Automatic Urban Water-Body Detection and Segmentation From Sparse ALSM Data via Spatially Constrained Model-Driven Clustering

Xiaohui Yuan, *Member, IEEE*, and Vaibhav Sarma

Abstract—Identifying hydrological features is important for urban planning and disaster assessment. Data spatial resolution poses challenges in automatic processing. In this letter, we present a novel spatially constrained model-driven clustering method that automatically detects and delineates water bodies in an urban area using airborne laser swath mapping (ALSM) data and imagery. Our method analyzes the modality of the sparseness histogram to decide the existence of water body, followed by clustering. Using the sparseness, clusters are decided by selecting candidate sites. In the iteration of clustering process, new sites are recruited within a close spatial vicinity of the boundary sites. Experiments were conducted using data sets from the city of New Orleans. Our method demonstrated superior robustness regardless of the density of ALSM sample and data discrepancy and very competitive accuracy in comparison with manual tracing, with an overall accuracy above 98%.

Index Terms—Image segmentation, sparse matrices, unsupervised learning, urban areas.

I. INTRODUCTION

HYDROLOGICAL features are key components in predicting flood risks, determining water balance, and estimating geomorphological changes [1]. Conventional approaches rely heavily on manual delineation, and the spatial contour of hydrological region is then georeferenced with elevation description to derive properties of water region [2].

Airborne laser swath mapping (ALSM) provides superior elevation precision, and research has been focused on classification and estimation of hydrological mass [3]. Challenges in ALSM data include the low spatial resolution and insufficient discriminative feature. After postprocessing, the average data-point distance can be as much as 4 m [4], which makes the water and land contact zone difficult to determine. Fig. 1 shows a manually traced water body from ALSM data. Large errors exist close to the boundaries and the regions with inconsistent data. In addition, ALSM systems operate in the near-infrared spectrum, and the absorption rates of water and asphalt are significantly higher than other ground features. Hence, ALSM by itself cannot accurately delineate stream channels or shorelines

The authors are with the Department of Computer Science and Engineering, University of North Texas, Denton, TX 76207 USA (e-mail: xiaohui.yuan@unt.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.
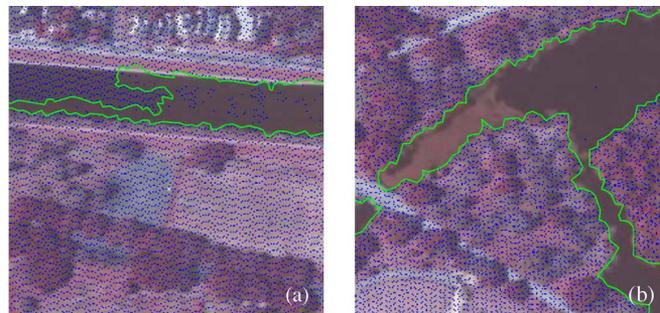
Fig. 1. Superimposed ALSM data on imagery. The contours show manual delineation of water body from ALSM data. (a) Discrepancy in different modalities. (b) ALSM measures of water body are absence.

visible on photographic images, which motivates the integration of other modalities.

Despite the advancement of using multimodal imagery, the integration of ALSM with photographic imagery is hindered by the discrepancies among data sets, including time delay in data acquisition. The time lag causes seasonal changes to appear in the collected data. As shown in Fig. 1(a), the reduction of water exposes the lake bottom that is visible in the image, whereas ALSM reports no such measures, which indicate possible water body. In contrast, Fig. 1(b) shows significant laser returns from water.

In this letter, we present a novel spatially constrained model-driven clustering method that automatically detects and delineates water bodies in urban areas using ALSM data and photographic imagery. Sparseness is computed from ALSM data followed by modality analysis to detect water bodies. Manually traced water bodies are used to construct a bank of models. Clustering is then performed using ALSM and photographic imagery. Sites are recruited within a spatial vicinity, and likelihood of the joint distribution is used for model compliance.

The remainder of this letter is organized as follows: Section II reviews the ALSM technology and the existing methods for water-body segmentation. Section III presents our method in detail. Section IV demonstrates and discusses the performance. Section V summarizes our studies.

II. RELATED WORK

An ALSM system acquires data with irregularly spaced monochromatic near-infrared laser pulse distributed in a pseudorandom fashion. Conversely, photographic images are

intensity values of radiation reflected from a contiguously recorded surface. The differences between these modalities cause complications for data fusion, and some recent studies have quantitatively related ALSM and reflectance data sets [5], [6]. Since the deployment of ALSM systems, research efforts have been devoted to improve the classification. In ALSM data sets, water is characterized as having low intensity values as well as having a low range in elevation or a small deviation from an average terrain model [7]. Brzank *et al.* proposed a method to detect reliable water regions and expand them with the use of elevation and intensity [8]. Local elevation minima were extracted as potential seed zones of the searched water areas followed by region growing. Fuzzy classifier was used to label raw data points [9]. Discrepancies were detected and corrected by comparing height differences between water and nonwater points in a neighborhood.

