Delineation of Individual Tree Crowns Using High Spatial Resolution Multispectral WorldView-3 Satellite Imagery

Fei Tong, Hengjian Tong, Rakesh Mishra, and Yun Zhang, *Senior Member, IEEE*

***Abstract*—Delineation of individual tree crowns can provide valuable information for sustainable forest management and environmental protection. However, it is hard to find a reliable tree crown delineation method that can continually generate expected results in high spatial resolution multispectral satellite images, because most of the existing methods need user-assigned parameters that greatly affect the quality of the delineation results. In this article, we propose a method based on the marker-controlled watershed segmentation to delineate individual tree crowns using high spatial resolution multispectral WorldView-3 satellite imagery. A gradient binarization process is proposed to accurately locate tree crown borders. The threshold for the binarization is determined by a supervised searching process. Markers used in marker-controlled watershed segmentation are spatial local maxima detected from the information provided by tree crown borders. Moreover, the definition of spatial local maxima from the literature is improved to eliminate false treetops. To validate the performance of the proposed delineation method, delineation results are compared with those obtained from the spectral angle segmentation (SAS) method that has been proposed in literature because the quality of delineation results generated by SAS does not rely on user-assigned parameters. The experiment results in two test images demonstrate that the proposed method outperforms SAS in terms of both delineation accuracy and visual quality of the delineation map. Moreover, it is proved that the modified spatial local maxima are more reliable for detecting treetops.**

***Index Terms*—Gradient binarization, individual tree crown delineation, spatial local maxima, tree crown borders, treetop detection, WorldView-3 satellite imagery.**

I. Introduction

As THE basis of many forest management applications, tree crown delineation is an important function in the field of remote sensing as applied to forestry. In recent years, the delineation of individual tree crowns using remote sensing data has therefore attracted a lot of attention. Various data sources are available for this purpose including multispectral aerial, multispectral satellite, Lidar, and hyperspectral data. A lot of recent studies focus on the use of Lidar data [1]–[8], in which a canopy height model (CHM) can be derived for delineation. However, compared with the passive optical (multispectral/hyperspectral) data, Lidar coverage is small, and the cost of acquisition and processing is high. High spatial resolution multispectral satellite data also have great potential for tree crown delineation [9], [10]. Although it cannot provide precise elevation information, the abundant spectral information from satellites can also provide valuable information for delineation. Therefore, the research described in this article is aimed at providing an effective delineation method with cheaper and more readily available data, and focuses on automatic tree crown delineation using high spatial resolution multispectral WorldView-3 satellite imagery.

Generally, tree crown delineation methods can be grouped into three categories [11]—valley following, region growing, and watershed segmentation. Valley following is the classic algorithm proposed for delineating tree crowns [12]–[17]. The algorithm has been shown effective for even-aged stands of a uniform species. However, if the automatic detection of valleys between trees is difficult, or some tree crowns are self-shaded, the algorithm will not perform well [11]. Region growing has also been widely applied for tree crown delineation [18]–[22]. In region growing methods, users need to provide both seed points and the criteria to stop growing. The criteria to stop growing always refer to thresholds. Although spectral information in the image can be effectively utilized in the region growing process, finding a single, unique, forest-wide spectral threshold in a complex forest is very difficult [9].

According to the analysis of challenges inherent in valley following and region growing, we turned our attention to watershed segmentation [23]. The watershed segmentation requires a grey-level image that is viewed as a topographic surface where the digital value for each pixel is considered as the elevation at the point. The disadvantage of watershed segmentation is that over-segmentation can be caused by image noise. In [24], a tree crown delineation method based on multiscale watershed segmentation was proposed to eliminate over-segmentation. However, three dominant scales calculated from scale analysis could not fully characterize the continuous multiscale of crowns and the associated morphological computation is

Manuscript received March 29, 2021; revised June 15, 2021; accepted July 17, 2021. Date of publication July 29, 2021; date of current version August 18, 2021. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) under the Discovery Grant CPI: 1240316, and in part by the National Natural Science Foundation of China under Grant 41171339. *(Corresponding author: Hengjian Tong.)*

Fei Tong, Rakesh Mishra, and Yun Zhang are with the Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, NB E3B 5A3, Canada (e-mail: ftong@unb.ca; rakesh.mishra@unb.ca; yunzhang@unb.ca).

Hengjian Tong is with the School of Computer Science, China University of Geosciences, Wuhan 430074, China (e-mail: thj26@cug.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3100748

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

complex. In [25], spectral angle segmentation (SAS) was proposed for tree crowns delineation based on watershed segmentation. Watershed segmentation was implemented in the spectral angle gradient image and a series of thresholds were applied to merge regions in different scales. The final delineation was generated by automatically selecting the optimal scale for each tree crown. SAS took full advantage of boundary information contained in the spectral contrast over multiple bands and a multiscale fitting method that could identify optimal parameters to best fit each tree crown was also proposed in SAS. It should be noted that parameters required by SAS can be easily assigned by users because users just need to identify a too coarse scale and a too fine scale to ensure that the optimal scale is between them. Nonetheless, the generated results always contain a lot of over-segmented tree crowns and some shadows are merged with tree crowns.

To avoid the over-segmentation issue, some studies chose to use marker-controlled watershed segmentation to eliminate the negative influence of noise [26]–[28]. Markers are considered as treetops for the tree crown delineation task. Therefore, the location and number of detected treetops will affect the quality of delineation results. For the task of treetop detection, some studies locate treetops by finding spectral local maxima based on the assumption that the peak of the tree crown is located at or very close to the treetop [11], [29]. Local brightest pixels are detected as spectral local maxima. This method performs well on coniferous forests while it performs poorly when dealing with forests with varying crown sizes and illumination conditions. Although some studies have proposed various methods [18], [26], [30], [31] to optimize the spectral local maxima, some demerits still cannot be eliminated. For example, the brightest pixel in a large, rounded crown may not appear at the point of maximum height, while for flat crowns there could be no difference in brightness at all [9]. Besides spectral local maximum, spatial local maximum was also adopted to detect treetops [26]–[28]. The spatial local maximum is defined based on the assumption that the treetop is located near the center of the tree crown [26]. In [26], the spatial local maxima detected from the geodesic distance map were used to refine the result obtained from spectral local maxima. After the refinement, false spectral local maxima were efficiently reduced. For the task of detecting spatial local maxima, the extraction of tree crown borders and the definition of spatial local maximum on the geodesic distance map are two important factors. In [26], tree crown borders were extracted using a Laplacian of Gaussian edge detection operator that may detect a lot of false borders in complex forests. Moreover, as explained in Section II-C of this article, the definition of spatial local maxima on the geodesic distance map may highlight potential tree locations that are unlikely to represent tree centers. In the following research, the authors in [28] applied the same tree crown delineation method as [26] while it was proposed to use a multiscale method to enhance tree crown boundaries for obtaining more accurate tree crown borders. Although the tree crown borders extraction process was improved, the problem about the definition of spatial local maximum still exists. In recently published research [32], a marker-controlled watershed segmentation method called SMS, which is based on morphological gradient image and spectral angle mapper (SAM), was proposed for delineating tree crowns. Although the SMS performed well on two plots containing dense crowns with different shapes, structures, and distinctive colors, the threshold value of SAM still need to be fine tuned by users.

In addition to commonly used valley following, region growing, and watershed segmentation, the authors in [9] proposed to delineate tree crowns using mathematical morphological operations. This method achieved 80% accuracy for tree crown delineation in a highly diverse tropical forest. However, it underestimates small trees and requires some parameters that are not easy for normal users to manually assign. Moreover, in a recently published research [10], the deep learning method was also applied to delineate tree crowns. Although the deep learning method can achieve good results, obtaining sufficient training samples to train the deep learning model is still time consuming.

