A Novel Image Classification Algorithm Using Overcomplete Wavelet Transforms

Soe W. Myint, Tong Zhu, and Baojuan Zheng

***Abstract*—A novel frequency-based classification framework and new wavelet algorithm (Wave-CLASS) is proposed using an overcomplete decomposition procedure. This approach omits the downsampling procedure and produces four-texture information with the same dimension of the original image or window at infinite scale. Three image subsets of QuickBird data (i.e., park, commercial, and rural) over a central region in the city of Phoenix were used to examine the effectiveness of the new wavelet overcomplete algorithm in comparison with a widely used classical approach (i.e., maximum likelihood). While the maximum-likelihood classifier produced < 78.29% overall accuracies for all three image subsets, the Wave-CLASS algorithm achieved high overall accuracies—95.05% for the commercial subset (Kappa $= 0.94$), 93.71% for the park subset (Kappa $= 0.93$), and 89.33% for the rural subset (Kappa $= 0.86$). Results from this study demonstrate that the proposed method is effective in identifying detailed urban land cover types in high spatial resolution data.**

***Index Terms*—Classification, high spatial resolution, infinite scale, overcomplete decomposition, urban land cover, wavelet transforms.**

I. Introduction

REMOTELY sensed imaging instruments are capable of collecting high spatial resolution data across multiple wavelength channels for a large geographic area. Advances in remote sensing technologies have enhanced our ability to identify land use and land cover (LULC) at local, regional, and global scales. Developing advanced classification techniques to improve classification accuracy and to automate classification procedures on remotely sensed images is indispensable for LULC assessment purposes.

Conventional digital classification techniques employ spectral responses from land cover types based on single pixel values without capturing spatial information. Several attempts, such as fractal, spatial autocorrelation, and spatial co-occurrence matrices, have been made to improve the spectral analysis of remotely sensed data [1]. Most of them primarily focus on coupling between features and objects at single scale and cannot effectively determine representative values of texture features according to their directionality, spatial arrangements, variations, edges, contrasts, and the repetitive nature of features. These methods do not provide satisfactory accuracy when dealing with complex spatial features and many different land cover types that share similar spatial pattern.

In recent years, wavelet transform has been investigated and applied in image processing due to its innovative mathematical framework for multiscale time–frequency signal analysis. Recent literature on image processing and analysis with wavelet transform generally focus on image compression [2], image fusion [3], image watermarking [4], face image detection [5], image noise removal [6], identification of tumorous regions and microcalcification clusters [7], image segmentation [8], etc. However, image classification, a type of categorization of image data using spectral, spatial, and temporal information that assigns the pixels of a continuous raster image to discrete categories concerning wavelet transforms, has not been well explored. Wavelet transform opens the possibility to capture image features at different scales, which have been shown to be more discriminative than the original spectral feature.

In this letter, a novel frequency-based classification framework and algorithm is proposed using an overcomplete decomposition procedure at multiple scales to examine if the proposed method can effectively identify detailed urban land cover types in high spatial resolution data (i.e., QuickBird). Since overcomplete wavelet can generate spatial arrangements of objects and features at any scale level (infinite-scale analysis), the proposed frequency-based multiscale classification algorithm using overcomplete wavelet is superior to the dyadic-based wavelet classifier approach introduced by Myint [9]. Furthermore, because the approach has the flexibility to use any window size, it is applicable to any LULC systems at any level and at either high spatial resolution (e.g., QuickBird and IKONOS) or medium-resolution satellite data (e.g., Landsat and MODIS). The newly developed algorithm is hereafter referred to as Wave-CLASS.

II. Overcomplete Wavelet Transform

Mallat initiated the multiresolution analysis theory using the orthonormal wavelet basis [10]. Wavelet decompositions can be categorized into dyadic and overcomplete approaches. The overcomplete decomposition approach employed in this study omits the downsampling procedure and produces four-texture information with the same dimension of the original image at an infinite scale. A set of feature vectors is used to identify detailed land cover classes. In the dyadic wavelet transform, the filtered

Manuscript received August 4, 2014; revised October 20, 2014; accepted November 5, 2014. This work was supported by the National Science Foundation through the project "Wavelet Analysis of High Spatial Resolution Imagery for Urban Mapping Using Infinite Scale Decomposition Techniques" under Grant 1154904.