The availability of multimodal data sets makes integration of ALSM with imagery a favored approach. Mundt *et al.* demonstrated improved classification accuracy by combining images and elevation data [6]. However, disagreement in features extracted from different data sets exists. Antonarakis *et al.* described a method that uses intensity and height data for classification [10]. Empirical thresholds were used. Matgen *et al.* combined flood boundary information with high spatial resolution high-precision light detection and ranging data using multiple linear-regression analysis to allow continuous derivation of remotely sensed water stages [11]. The information required in these algorithms, however, is usually unavailable. In addition, the discrepancy induced from time lag and user initialization requires further investigation.

III. METHODOLOGY

A. Sparseness Metrics

The sparseness Ω is defined as the l^0 norm of the nonzero element count [12]: $\Omega = \|x\|_0 = \#\{i, x_i \neq 0\}$, where x is a signal vector and i denotes a component. This metric, however, is not practical for measuring the sparseness of ALSM data points due to the irregular space between data points. Three alternative metrics are proposed as follows.

1) Window-Based Sparseness: By embedding data points v into a 2-D matrix X, elements with no corresponding ALSM measures are assigned a void value $\emptyset$. Let $\mathcal{W}_{(p,q)}$ denote a $w \times w$ window centered at (p,q). The sparseness $\Omega(p,q)_{\mathcal{W}}$ is the normalized count of the nonvoid elements in $\mathcal{W}_{(p,q)}$

$$\Omega_{\mathcal{W}}(p,q) = \frac{\#\left\{(i,j) X(i,j) \neq \emptyset, X(i,j) \in \mathcal{W}_{(p,q)}\right\}}{w^2 - 1} \quad (1)$$

where $\Omega_{\mathcal{W}} \in [0,1]$. The mapping from the matrix to the irregular data cloud can be achieved with a lookup table.

2) kNN-Based Sparseness: An alternative sparseness can be formulated as an inverse function of the distance to the k nearest neighbors (kNN) Ω_N

$$\Omega_N(p,q) = \frac{1}{\sum_{(i,j)} D\left(A(p,q), A(i,j)\right)} \quad (2)$$

where $A(i,j)$ is one of the k neighbors. The range of Ω_N is in $[1/\mathcal{D}, 1]$, where $\mathcal{D}$ is the farthest distance of two data points.
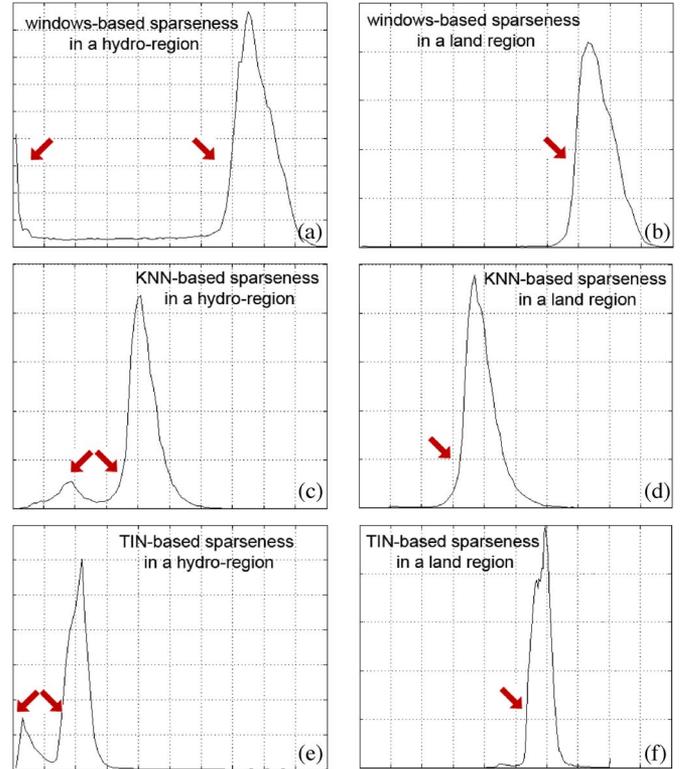

Fig. 2. Sparseness histograms of regions with and without water body.

3) TIN-Based Sparseness: An alternative distance-based sparseness is based on triangulated irregular network (TIN) model. The total distance to the connected data points in the TIN model to represent the sparseness Ω_T is formulated as follows:

$$\Omega_T(p,q) = \frac{l}{\sum_{(i,j)} D\left(A(p,q), A(i,j)\right)} \quad (3)$$

where $A(p,q)$ and $A(i,j)$ share an edge of a triangle and l is the degree of $A(p,q)$.

B. Water-Body Detection Using Modality Analysis

A histogram of sparseness is computed from ALSM data set. Fig. 2 shows typical sparseness histograms for a region with significant amount of water body (left column) and one without water body (right column). For regions that contain significant amount of hydrological features, the histograms show bimodal distributions (as highlighted with arrows in the plots). The left mode in the bimodal distribution is originated from high absorption areas, which implies possible significant water body. The magnitude of the mode indicates the size of the high-absorption areas.