From our review of the relevant research, it can be concluded that to guarantee good delineation performance, most existing tree crown delineation methods need crucial parameters which are hard to be assigned by users, or a large number of training samples that are hard to be obtained by users. To make it easier for users to get a good delineation result, we propose a tree crown delineation method that can generate satisfactory tree crown delineation results without crucial parameters in a mixed forest captured by WorldView-3 satellite imagery. The proposed method is developed from the marker-controlled watershed segmentation in [26]. Contributions are made on two aspects. 1) Instead of obtaining tree crown borders by edge detection, the proposed method extracts tree crown borders by an automatic gradient binarization process in which the search of the optimal threshold is supervised by a simple classification result which can be easily obtained by normal users. 2) The existing definition of spatial local maximum for detecting treetops in [26] is improved so that fewer false treetops will be detected.

II. Materials and Methods

A. Study Site and Dataset

To verify the robustness of the proposed method, two test images with various tree species and various sizes of tree crowns were selected for experiments. These two test images are located in the district Black Brook, New Brunswick, Canada. The WorldView-3 imagery of the study site was collected on July 27, 2016 and provided by $\mathrm{Maxar}^{\mathrm{TM}}$. There are three main tree species in two test images—spruce (mostly black), birch (mostly white), and maple(mostly red). The collected WorldView-3 satellite image provides a Panchromatic band with a spatial resolution of 0.31 m, and eight visible and near-infrared (VNIR) bands—coastal blue, blue, green, yellow, red, red edge, NIR-1, and NIR-2, each with a spatial resolution of 1.24 m. The Fuze $\mathrm{Go}^{\mathrm{TM}}$ software tool using the UNB Pansharp algorithm [33]–[36] was applied to pan-sharpen the WorldView-3 multispectral and Panchromatic images in order to produce high resolution multispectral image. The Fuze $\mathrm{Go}^{\mathrm{TM}}$ software tool was selected because the UNB Pansharp algorithm is used by $\mathrm{Maxar}^{\mathrm{TM}}$ as the standard fusion tool to pan-sharpen its satellite images. As a result, the 1.24 m resolution of the eight VNIR bands is increased to 0.31 m. After pan-sharpening, the RGB composites

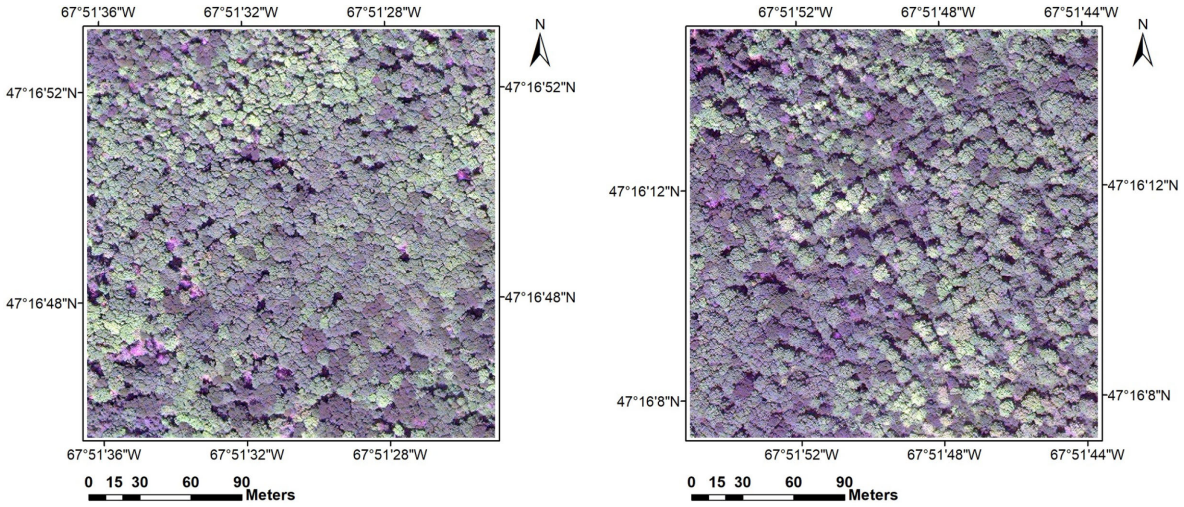


Fig. 1. Multispectral images obtained from the composition of R, G, and B spectral bands with spatial resolution of 0.31 m. (a) Test image 1. (b) Test image 2.

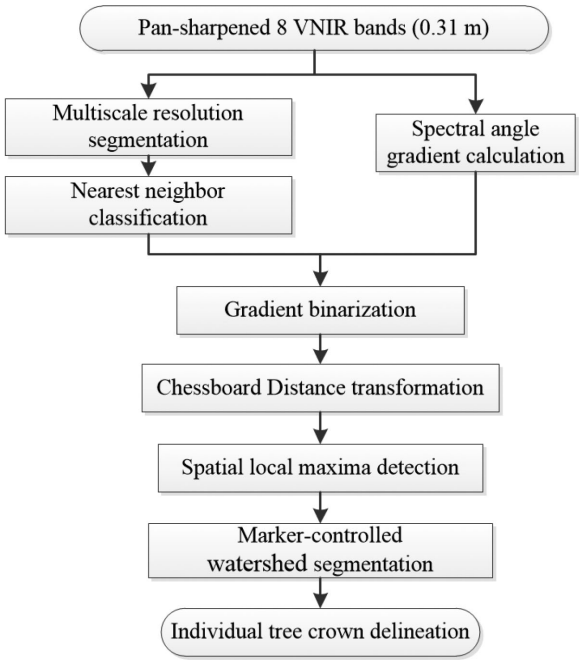


Fig. 2. Flowchart of the proposed delineation method.

of the test images with the size of 800×800 pixels are shown in Fig. 1. Reference delineations used in experiments were drawn manually by an independent, experienced researcher. The numbers of reference tree crowns selected for the two test images were 410 and 303. In the following sections, the detail of the proposed delineation method will be introduced. The flowchart of the proposed method is shown in Fig. 2.

B. Tree Crown Borders Extraction

1) Classification for Identifying Shadows: Much previous research by others has proven that shadows provide useful information for extracting tree crown borders. In this section, a coarse, object-based classification is applied to generate a binary map in which pixel values of shadows and tree crowns are labeled as 0 and 1, respectively. First, the multiresolution segmentation [37] is applied to segment the image. The segmentation scale should be fine so that the obvious shadows between tree crowns will not be merged with tree crowns. Next, a simple supervised classification is executed based on the generated segmentation map, to create a binary map. The features selected for classification are the mean values and standard deviations of all VNIR bands. The nearest neighbor classification method is applied for the classification. Generally, two classes, shadows and tree crowns, are labeled in the classification result. For each class, several regions that can be easily distinguished by the human eye are selected as samples. After classification, the expected binary map will be generated. It should be noted that some images may contain some classes such as grassland, bare land, and water, which are not intersected in the tree crown delineation task. For such images, three classes are labeled in the classification result—shadow, tree crown, and a special class containing all other classes. After classification, all regions labeled as the special class in the classification result will not be involved in the subsequent process. After excluding all special regions, the classification result can be transformed into a binary map in which shadow areas and tree crowns are represented by 0 and 1, respectively. As spectral characteristics of shadows are obviously different from tree crowns, it is easy to discriminate most shadows in the image using the supervised classification method. However, some small shadow regions between tree crowns are still misclassified.