S. W. Myint and B. Zheng are with the School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287 USA (e-mail: soe.myint@asu.edu; bzheng11@asu.edu).

T. Zhu is with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287 USA (e-mail: zhu@asu.edu).

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

Digital Object Identifier 10.1109/LGRS.2015.2390133

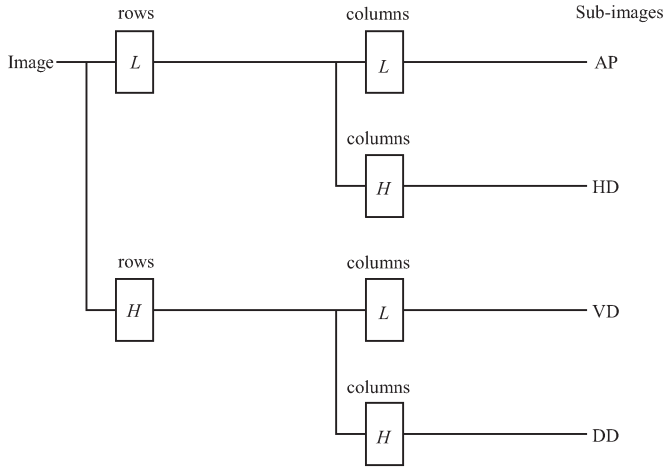

Fig. 1. Overcomplete wavelet decomposition: L = lowpass filter; H = highpass filter; AP = approximation subimage; HD = horizontal detail subimage; VD = vertical detail subimage; DD = diagonal detail subimage.

versions of each subimage are downsampled by a factor of 2. The nonlinear nature of the decimator in dyadic decomposition may cause the filtering-decimation operation to become shift invariant [7].

The idea of employing overcomplete wavelet decomposition is motivated by the fact that it can provide translational invariant features [11]. Another advantage of using the overcomplete approach is that it can describe features with a smaller characteristic scale [12] since any size of window can be effectively employed without downsampling (see Fig. 1). This is crucial for classification a higher decomposition level can be expected to improve accuracy.

III. Proposed Multiscale Classification Algorithm

The research design for the wavelet-based overcomplete classification algorithm is shown in Fig. 2.

An entire image, which consists of both unknown blocks (blocks that need to be classified) and sample blocks (training samples), undergoes multiscale overcomplete wavelet transform. Several training samples need to be selected for each LULC class. A spatial measure can be used as the feature to represent unknown blocks and sample blocks. After that, the distance between the feature vectors leads to supervised classification, such as Euclidean distance classifier.

A. Wavelets

One of the Daubechies wavelets (i.e., Daub4) (Strang and Nguyen, 1998) that has coefficients $c(0) = 1 + \sqrt{3},\ c(1) = 3 + \sqrt{3},\ \ c(2) = 3 - \sqrt{3},\ \ c(3) = 1 - \sqrt{3},\ \ d(0) = 1 - \sqrt{3},\ d(1) = -(3 - \sqrt{3}),\ d(2) = 3 + \sqrt{3},$ and $d(3) = -(1 + \sqrt{3})$ multiplied by $1/4\sqrt{2}$ was used in this study. Its dilation equation and wavelet equation can be expressed as

