shape features. We assume that the Object-SVMCV with underestimated classification performance preferred to choose features within higher importance scores. Subsequently, the use of Object-SVMCV method led to low accuracy due to the lack of texture and shape features. This reinforces that Polygon-SVMCV is a more appropriate

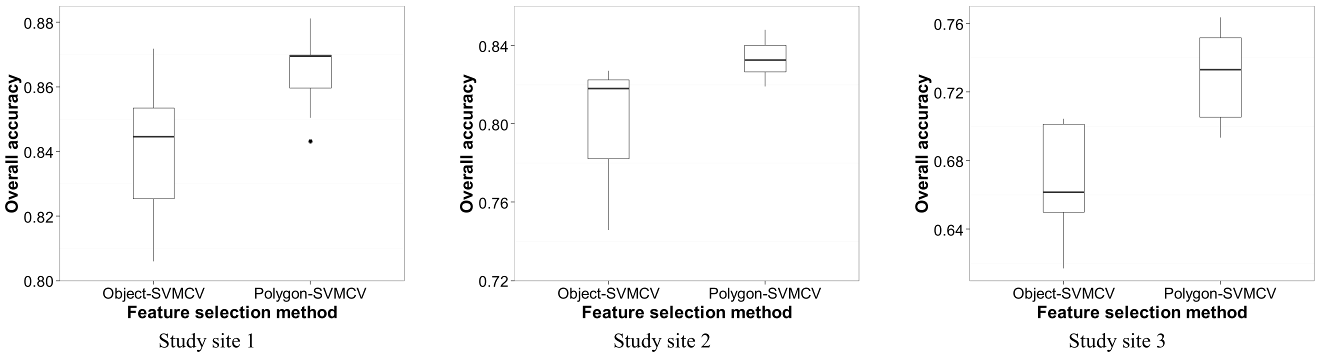

Fig. 3. Comparison between classification performances with Polygon-SVMCV or Object-SVMCV.

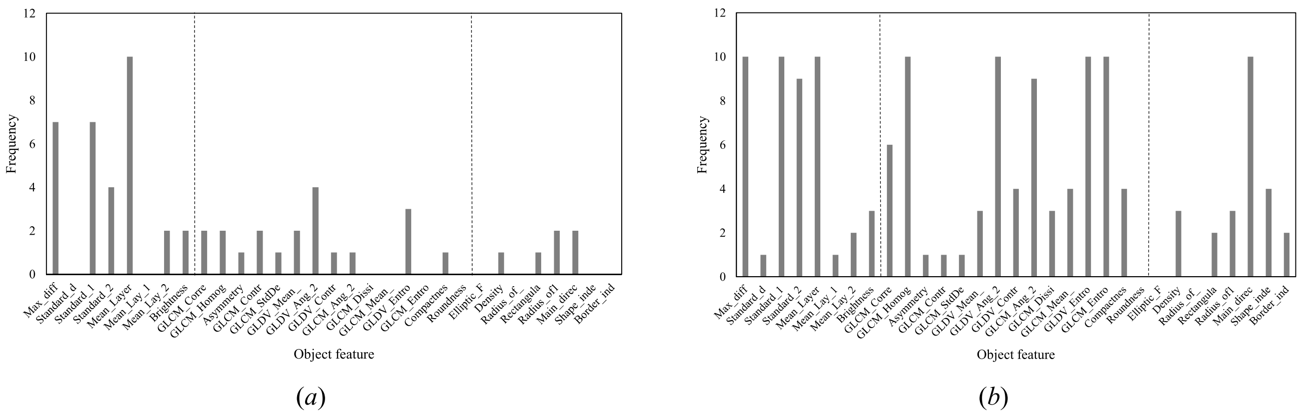

Fig. 4. Selection frequency of feature by (a) Object-SVMCV and (b) Polygon-SVMCV using ten times independent sampling for study site 1. The black vertical dotted lines distinguish three types of features. The left first part is spectral features, the second part is texture features, and the last part is shape features.

feature selection method, a conclusion that is further supported by multiple studies suggesting that texture or shape features of segmented object improve object-based classification performance [7], [22].

IV. Conclusion

The ultimate objective of this letter is to evaluate wrapper feature selection combined with polygon-based CV in object-based classification. Statistical results of classification accuracies demonstrate that the proposed Polygon-SVMCV wrapper feature selection method is valuable for object-based classification. Furthermore, the proposed method could equally select three types of features without any bias, while the Object-SVMCV method focused more on specific spectral or texture features.