2) Extracting Tree Crown Borders by Gradiant Binarization: Generally, tree crown borders tend to have high spectral gradient magnitude while pixels in tree crowns have low spectral gradient magnitude. To accurately extract tree crown borders, a binarization process on the gradient image is proposed in this article. The spectral gradient angle, which can capture the gradient within multiple spectral bands, is adopted for calculating the gradient image. Assuming that the multispectral image has n bands, the

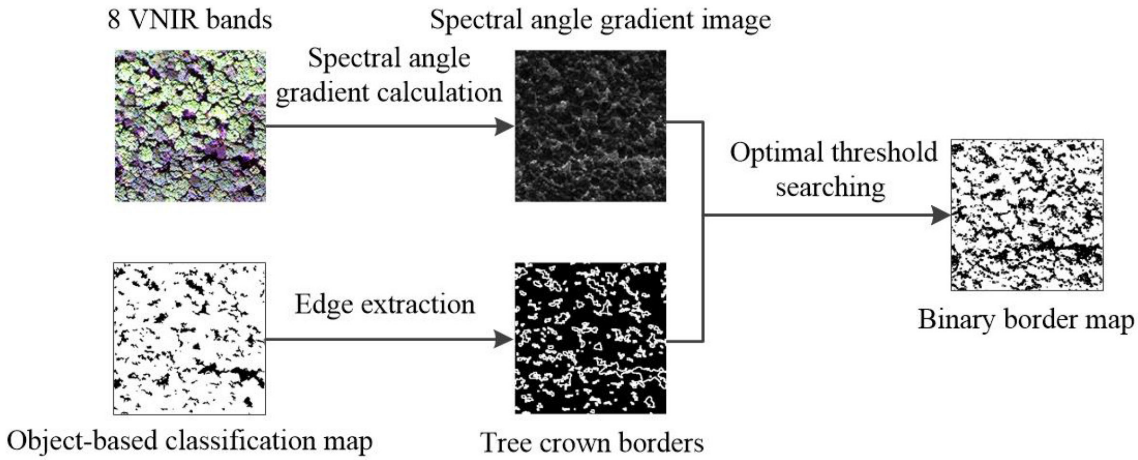


Fig. 3. Gradient binarization process.

spectral angle between two pixels is given by [38]

$$\Theta_{ab} = \cos^{-1} \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2}} \tag{1}$$

where a_i and b_i are pixel values of ith band from two different pixels. The spectral gradient angle for each pixel is defined as the maximum spectral angle between any pair of pixels within a moving window W of a fixed size. The window size used in this article is 3×3.

To execute binarization on the gradient map, a gradient threshold is required. It is hard to manually find the optimal threshold for the gradient. To automatically select the threshold, reference information is required. As a binary map has been generated from the previous classification process, border pixels can be extracted from the binary map. Although all identified border pixels are related to shadows, these border pixels still occupy a large percentage of all tree crown border pixels in the image. Therefore, it is assumed that the gradient values of identified border pixels cover the whole range of possible gradient values for borders. Identified border pixels are considered as reference information, and applied to find the optimal gradient threshold. To make the search of the threshold more general, pixel values in the gradient image are transformed to the range $0 \sim 255$ first. Then a search process is conducted to determine the gradient threshold. The search for the optimal threshold starts at 255 and goes to 1 with an interval of -2. The similarity $\mathbf{Sim}$ between border pixels identified from the classification result and the border pixels generated from a selected threshold is

$$\mathbf{Sim} = \frac{h_{b,b}}{h_{b,i} + h_{i,b}} \tag{2}$$

where $h_{b,b}$ represents the number of pixels identified as "border" by both the classification and gradient binarization, $h_{b,i}$ represents the number of pixels identified as "border" by the classification but identified as "internal pixels" by the gradient binarization, and conversely $h_{i,b}$ represents the number of pixels identified as "internal pixels" by the classification but identified as "border" by the gradient binarization. In the search process, the threshold that receives the greatest value for $\mathbf{Sim}$ will be chosen as the optimal threshold. The final binarization will then be executed based on this optimal threshold. The procedure for gradient binarization is shown in Fig. 3.

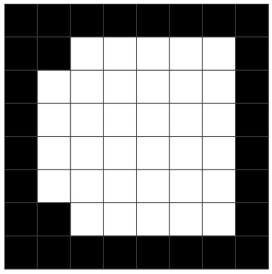
C. Spatial Local Maxima Treetop Detection

As our proposed method is developed based on the method proposed in [26], in this section, the procedure of detecting spatial local maxima in [26] is introduced first and then the improvements made by our proposed method will be explained.

In [26], treetops are detected by locating spatial local maxima in a morphologically transformed distance map. A binary map in which tree crown pixels are labeled as 1 and background pixels are labeled as 0 is required. Tree crown border pixels are background pixels that are adjacent to tree crown pixels. Then the distance map will be calculated. In the distance map, the values of background pixels are assigned to 0, and the value of each tree crown pixel is the distance between it and its nearest tree crown border pixel. The distance between two pixels is calculated based on the properties of a structure element (SE), which is a group of connected pixels that resembles the geometry of the object to be measured. Distance can be measured only along connected paths as defined by the SE, and the length of each step is determined by the value of each pixel in the SE. The elementary disk SE is a 3×3 window of pixels whose values are all equal to 1. In fact, the distance can be called Chebyshev Distance [39]. The distance between two pixels $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ is defined as

$$\mathbf{Dis} = \max(|x_1 - x_2|, |y_1 - y_2|). \tag{3}$$

After transformation, a distance map with integer pixel values is generated. The local maxima can then be extracted based on the distance map. The spatial local maximum in [26] is defined as—a connected group of pixels with a single distance value such that each pixel in the group has a value greater than or equal to all the pixel values within the surrounding eight-connectivity neighborhood. The local maximum is then labeled as a treetop based on the assumption that treetops are usually located around centers of tree crowns.



(a)

0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	0
0	1	1	2	2	2	1	0
0	1	2	2	3	2	1	0
0	1	2	2	3	2	1	0
0	1	1	2	2	2	1	0
0	0	1	1	1	1	1	0
0	0	0	0	0	0	0	0

(b)

Fig. 4. Example of spatial local maxima detection. (a) Binary image for an example tree crown. (b) Local maxima detected by the definition in [26].

In the proposed method, tree crown pixels are identified based on the classified binary map. Tree crown borders are border pixels that are adjacent to internal pixels in the binary border map generated from gradient binarization. Then distance map indicating distances between tree crown pixels and their nearest borders is calculated. Spatial local maxima will be detected based on this distance map. According to the definition of spatial local maximum in [26], two kinds of local maxima will be detected, however, one of them is false. In Fig. 4(a), there is an image for example tree crown in which each grid represents one pixel (tree crown pixels are white and tree crown borders are black). It is clear that only one local maximum should be detected from the tree crown in Fig. 4(a). However, as shown in Fig. 4(b), in the distance map calculated from (3), two spatial local maxima surrounded by red frames are detected according to the definition in [26]. In fact, the spatial local maximum with the value of 3 is real while the one with the value of 2 is false. In a real forest image, a lot of tree crowns similar to the one shown in Fig. 4(a) can be found. Therefore, the quality of the tree crown delineation result will be negatively influenced. To solve this problem, we propose a modification of the definition of the spatial local maximum. The new local maximum is defined as a connected group of pixels with a single distance value that is greater than the distance values of all neighbor pixels surrounding the group. The purpose of this modification is to exclude false local maximum such as the one shown in Fig. 4(b).