$$\varphi(t) = \varphi(2t) + \varphi(2t-1) \tag{1}$$

$$\psi(t) = \varphi(2t) + \varphi(2t-1). \tag{2}$$

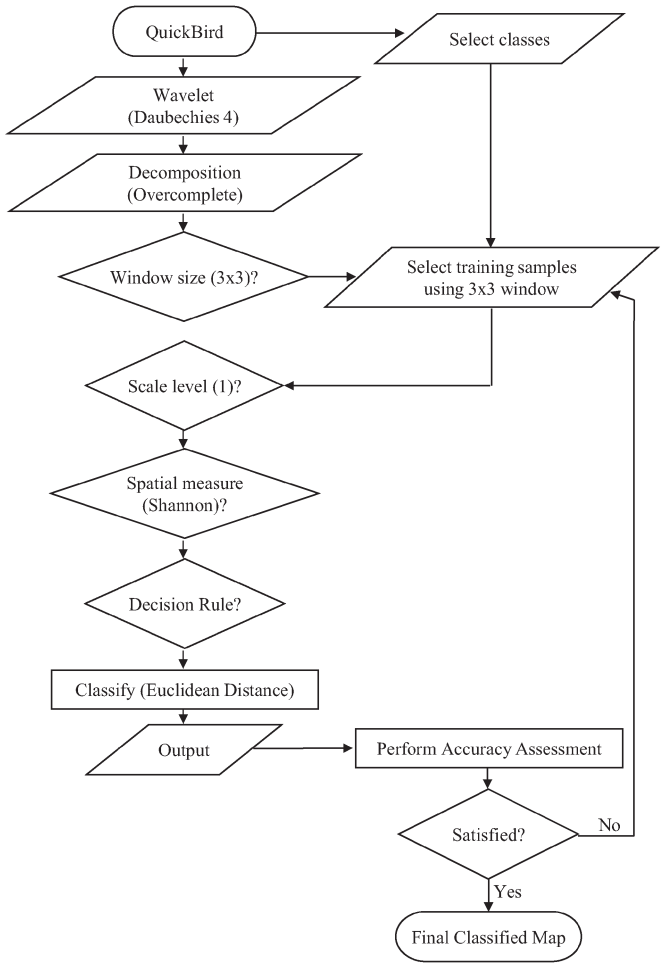

Fig. 2. Proposed framework with overcomplete wavelet decomposition approach (Wave-CLASS).

B. Spatial Measures

In this letter, Shannon index (SHAN) was used in Wave-CLASS to identify the spatial arrangements of the selected LULC classes at different scales, i.e.,

$$\text{SHAN} = -\sum_{i=1}^{M}\sum_{j=1}^{N} c(i,j)^* \log c(i,j)) \tag{3}$$

where $c(i,j)$ is a wavelet coefficient of a subimage with M rows and N columns at locations i and j at one level. Wavelet decomposition and computation of feature vectors are done for all selected training samples before starting the decomposition of the local window at the top left corner of the image and computing the texture measure values of subimages of that local window. If a sample image or a local window of an image is decomposed up to level m, the feature vector used in this study can be described as

$$[f_{LL-1}, f_{LH-1}, f_{HL-1}, f_{HH-1}, \ldots, f_{LH-m}, f_{LH-m}, f_{HL-m}, f_{HH-m}]^T. \tag{4}$$

The distance between the feature vectors of a local window and a set of certain texture samples of their feature vectors lead

to the supervised classification using a classifier. We used the Shannon index to test the effectiveness of the algorithm.

C. Classifier

The performance of a Euclidean distance classifier was evaluated for texture classification using the computed texture feature value of the subimages (decomposed images at different levels). Each pattern class C_k is represented by a prototype pattern P_k. In this letter, P_k $(k = 1, 2, \ldots, 7)$ is a set of total feature vectors of the training samples. The minimum distance classifier assigns an unknown class pattern Q to the class S_k, if the distance R_k between Q and P_k is minimum among all possible class prototypes. The Euclidean distance is defined as

$$R_k = \|Q - P_k\| = \sqrt{\sum_{j=1}^{N}(q_j - p_{k,j})^2} \tag{5}$$

where q_j is the jth feature of unknown class pattern Q, and $p_{k,j}$ is the jth feature of prototype P_k. N is the number of feature vectors.