References

[1] U. C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen, "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, nos. 3–4, pp. 239–258, Jan. 2004.

[2] A. S. Laliberte, D. M. Browning, and A. Rango, "A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 15, pp. 70–78, Apr. 2012.

[3] L. Ma, L. Cheng, M. Li, Y. Liu, and X. Ma, "Training set size, scale, and features in geographic object-based image analysis of very high resolution unmanned aerial vehicle imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 102, pp. 14–27, Apr. 2015.

[4] D. Rodrigues *et al.*, "A wrapper approach for feature selection based on bat algorithm and optimum-path forest," *Expert Syst. Appl.*, vol. 41, no. 5, pp. 2250–2258, Apr. 2014.

[5] F. M. B. van Coillie, L. P. C. Verbeke, and R. R. De Wulf, "Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium," *Remote Sens. Environ.*, vol. 110, no. 4, pp. 476–487, Oct. 2007.

[6] T. Novack, T. Esch, H. Kux, and U. Stilla, "Machine learning comparison between WorldView-2 and QuickBird-2-simulated imagery regarding object-based urban land cover classification," *Remote Sens.*, vol. 3, no. 10, pp. 2263–2282, Oct. 2011.

[7] A. S. Laliberte and A. Rango, "Texture and scale in object-based analysis of subdecimeter resolution unmanned aerial vehicle (UAV) imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 761–770, Mar. 2009.

[8] M. Pal, "Hybrid genetic algorithm for feature selection with hyperspectral data," *Remote Sens. Lett.*, vol. 4, no. 7, pp. 619–628, Jul. 2013.

[9] F. Löw, U. Michel, S. Dech, and C. Conrad, "Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines," *ISPRS J. Photogramm. Remote Sens.*, vol. 85, pp. 102–119, Nov. 2013.

[10] N. K. Poona, A. van Niekerk, R. L. Nadel, and R. Ismail, "Random forest (RF) wrappers for waveband selection and classification of hyperspectral data," *Appl. Spectrosc.*, vol. 70, no. 2, pp. 322–333, Feb. 2016.

[11] D. C. Duro, S. E. Franklin, and M. G. Dub, "Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests," *Int. J. Remote Sens.*, vol. 33, no. 14, pp. 4502–4526, Jul. 2012.

[12] T. G. Whiteside, S. W. Maier, and G. S. Boggs, "Area-based and location-based validation of classified image objects," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 28, pp. 117–130, May 2014.

[13] J. Radoux and P. Bogaert, "Accounting for the area of polygon sampling units for the prediction of primary accuracy assessment indices," *Remote Sens. Environ.*, vol. 142, pp. 9–19, Feb. 2014.

[14] M. Li, L. Ma, T. Blaschke, L. Cheng, and D. Tiede, "A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 49, pp. 87–98, Jul. 2016.

[15] B. A. Johnson, "Scale issues related to the accuracy assessment of land use/land cover maps produced using multi-resolution data: Comments on 'The improvement of land cover classification by thermal remote sensing,"' *Remote Sens.*, vol. 7, no. 10, pp. 13436–13439, Oct. 2015.

[16] F. Rottensteiner, G. Sohn, M. Gerke, J. D. Wegner, U. Breitkopf, and J. Jung, "Results of the ISPRS benchmark on urban object detection and 3D building reconstruction," *ISPRS J. Photogramm. Remote Sens.*, vol. 93, pp. 256–271, Jul. 2014.

[17] M. Baatz and A. Schäpe, "Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation," in *Angewandte Geographische Informations-Verarbeitung*, J. Strobl, T. Blaschke, and G. Griesebner, Eds. Karlsruhe, Germany: Wichmann Verlag, 2000, pp. 12–23.

[18] X. Zhang, X. Feng, P. Xiao, G. He, and L. Zhu, "Segmentation quality evaluation using region-based precision and recall measures for remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 102, pp. 73–84, Apr. 2015.

[19] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Elect. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014.

[20] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. San Mateo, CA, USA: Morgan Kaufman, 2011.

[21] R. G. Pontius, Jr., and M. Millones, "Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment," *Int. J. Remote Sens.*, vol. 32, no. 15, pp. 4407–4429, Aug. 2011.

[22] M. Kim, T. A. Warner, M. Madden, and D. S. Atkinson, "Multiscale GEOBIA with very high spatial resolution digital aerial imagery: Scale, texture and image objects," *Int. J. Remote Sens.*, vol. 32, no. 10, pp. 2825–2850, May 2011.