The new process of detecting the spatial local maxima on the transformed distance map is illustrated in Fig. 5. In the process of eliminating false local maxima, if the value of any neighbor pixel is greater than or equal to the local maximum, the local maximum will be deleted. The local maxima detected by the proposed definition will be recognized as treetops that will be used as markers in marker-controlled watershed segmentation.

D. Marker-Controlled Watershed Segmentation

After treetops are marked, marker-controlled watershed segmentation is applied to delineate the tree-crown boundary for each individual tree within each crown object. The resultant process for marker-controlled watershed segmentation [40] can be summarized as follows:

1) Inverts the value of each pixel so that the local maximum is transformed into local minimum.
2) Introduces the virtual water to the system.
3) Each valley collects water starting from predefined marker until the water spills over the watershed into an adjacent valley.

The key points of the marker-controlled watershed segmentation are the selection of markers, and the selection of the base image. In this article, the markers are spatial local maxima detected by the proposed method and the base image is the calculated distance map. After the marker-controlled watershed segmentation, the final delineation result will be generated.

E. Accuracy Assessment

To validate the performance of the proposed delineation method, two metrics were selected to assess the accuracy of tree crown delineation, the object recognition rate (ORR) [27] and the segmentation evaluation index (SEI) used in [25]. ORR is a two-sided metric that identifies the rate of correctly delineated tree crowns

$$\mathbf{ORR} = \frac{\mathrm{N}_{correct}}{\mathrm{N}_{total}} \tag{4}$$

where the $N_{correct}$ refers to the number of correctly delineated tree crowns and N_{total} is the number of total reference tree crowns. The correctly delineated tree crowns are those which overlap reference crowns by more than 50% of both the segment and the reference polygon. The value of ORR ranges from 0 to 1, with higher values indicating higher accuracy. The SEI was also selected as a metric to further quantify overlaps of segments and reference crowns. It measures the goodness of fit for the delineated tree crowns

$$\mathbf{SEI} = \frac{1}{I}\sum_{i=1}^{I}\mathbf{SEI}_{\mathrm{local}}(i) \tag{5}$$

where I is the number of reference crowns and $\mathbf{SEI}_{\mathrm{local}}(i)$ represents the SEI for ith reference crown. To calculate $\mathbf{SEI}_{\mathrm{local}}(i)$, if the corresponding segment s_i for ith reference crown r_i cannot be found, the value of $\mathbf{SEI}_{\mathrm{local}}(i)$ is 0.71, otherwise $\mathbf{SEI}_{\mathrm{local}}(i)$ is calculated as [25]

$$\mathbf{SEI}_{\mathrm{local}}(i) = \sqrt{\frac{(1-\frac{\mathrm{area}(r_i\cap s_i)}{\mathrm{area}(r_i)})^2 + (1-\frac{\mathrm{area}(r_i\cap s_i)}{\mathrm{area}(s_i)})^2}{2}}. \tag{6}$$

A segment is labeled as the corresponding segment of a reference polygon once the overlapping area is more than 50% of both the reference polygon and the segment. The value of SEI ranges from 0 to 0.71, with lower values indicating higher accuracy.

In addition to metrics mentioned above, to better illustrate the weaknesses of the delineation method, the number of reference crowns which could be put into the following two categories were counted. Each category indicates different kinds of errors according to their spatial relationships with the generated target segments [1], [24].

1) Merged: If more than half the area of multiple reference crowns was covered by a single target segment, the multiple reference crowns are taken as crowns merged in the automatic delineation.

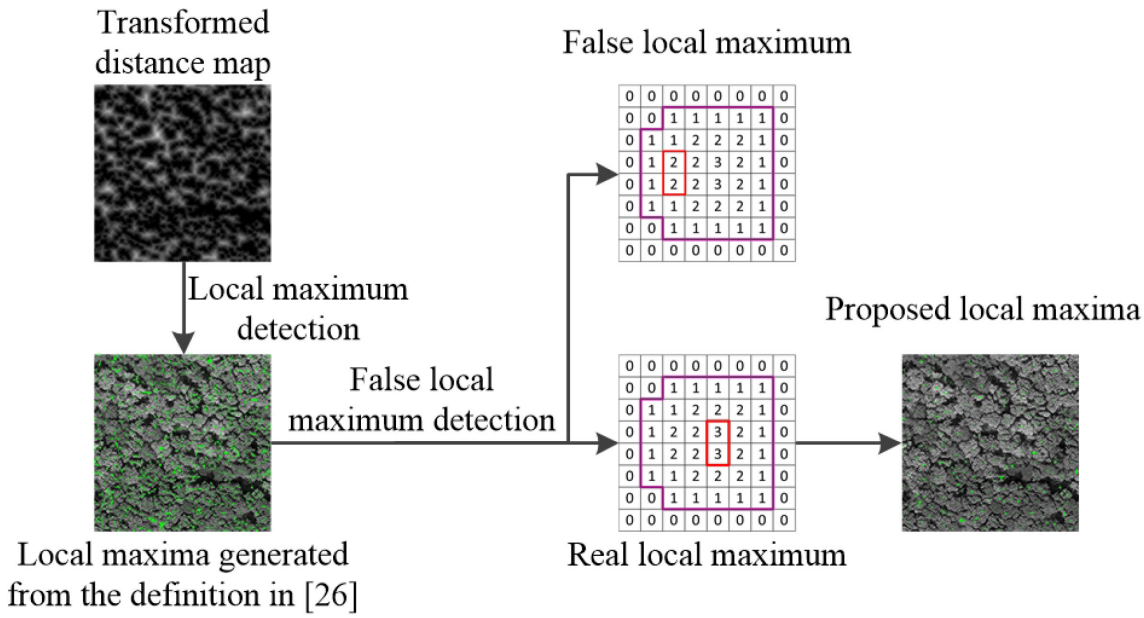


Fig. 5. Process of detecting spatial local maxima.

2) Split: If more than half the area of multiple target segments were covered by a single reference crown, the reference crown is considered as a crown split in the automatic delineation.

It should be noted that the tree crowns identified as correctly delineated in calculating ORR may be labeled as "Merged" or "Split" according to the above definition.

III. Experiments

A. Comparison With the SAS on Delineation Results

To demonstrate the effectiveness of the proposed tree crown delineation method, the SAS in [25] was implemented for comparison because the quality of delineation results generated by SAS does not rely on user-assigned parameters. For the SAS method, all eight VNIR bands were used in the calculation of spectral angle gradient. The ranges of seed-to-saddle thresholds in SAS were: 0.2 to 1.1 with an interval of 0.1 for test image 1; and 0.2 to 1.0 with an interval of 0.1 for test image 2. The upper bound and lower bound of the thresholds were determined by the delineation generated from a single scale threshold. This was done to make sure that the upper bound was coarse enough and the lower bound was fine enough so that the optimal scale was between them. For the proposed method, parameters for multiresolution segmentation in two test images were the same (scale: 50, the weight of shape: 0.1, the weight of compactness: 0.9) and all 8 VNIR bands participated in the multiresolution segmentation. In the process of object-based classification on eCognition, 20 shadow regions and 20 tree crown regions were selected for the classification of two test images. All eight VNIR bands were used in the calculation of the spectral angle gradient.