D. Decision Rule

The classification decision rule in this study computes one distance for each class among all bands; hence, five distances are computed for the classification decision. The training sample with the shortest distance to the unknown feature vector wins the unknown texture. In other words, the training class that is closest to the local window is assigned to the center of that window.

IV. Experimental Approach

A. Data and Study Area

Three image subsets of QuickBird data over a central region in the city of Phoenix acquired on May 29, 2007, were used to examine the effectiveness of the new Wave-CLASS wavelet overcomplete algorithm. One image subset covers commercial and industrial areas with large buildings, roads, and parking lots (see Fig. 3(a); 570×267 pixels). The second image subset includes a park and a small portion of a residential area with plenty of trees, shrubs, and grass (see Fig. 3(b); 334×441 pixels). The third subset covers a rural urban fringe dominated by open soil or desert landscape (see Fig. 3(c); 378×178 pixels). The three image subsets give a diversity of urban classes. The data set has a spatial resolution of 2.4 m, with four channels: blue—B1 (0.45–0.52 μm); green—B2 (0.52–0.60 μm); red—B3 (0.63–0.69 μm); and near infrared—B4 (0.76–0.90 μm). The radiometric resolution of the data set is 16 bit.

B. Classification and Training Sample Selection

We identified seven LULC classes for this study, including buildings, other impervious surfaces (e.g., roads and parking lots), unmanaged exposed soil, trees and shrubs, grass, swimming pool, and other water bodies. Training samples for the preceding classes were selected using local area knowledge, visual identification, and Google Earth. The number of samples per class ranges from 2 to 10, depending on the spectral response of different surface features within a land cover type (e.g., five spectrally different impervious surfaces). Local window sizes of 3×3 centered at the selected homogenous pixels that represent urban classes were used to subset training samples. The number of samples selected for each land cover class is presented in Table I.

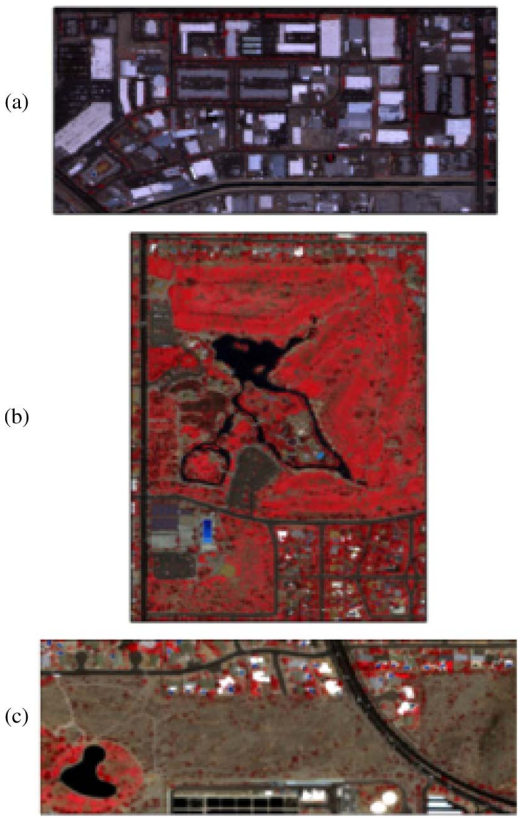

Fig. 3. QuickBird images displaying near-infrared band in red, red band in green, and green band in blue. (a) Commercial. (b) Park. (c) Rural.