The reference tree crowns and delineation results of two test images generated from the proposed method are shown in Figs. 6 and 7, respectively. The related accuracy metrics of delineations for two test images are tabulated in Table I. To better show the delineation results, the subsets of the results with the size of 256×256 pixels from two test images are shown in Fig. 8. As can be observed from Table I, the proposed method performed better than the delineations using SAS in both metrics (ORR and SEI), for both of the two test images. Moreover, from the data listed in Table I and delineation maps shown in Fig. 8, it can be observed that the results from SAS had more over-segmentations and under-segmentation compared to that from the proposed method.

TABLE I
Delineation Accuracy of Two Test Images

Image	Method	Merged	Split	ORR(%)	SEI
1	**SAS**	39	220	52.20	0.47
	Proposed	23	158	73.41	0.35
2	**SAS**	42	178	47.19	0.50
	Proposed	8	139	71.62	0.37

B. Comparison With Other Methods

To further validate the performance of the proposed delineation method, two other delineation methods that could be applied to high spatial resolution multispectral imagery were also implemented for comparison. The first one is the multiscale segmentation method proposed in [24], called MSS for brevity. The gray image used for delineation was the first principal components analysis (PCA) band which was extracted from all eight VNIR bands using the PCA algorithm [41]. Three dominant tree crown sizes calculated based on scale analysis for both two test images were 9, 17, and 23 pixels. The thinness ratio threshold k was set as 0.85 based on the fact that a tree crown normally has a circularity value higher than 0.85 [42]. As area ratio threshold

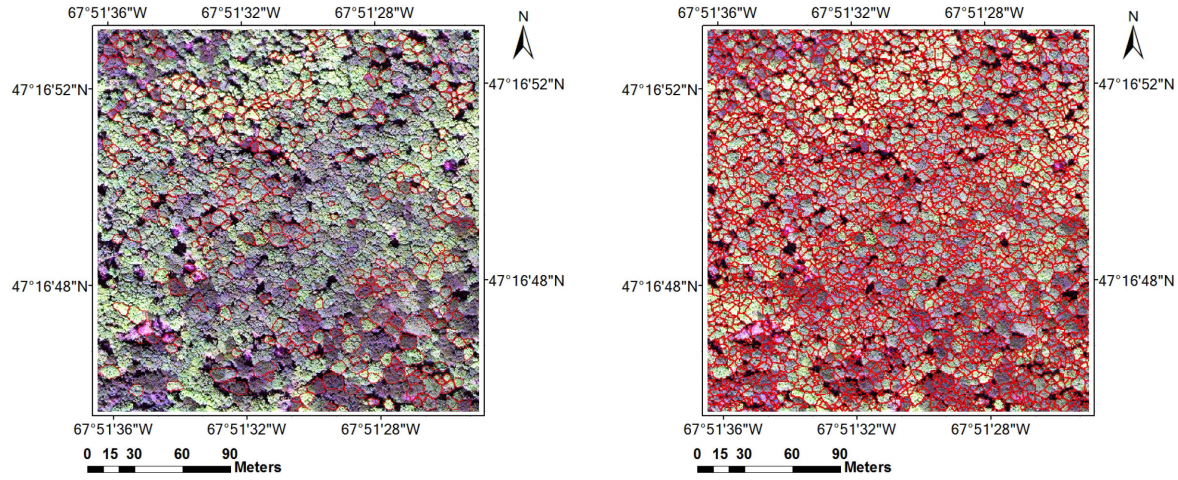


Fig. 6. (a) Reference map of test image 1. (b) Delineation result of test image 1.

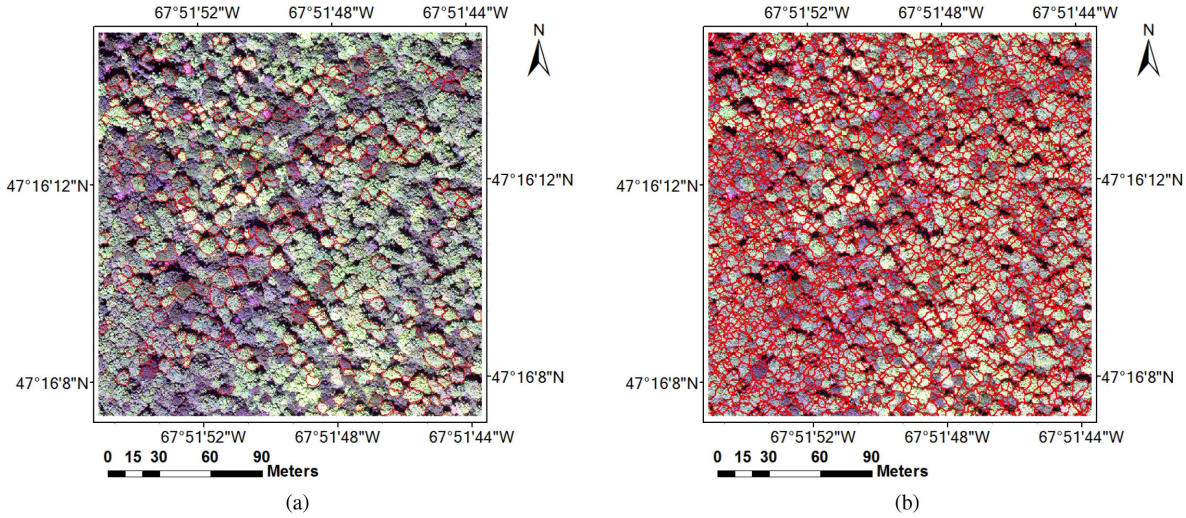


Fig. 7. (a) Reference map of test image 2. (b) Delineation result of test image 2.

T_A would affect the delineation accuracy, the T_A was tuned to achieve the best delineation accuracy. The optimal T_A for both two test images was 0.3. The second method is the SMS proposed in [32]. The threshold for roundness was 0.6, which was the same as that used in [32]. The initial SAM thresholds for both images were obtained based on the statistical results of all SAM values calculated on the image (0.995 for test image 1 and 0.9974 for test image 2). After fine tuning according to delineation accuracy, the SAM threshold for test image 1 was set as 0.999 and for test image 2 was set as 0.9994. The accuracy metrics and time cost of delineations for two test images are tabulated in TABLE II. It should be noted that the time cost for parameters tuning in MSS and SMS was not counted. Besides, the time cost for multiresolution segmentation and nearest neighbor classification in the proposed method was also not counted. All delineations were implemented using MATLAB programming language on a desktop with Intel(R) Core(TM) i7-2600 CPU @ 3.40 GHz. From the perspective of delineation accuracy, it can be observed that the proposed method achieved higher delineation accuracy than that of the other two methods. From the perspective of time cost, the proposed method required less time than MSS and SMS. Although the time cost for multiresolution segmentation and nearest neighbor classification in the proposed method cannot be ignored, the parameter tuning process that is difficult without the ground truth is not required.

C. *Effects of Gradient Binarization*

As tree crown borders can also be extracted from the classification result, to validate that tree crown borders extracted from gradient binarization result are more accurate than that extracted from the classification result, two delineations that used different tree crown borders were conducted and their delineation accuracy was compared. The results of the comparison are tabulated in Table III. From Table III, it can be found that the delineation based on gradient binarization obviously outperformed the delineation based on classification. In addition, as more borders were not detected in the classification, the delineation based on

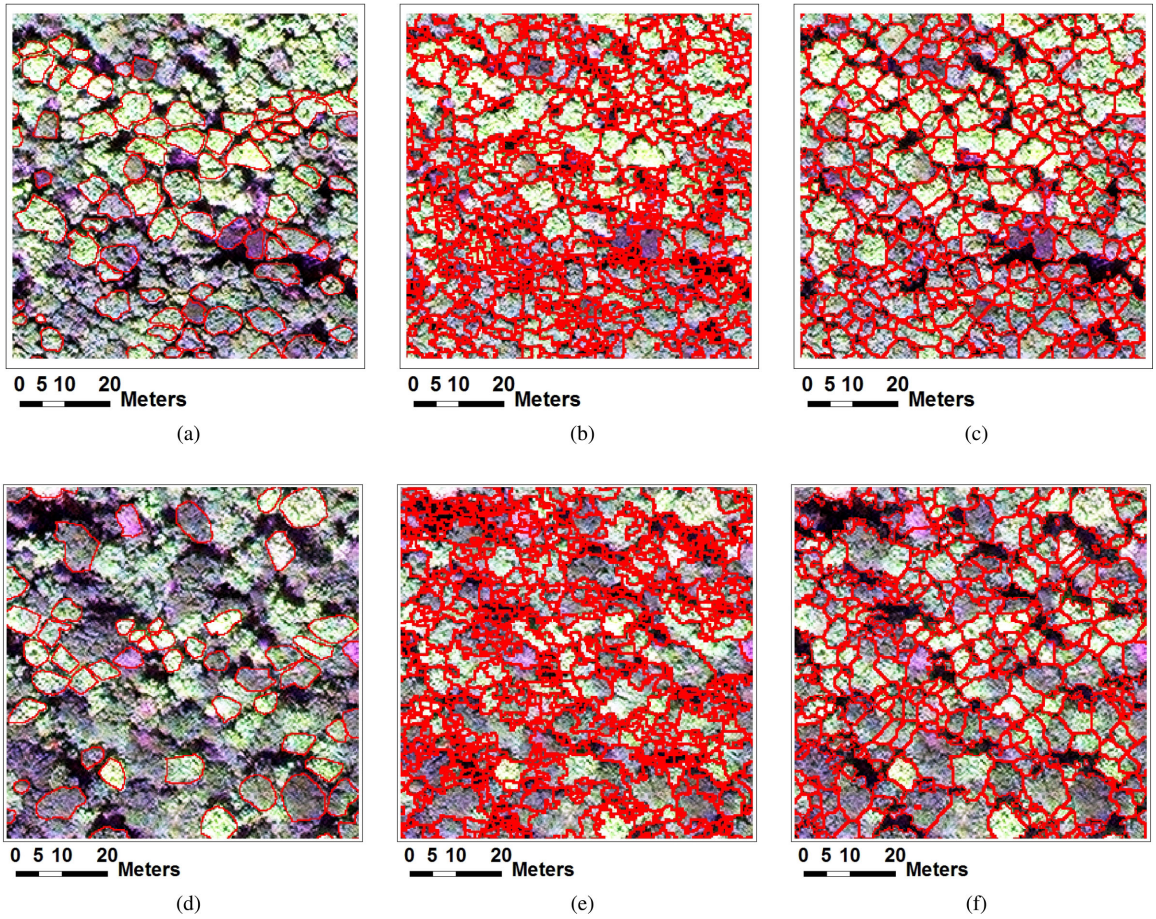


Fig. 8. Subset in test image 1. (a) Reference map. (b) SAS delineation map. (c) Delineation map with the proposed method. Subset in test image 2. (d) Reference map. (e) SAS delineation map. (f) Delineation map with the proposed method.

TABLE II
DELINEATION ACCURACY AND TIME COST FOR DIFFERENT METHODS

Image	Method	ORR(%)	SEI	Time Cost (s)
1	**MSS**	41.71	0.53	73
	SMS	57.07	0.44	35
	Proposed	73.41	0.35	13
2	**MSS**	34.65	0.56	65
	SMS	40.92	0.52	37
	Proposed	71.62	0.37	13

TABLE III
DELINEATION ACCURACY OF TWO TEST IMAGES USING DIFFERENT BORDERS

Image	Borders	Merged	Split	ORR(%)	SEI
1	**Classification**	154	28	30.00	0.57
	Binarization	23	158	73.41	0.35
2	**Classification**	112	30	32.34	0.56
	Binarization	8	139	71.62	0.37

classification had more merged tree crowns that will negatively affect subsequent processes like tree species classification.

D. *Effects of Different Treetop Detection Methods*

Another improvement inherent in the proposed method is the modification of the definition of the spatial local maximum. Here, we implemented three delineations using different treetop detection methods and compared their delineation accuracy with that of the proposed method. The first delineation detected treetops using the spectral maxima. The spectral maxima were detected on the first PCA band. To get a good gray-scale image for detecting spectral local maxima, the PCA band was smoothed by a Gaussian filter to remove noises. Three filter sizes 3×3, 5×5, and 7×7 were adopted to smooth the filter, and the corresponding standard deviation $\boldsymbol{\sigma}$ for each filter was set as $\frac{1}{3}$ of the filter size. Finally, 5×5 was selected to smoothed the image because the delineation with this filter size got the highest accuracy. The second delineation detected treetops using the spatial local maxima defined in [26]. The third delineation detected treetops using intersected local maxima that combined spectral and spatial local maxima as described in [26].

The distance map used for detecting spatial local maxima was as same as that of the proposed method. The delineation accuracy of these delineations is listed in Table IV. In Table IV, it can be found that the proposed method achieved much higher delineation accuracy compared to other delineations. Too many

TABLE IV
DELINEATION ACCURACY OF TEST IMAGES WITH DIFFERENT TREETOPS

Image	Treetop	Merged	Split	ORR(%)	SEI
1	Spectral local maxima	28	275	46.59	0.50
	Original spatial local maxima	0	406	5.37	0.69
	Intersected local maxima	65	192	53.17	0.47
	Proposed spatial local maxima	24	136	73.90	0.35
2	Spectral local maxima	14	240	50.50	0.49
	Original spatial local maxima	0	301	2.97	0.70
	Intersected local maxima	42	144	47.85	0.49
	Proposed spatial local maxima	8	139	71.62	0.37

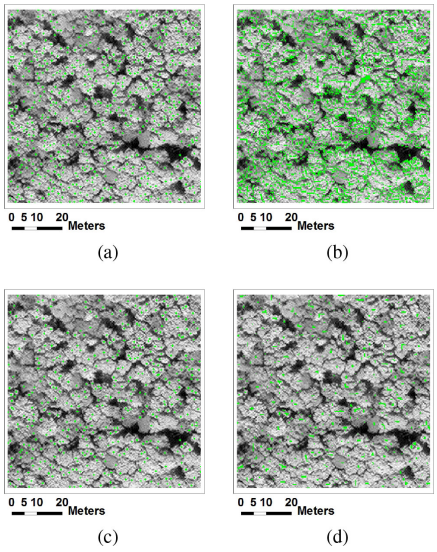


Fig. 9. Four kinds of treetops on PCA band. (a) Spectral local maxima. (b) Original spatial local maxima. (c) Intersected local maxima. (d) Proposed spatial local maxima.

split tree crowns were generated in the delineation using spatial local maxima defined in [26] because a lot of false local maxima were reserved. Although the best filter size was selected to smooth the PCA band, the accuracy for the delineation using spectral local maxima was still low because a single filter size could not perfectly work on various sizes of tree crowns and some spectral local maxima could not accurately locate positions of treetops. To better show the performance of different treetop detection methods, four different kinds of treetops detected on a plot cropped from test image 1 are shown in Fig. 9. As shown in Fig. 9, there are multiple spectral local maxima within most tree crowns. The proposed spatial local maxima successfully eliminated a lot of false treetops compared to spatial local maxima defined in [26]. Moreover, multiple intersect local maxima were found in some tree crowns and part of the intersected treetops were not located at the center of tree crowns.