TABLE I
Number of Training Samples Selected For Each Land Cover Class

	Commercial	Park	Rural
Building	10	11	13
Soil	14	14	19
Grass	6	15	5
Imp	28	14	18
Pool	3	4	3
Trees	10	26	6
Water	8	10	6
Total	79	94	70

C. Comparison With the Maximum-Likelihood Classifier

For comparison purposes and to better evaluate the effectiveness of the Wave-CLASS algorithm, we also employed maximum-likelihood classifier, one of the most widely used classical techniques, to identify the selected classes. To be

TABLE II
OVERALL ACCURACY, PRODUCER'S ACCURACY, USER'S ACCURACY, AND KAPPA PRODUCED BY WAVE-CLASS AND MAXIMUM LIKELIHOOD (COMMERCIAL SUBSET)

	Wave-CLASS		Maximum Likelihood	
	Pro Acc	Use Acc	Pro Acc	Use Acc
Building	94%	99%	62%	97%
Soil	97%	89%	17%	56%
Grass	85%	93%	72%	29%
Imp	97%	95%	83%	71%
Pool	100%	100%	100%	69%
Trees	91%	93%	67%	63%
Water	98%	100%	80%	100%
	Ove Acc = 95.05%; Kappa = 0.9389		Ove Acc = 65.90%; Kappa = 0. 5861	

TABLE III
OVERALL ACCURACY, PRODUCER'S ACCURACY, USER'S ACCURACY, AND KAPPA PRODUCED BY WAVE-CLASS AND MAXIMUM LIKELIHOOD (PARK SUBSET)

	Wave-CLASS		Maximum Likelihood	
	Pro Acc	Use Acc	Pro Acc	Use Acc
Building	88%	75%	47%	26%
Soil	86%	87%	66%	52%
Grass	98%	99%	70%	98%
Imp	88%	96%	58%	88%
Pool	100%	100%	56%	34%
Trees	99%	95%	81%	75%
Water	95%	100%	32%	100%
	Ove Acc = 93.71%; Kappa = 0.9252		Ove Acc = 61.71%; Kappa = 0. 5502	

TABLE IV
OVERALL ACCURACY, PRODUCER'S ACCURACY, USER'S ACCURACY, AND KAPPA PRODUCED BY WAVE-CLASS AND MAXIMUM LIKELIHOOD (RURAL SUBSET)

	Wave-CLASS		Maximum Likelihood	
	Pro Acc	Use Acc	Pro Acc	Use Acc
Building	83%	71%	87%	55%
Soil	92%	93%	70%	94%
Grass	93%	100%	89%	79%
Imp	85%	81%	55%	89%
Pool	83%	94%	100%	78%
Trees	90%	92%	96%	59%
Water	94%	98%	90%	100%
	Ove Acc = 89.33%; Kappa = 0.8629		Ove Acc = 78.29%; Kappa = 0. 7315	

consistent, we used the same training samples and the same accuracy check points.

D. Accuracy Assessment

The overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient were generated using the error matrix. A minimum of 50 sample points for each LULC category is generally suggested for the accuracy assessment of any image classification [13]. Thus, a total of 525 samples was generated using a stratified random sampling approach with a minimum per-size sample size of 35 points, resulting an average of 75 points per class (a total of seven classes). Tables II–IV show the overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient produced by the Wave-CLASS algorithm in comparison with the maximum-likelihood classifier for the commercial, park, and rural subsets, respectively.

V. RESULTS AND DISCUSSION

A. Wave-CLASS Algorithm

In general, the producer's and user's accuracies for all the categories are about equal for all image subsets, which is one of the key criteria for systematic LULC mapping [14], [15]. The producer's and user's accuracies of all the selected classes in the commercial, park, and rural subsets achieve above 80% accuracy, except the user's accuracy of building class for the park (75%) and rural subsets (71%). User accuracy of 75% indicates that only 75% of the areas identified as building within the classification in the park subset are truly of that category, although the producer's accuracy of the same category reached 88%, which can be considered to be effective. In other words, remote sensing analysis of the algorithm shows that 88% of the time an area that was identified as a building in the park subset was actually a building, whereas a user of the output map would argue that only 75% of the time the classification correctly identifies "building" as actually being a building. In contrast, building class in the commercial subset achieved high producer's accuracy (94%) and user's accuracy (99%).