IV. DISCUSSION

In this article, we proposed a tree crown delineation method that took advantage of both spectral information and spatial information in the satellite image. Tree crown borders were extracted from the gradient image calculated from multiple spectral bands. Treetops were located on a distance map measuring spatial distances between tree crown pixels and tree crown borders. Instead of setting the parameters manually, the proposed method automatically searched the optimal parameter supervised by the tree crown borders extracted from a classification result. As spectral features of shadows are obviously different from tree crowns, it is easy to get a good classification result by normal users. It is acceptable that the classification result is not perfect because the classification result just provides reference information for gradient binarization. Therefore, we just applied a simple nearest neighbor classifier to get the classification result. The experiment results (see Table III) showed that with the tree crown borders extracted from the classification result, the delineation accuracy was low because some tree crown borders were missed in the classification. However, the delineation using tree crown borders extracted from gradient binarization achieved much higher accuracy because more tree crown borders were detected in the binarization process. Although the classification result was not perfect, after the gradient binarization, the final delineation result was still satisfactory. Therefore, compared with those methods that require crucial parameters, our method can continually generate good delineation results by normal users.

In the comparison with the SAS method that also does not need crucial parameters, it was apparent that the proposed method generated more accurate delineation results. Although our method demonstrates great potential for tree crown delineation task, over-segmentations and a few under-segmentations were also found in the generated products. Most over-segmentations were caused by noise pixels or inner shadows within tree crowns. To solve this problem, most existing methods require user-defined parameters. As we were trying to provide a method that does not rely on user-defined parameters, the smoothing filter that can remove noises and the area threshold that can eliminate inner shadows were not applied in the proposed delineation method. Under-segmentations were caused by missing some tree crown borders in the gradient binarization process. As reference tree crown borders from the classification result were strong borders connecting tree crowns and shadows, some weak borders between overlapped tree crowns might be ignored in the gradient binarization process. Actually, such weak borders between overlapped tree crowns were difficult to be accurately detected using spectral information. From the perspective of the application, the proposed method was not suitable to count the number of tree crowns in the forest because over-segmented and under-segmented tree crowns would negatively affect the accuracy. However, the proposed delineation method could benefit the individual tree-based species classification. Although over-segmented tree crowns existed, they would not negatively affect the classification accuracy. Under-segmented tree crowns would lower the classification accuracy while under-segmented tree crowns could be refined according to the spectral information [43] or the shape information [44], [45].

V. CONCLUSION

In this article, a novel method for automatically delineating individual tree crowns in WorldView-3 satellite imagery is proposed, which can effectively detect tree crown borders and

also accurately locate the treetops. In terms of the quantitative assessment and visual interpretation, the proposed tree crown delineation method outperformed the SAS method. Moreover, the comparison among different kinds of treetops demonstrated that the proposed spatial local maxima were more reliable for locating treetops.

ACKNOWLEDGMENT

The authors would like to thank anonymous reviewers for providing their valuable comments and suggestions.

REFERENCES

[1] L. Jing, B. Hu, J. Li, and T. Noland, "Automated delineation of individual tree crowns from LiDAR data by multi-scale analysis and segmentation," *Photogramm. Eng. Remote Sens.*, vol. 78, no. 12, pp. 1275–1284, 2012.

[2] V. F. Strîmbu and B. M. Strîmbu, "A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data," *ISPRS J. Photogramm. Remote Sens.*, vol. 104, pp. 30–43, 2015.

[3] F. Fang, J. Im, J. Lee, and K. Kim, "An improved tree crown delineation method based on live crown ratios from airborne LiDAR data," *GISci. Remote Sens.*, vol. 53, no. 3, pp. 402–419, 2016.

[4] B. Wu, B. Yu, Q. Wu, Y. Huang, Z. Chen, and J. Wu, "Individual tree crown delineation using localized contour tree method and airborne LiDAR data in coniferous forests," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 52, pp. 82–94, 2016.

[5] Z. Zhen, L. J. Quackenbush, and L. Zhang, "Trends in automatic individual tree crown detection and delineation-evolution of LiDAR data," *Remote Sens.*, vol. 8, no. 4, 2016, Art. no. 333.

[6] C. Barnes, H. Balzter, K. Barrett, J. Eddy, S. Milner, and J. C. Suárez, "Individual tree crown delineation from airborne laser scanning for diseased larch forest stands," *Remote Sens.*, vol. 9, no. 3, 2017, Art. no. 231.

[7] W. M. Jaafar *et al.*, "Improving individual tree crown delineation and attributes estimation of tropical forests using airborne LiDAR data," *Forests*, vol. 9, no. 12, 2018, Art. no. 759.

[8] F. Naveed, B. Hu, J. Wang, and G. B. Hall, "Individual tree crown delineation using multispectral LiDAR data," *Sensors*, vol. 19, no. 24, 2019, Art. no. 5421.

[9] F. H. Wagner *et al.*, "Individual tree crown delineation in a highly diverse tropical forest using very high resolution satellite images," *ISPRS J. Photogramm. Remote Sens.*, vol. 145, pp. 362–377, 2018.

[10] J. R. Braga *et al.*, "Tree crown delineation algorithm based on a convolutional neural network," *Remote Sens.*, vol. 12, no. 8, 2020, Art. no. 1288.

[11] Y. Ke and L. J. Quackenbush, "A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing," *Int. J. Remote Sens.*, vol. 32, no. 17, pp. 4725–4747, 2011.

[12] F. A. Gougeon, "A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images," *Can. J. Remote Sens.*, vol. 21, no. 3, pp. 274–284, 1995.

[13] F. A. Gougeon *et al.*, "Automatic individual tree crown delineation using a valley-following algorithm and rule-based system," in *Proc. Int. Forum Automated Interpretation High Spatial Resolution Digit. Imagery Forestry*, Victoria, BC, Canada, 1998, pp. 11–23.

[14] F. A. Gougeon *et al.*, "Forest information extraction from high spatial resolution images using an individual tree crown approach," Natural Resour. Canada, Ottawa, ON, Canada, Inf. Rep. BC-X-396, 2003.

[15] D. G. Leckie, F. A. Gougeon, N. Walsworth, and D. Paradine, "Stand delineation and composition estimation using semi-automated individual tree crown analysis," *Remote Sens. Environ.*, vol. 85, no. 3, pp. 355–369, 2003.

[16] D. G. Leckie, F. A. Gougeon, S. Tinis, T. Nelson, C. N. Burnett, and D. Paradine, "Automated tree recognition in old growth conifer stands with high resolution digital imagery," *Remote Sens. Environ.*, vol. 94, no. 3, pp. 311–326, 2005.

[17] F. A. Gougeon and D. G. Leckie, "The individual tree crown approach applied to Ikonos images of a coniferous plantation area," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 11, pp. 1287–1297, 2006.

[18] D. S. Culvenor, "TIDA: An algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery," *Comput. Geosci.*, vol. 28, no. 1, pp. 33–44, 2002.

[19] M. Erikson, "Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules," *Can. J. Forest Res.*, vol. 33, no. 8, pp. 1557–1563, 2003.

[20] Z. Zhen, L. J. Quackenbush, and L. Zhang, "Impact of tree-oriented growth order in marker-controlled region growing for individual tree crown delineation using airborne laser scanner (ALS) data," *Remote Sens.*, vol. 6, no. 1, pp. 555–579, 2014.

[21] M. P. Ferreira *et al.*, "Automatic tree crown delineation in tropical forest using hyperspectral data," in *Proc. IEEE Geosci. Remote Sens. Symp.*, 2014, pp. 784–787.

[22] Z. Zhen, L. J. Quackenbush, S. V. Stehman, and L. Zhang, "Agent-based region growing for individual tree crown delineation from airborne laser scanning (ALS) data," *Int. J. Remote Sens.*, vol. 36, no. 7, pp. 1965–1993, 2015.

[23] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, Jun. 1991.

[24] L. Jing, B. Hu, T. Noland, and J. Li, "An individual tree crown delineation method based on multi-scale segmentation of imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 70, pp. 88–98, 2012.

[25] J. Yang, Y. He, J. P. Caspersen, and T. A. Jones, "Delineating individual tree crowns in an uneven-aged, mixed broadleaf forest using multispectral watershed segmentation and multiscale fitting," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 4, pp. 1390–1401, Apr. 2017.

[26] L. Wang, P. Gong, and G. S. Biging, "Individual tree-crown delineation and treetop detection in high-spatial-resolution aerial imagery," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 3, pp. 351–357, 2004.

[27] W. R. Lamar, J. B. McGraw, and T. A. Warner, "Multitemporal censusing of a population of eastern hemlock (Tsuga canadensis L.) from remotely sensed imagery using an automated segmentation and reconciliation procedure," *Remote Sens. Environ.*, vol. 94, no. 1, pp. 133–143, 2005.

[28] L. Wang, "A multi-scale approach for delineating individual tree crowns with very high resolution imagery," *Photogramm. Eng. Remote Sens.*, vol. 76, no. 4, pp. 371–378, 2010.

[29] A. J. Pinz, "A computer vision system for the recognition of trees in aerial photographs," in *Proc. Int. Assoc. Pattern Recognit. Workshop*, 1991, pp. 111–124.

[30] M. Wulder, K. O. Niemann, and D. G. Goodenough, "Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery," *Remote Sens. Environ.*, vol. 73, no. 1, pp. 103–114, 2000.

[31] D. Pouliot and D. King, "Approaches for optimal automated individual tree crown detection in regenerating coniferous forests," *Can. J. Remote Sens.*, vol. 31, no. 3, pp. 255–267, 2005.

[32] L. Qiu, L. Jing, B. Hu, H. Li, and Y. Tang, "A new individual tree crown delineation method for high resolution multispectral imagery," *Remote Sens.*, vol. 12, no. 3, 2020, Art. no. 585.

[33] Y. Zhang, "Understanding image fusion," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 6, pp. 657–661, 2004.

[34] Y. Zhang and R. K. Mishra, "A review and comparison of commercially available pan-sharpening techniques for high resolution satellite image fusion," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2012, pp. 182–185.

[35] Y. Zhang and R. K. Mishra, "From UNB PanSharp to Fuze Go - The success behind the pan-sharpening algorithm," *Int. J. Image Data Fusion*, vol. 5, no. 1, pp. 39–53, 2014. [Online]. Available: https://doi.org/10.1080/19479832.2013.848475

[36] Y. Zhang, A. Roshan, S. Jabari, S. A. Khiabani, F. Fathollahi, and R. K. Mishra, "Understanding the quality of pansharpening-A lab study," *Photogramm. Eng. Remote Sens.*, vol. 82, no. 10, pp. 747–755, 2016.

[37] M. Baatz, "Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation," in *Proc. Angewandte Geographische Inf. Sverarbeitung XII*, 2000, pp. 12–23.

[38] F. A. Kruse *et al.*, "The spectral image processing system (SIPS)-Interactive visualization and analysis of imaging spectrometer data," in *Proc. AIP Conf.*, 1993, pp. 192–201.

[39] C. D. Cantrell, *Modern Mathematical Methods for Physicists and Engineers*. New York, NY, USA: Cambridge Univ. Press, 2000.

[40] B. Sagar, K. Parvati, B. S. Prakasa Rao, and M. Mariya Das, "Image segmentation using gray-scale morphology and marker-controlled watershed transformation," *Discrete Dyn. Nat. Soc.*, vol. 2008, 2008, Art. no. 384346. [Online]. Available: https://doi.org/10.1155/2008/384346

[41] R. Vidal, Y. Ma, and S. S. Sastry, *Principal Component Analysis*. New York, NY, USA: Springer, 2016, pp. 25–62.

[42] B.-M. Wolf and C. Heipke, "Automatic extraction and delineation of single trees from remote sensing data," *Mach. Vis. Appl.*, vol. 18, no. 5, pp. 317–330, 2007.

[43] W. Zhang, L. J. Quackenbush, J. Im, and L. Zhang, "Indicators for separating undesirable and well-delineated tree crowns in high spatial resolution images," *Int. J. Remote Sens.*, vol. 33, no. 17, pp. 5451–5472, 2012.
[44] D. G. Leckie, N. Walsworth, and F. A. Gougeon, "Identifying tree crown delineation shapes and need for remediation on high resolution imagery using an evidence based approach," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 206–227, 2016.
[45] D. G. Leckie *et al.*, "Automated individual tree isolation on high-resolution imagery: Possible methods for breaking isolations involving multiple trees," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 7, pp. 3229–3248, Jul. 2016.

Fei Tong received the B.S. and M.S. degrees in computer science and technology from the China University of Geosciences, Wuhan, China, in 2015 and 2018, respectively. He is currently working toward the Ph.D. degree in AI based remote sensing environmental information extraction with the University of New Brunswick, Fredericton, NB, Canada.

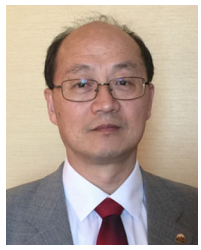
His main research interests include remote sensing image processing and artificial intelligence.

Hengjian Tong received the B.S. degree in software engineering from Northeast Petroleum University, Daqing, China, in 1993, the M.S. degree in computer application, and Ph.D. degrees in cartography and geography information system from the China University of Geosciences, Wuhan, China, in 1999 and 2003, respectively.

He is currently a Professor with the China University of Geosciences. His main research interests include remote sensing image segmentation and geographic object-based image analysis (GEOBIA).

Rakesh Mishra received the Ph.D. degree in remote sensing data fusion and high resolution information extraction from University of New Brunswick, Fredericton, NB, Canada, in 2017.

He is an Adjunct Professor with the Department of Geodesy and Geomatics Engineering, University of New Brunswick and CTO with SceneSharp Technologies Inc. Fredericton, NB, Canada. He has authored or coauthored numerous research articles on image fusion, object detection, and LiDAR. He is an expert in multisensor data fusion, computer vision, and AI-based object detection using multisensor data.



Yun Zhang (Senior Member, IEEE) received the Ph.D. degree in Urban remote sensing and change detection from the Free University of Berlin, Berlin, Germany, in 1997.

He is a Professor in Remote Sensing with the University of New Brunswick, Fredericton, NB, Canada. He was a Visiting Professor with the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, and Research Scientist with the German Aerospace Center (DLR), Cologne, Germany. He has authored or coauthored over 260 research papers. The research outcomes from his laboratory have resulted 15 patents, 3 new patent applications, and more than ten technology licenses. Users of the technologies include NASA, Google, Natural Resources of Canada, DigitalGlobe, and many others.

Dr. Zhang is a fellow of Canadian Academy of Engineering (FCAE). He was a Canada Research Chair in Advanced Geomatics Image Processing. He was the recipient many prestigious national and international research awards.