We anticipated that building class in the park and rural subsets could potentially lower the overall accuracy. The relatively low user's accuracy could be due to the spectral similarity among building rooftops, soil types, and impervious surfaces. Another important factor is that almost all buildings in these subsets are residential-type buildings. These buildings are generally much smaller than those in the commercial subset where most are large office and industrial buildings. Most commercial buildings are bright (white to gray), surrounded by dark impervious surfaces (asphalt roads and parking lots), whereas residential buildings have many different colors of rooftops that are mixed or interlocked with diverse materials or complex background objects and features (e.g., soil, grass, trees, shrubs, pool, streets, and parking lots).

Results for the pool category achieved 100% for both producer's and user's accuracies in both the commercial and park subsets, because the spectral signatures of swimming pools are highly distinguishable from those of other land cover features. The relatively low classification accuracy for pools in the rural subset is likely due to its smaller size in general that inevitably causes mixed pixels in the training samples. The next highest producer's accuracy is 98% for the commercial subset, 99% for the park subset, and 94% for the rural subset.

Since most urban objects and LULC types can easily be identified in high-resolution data (i.e., QuickBird), we also qualitatively assessed the effectiveness of the Wave-CLASS by comparing the output maps and the original satellite data side by side. We found that the outputs match well with the image data (see Fig. 4). As a rule of thumb, the minimum mapping accuracy of 85% is generally required for most mapping activities [15]. We believe that the Wave-CLASS algorithm is effective since it not only produced high overall accuracies for all three image data—95.05% for the commercial subset (Kappa $= 0.94$), 93.71% for the park subset (Kappa $= 0.93$), and 89.33% for the rural subset (Kappa $= 0.86$)—but also generated similar producers' and users' accuracies for all classes.

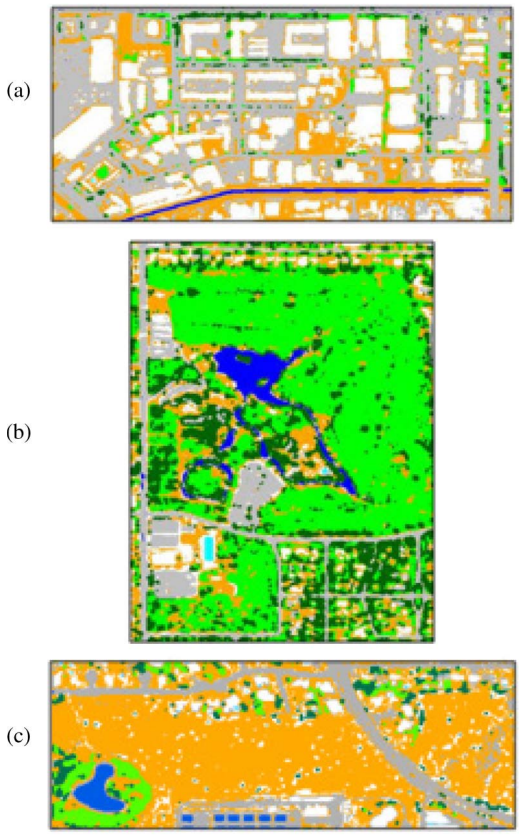

Fig. 4. Output maps generated by Wave-CLASS. (a) Commercial. (b) Park. (c) Rural. Note: building = white; grass = light green; impervious = gray; pool = cyan; soil = orange; trees/shrubs = dark green; water = blue.

B. Comparison With the Maximum-Likelihood Classifier

Overall accuracies and kappa coefficients generated by the maximum-likelihood classifier for the commercial, park, and rural subsets were 65.90% (Kappa = 0.59), 61.71% (Kappa = 0.55), and 78.29% (Kappa = 0.73), respectively, which are significantly lower than those of the Wave-CLASS algorithm. Moreover, the producer's and user's classification accuracies generated by the maximum-likelihood classifier vary substantially, as compared with those of Wave-CLASS. Therefore, the wavelet-based approach produced more accurate classifications and presented better capabilities to generate higher classification accuracies for rare classes than the traditional classifier.

VI. Conclusion

It can be expected that higher spatial resolution imagery will result in higher classification accuracy. However, in reality, higher resolution imagery acquired over urban areas produce lower classification accuracies, because it becomes more difficult to train computers to identify many detailed features and small objects that share similar spectral responses.

Our newly developed Wave-CLASS algorithm produced higher overall accuracies than the maximum-likelihood classifier. In addition, the Wave-CLASS algorithm resulted in similar producers' and users' accuracies for all the selected classes in all images, which demonstrates its robustness for an accurate LULC classification. Results from this study suggest that the Wave-CLASS method is effective in identifying detailed urban land cover types in high spatial resolution data.

References

[1] S. W. Myint, "Fractal approaches in texture analysis and classification of remotely sensed data: Comparisons with spatial autocorrelation techniques and simple descriptive statistics," *Int. J. Remote Sens.*, vol. 24, no. 9, pp. 1925–1947, Jan. 2003.

[2] X. Yu, "Image compression method based on the sub-band similarity of different 191 scales by wavelet transform decomposition," in *Proc. Int. Comput. Congr. Wavelet Anal. Appl., Active Media Technol.*, 2005, vol. 1, pp. 191–197.

[3] M. Xia, Y. He, F. Su, and W. Ouyang, "A novel image fusion algorithm using multiscale edges information," in *Proc. Int. Comput. Congr. Wavelet Anal. Appl., Active Media Technol.*, 2004, vol. 1, pp. 306–311.

[4] Y. Yuan, Y. Ding, and B. Li, "Aimed attacked method in digital image watermarking based on discrete wavelet transform," in *Proc. Int. Comput. Congr. Wavelet Anal. Appl., Active Media Technol.*, 2004, vol. 1, pp. 318–324.

[5] K. He, J. Zhou, D. He, and F. Zhong, "Face detection in color images based on multi-resolution sub-image fusion," in *Proc. Int. Comput. Congr. Wavelet Anal. Appl., Active Media Technol.*, 2005, vol. 1, pp. 222–228.

[6] W. Wang, F. Xing, Y. Dong, and G. Rui, "Despeckling SAR images based on statistically modeling local wavelet coefficients," in *Proc. 6th Int. Congr. Wavelet Anal. Active Media Technol.*, 2005, vol. 1, pp. 255–260.

[7] A. K. Chan, and C. Peng, *Wavelets for Sensing Technologies*. Norwood, MA, USA: Artech House, 2003, p. 237.

[8] L. D. Wang, Z. L. Shi, and S. B. Huang, "Image segmentation method using fractal dimension and wavelet decomposition," in *Proc. 6th Int. Congr. Wavelet Anal. Active Media Technol.*, 2005, vol. 2, pp. 783–787.

[9] S. W. Myint, "Multi-resolution decomposition in relation to characteristic scales and local window sizes using an operational wavelet algorithm," *Int. J. Remote Sens.*, vol. 31, no. 10, pp. 2551–2572, May 2010.

[10] S. G. Mallat, "A theory for multi-resolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, 674–693, Jul. 1989.

[11] S. Fukuda and H. Hirosawa, "Suppression of speckle in synthetic aperture radar images using wavelet," *Int. J. Remote Sens.*, vol. 19, no. 3, pp. 507–519, Jan. 1998.

[12] R. M. Lark, "Geostatistical description of texture on an aerial photograph for discriminating classes of land cover," *Int. J. Remote Sens.*, vol. 17, no. 11, pp. 2115–2133, Jul. 1996.

[13] R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Boca Raton, FL, USA: Lewis Publishers, 1999, p. 137.

[14] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, *Remote Sensing and Image Interpretation*. New York, NY, USA: Wiley, 2008.

[15] J. Anderson, E. E. Hardy, J. T. Roach, and R. E. Witmer, *A Land Use and Land Cover Classification System for Use with Remote Sensor Data, Geological Survey Professional Paper 964*. Washington, DC, USA: U.S. Gov. Printing Off., 1976.