The likelihood of a unimodal distribution can be evaluated with coefficient of bimodality b [13]

$$b = \frac{\gamma^2 + 1}{\beta + \frac{3(n-1)^2}{(n-2)(n-3)}} \quad (4)$$

where γ is the skewness and an estimator of γ is given as $\gamma = (1/n)\sum_{i=1} n(x_i - \bar{x})^3/((1/n)\sum_{i=1}^n (x_i - \bar{x})^2)^{3/2}$, and β is

the kurtosis and an estimator of β is given as $\beta = \sum_{i=1}^{n}(x_i - \bar{x})^4/((1/n)\sum_{i=1} n(x_i - \bar{x})^2)^2 - 3$. n denotes the number of samples, and $\bar{x}$ denotes the mean. Values of b that are less than 0.555 indicate a unimodal distribution with heavy tails; otherwise, it is a bimodal or multimodal distribution.

C. Spatially Constrained Model-Driven Clustering

We use site $s(x, y)$ to represent a data sample at spatial coordinates (x, y) in the feature space of ALSM data set and imagery. The idea of our clustering method is that the feature vectors (i.e., sites) in a closed water region are homogeneous with small variations. Using manually delineated water bodies, feature distributions can be modeled and used to guide automatic segmentation. In our method, multivariate Gaussian functions are used as the model function $G(s; \Sigma, \mu) = (1/T)\exp(-1/2(s - \mu)'\Sigma^{-1}(s - \mu))$, where Σ is the covariance matrix and μ is the mean vector. With manually traced water bodies in the training data, a model bank $\mathcal{G}$, $\mathcal{G} = \{G_1, G_2, \ldots, G_N\}$, is created. Each model G_i represents a distinct water body. Mahalanobis distance [14] is used to evaluate the divergence between models $D_M = ((\bar{x} - \bar{y})^T\Sigma^{-1}(\bar{x} - \bar{y}))^{1/2}$, where $\bar{x}$ and $\bar{y}$ are the mean vectors of the two distributions. Similar models are joined with parameters computed using samples from all regions that were used to derive these models.

In segmentation, our method starts with seed sites with minimum sparseness values ν. The centroids and clusters are assumed independent, and sites in a close spatial vicinity are selected. The selected group of sites are the members of that seed and is denoted with M_{ν_i}

$$M_{\nu_i} = \left\{ s_j D(s_j, s_b = \nu_i) < d, s_j = \arg\min_{s_j} G(s_j) \right\} \quad (5)$$

where s_b denotes sites on the boundary of the cluster M_{ν_i}. The distance d ensures spatial continuity. At initialization, the seed site ν_i is the only boundary site. The boundary sites of M satisfy the following conditions: $s_b \in M_{\nu_i}$ and $\exists s \notin M_{\nu_i}$, such that $D(s_b, s) < 2$. Function $G(s_j)$ is the likelihood of s_j. A model G is selected for ν_i if M_{ν_i} exhibits the greatest joint likelihood

$$G_{\nu_i} = \arg\max_G L(G, M_{\nu_i}) = \arg\max_G G(\nu_i)\Pi_{i=p}^{q}G(s_i). \quad (6)$$

The divergence of M_{ν_i} to the model G_{ν_i} is verified in every iteration using D_M, and the iteration terminates when the divergence exceeds a threshold. Algorithm 1 summarizes our method.

Algorithm 1 Spatially constrained model-driven clustering
1: **INPUT**: image I, seeds ν, and model bank $\mathcal{G}$
2: **OUTPUT**: binary segmentation result W
3: $W \leftarrow 0$
4: **for all** $\nu_i \in \nu$ **do**
5:     Initialize M_{ν_i} using (5)
6:     Select G_{ν_i} from $\mathcal{G}$ using (6)
7:     **for all** $s_b \in M$ **do**
8:         Update M_{ν_i} using (5)

9:         **if** $D_{Maha}(G_{\nu_i}, M_{\nu_i}) > L$ **then**
10:           Continue to the next ν_i
11:         **end if**
12:     **end for**
13:     $W(s) \leftarrow 1, \forall s_i \in M_{\nu_i}$
14: **end for**

Algorithm 2 Efficient search for nearest neighbors
1: **INPUT**: Gridded ALSM X, window size s_0, and k
2: **OUTPUT**: S_N
3: **for all** elements in X **do**
4:     $s \leftarrow s_0$
5:     **if** $X(i, j) = 1$ **then**
6:         **while** $\sum w_s < k$ **do**
7:           $s \leftarrow 2 \times s$
8:         **end while**
9:         **for all** $L(p, q) \in w_s$ **do**
10:           **if** $D_{(i,j),(p,q)} \neq \phi$ **then**
11:             $D_{(i,j),(p,q)} = Dist(X(i, j), X(p, q))$
12:           **end if**
13:         **end for**
14:         $S_N(i, j) = (1/k)D_{(i,j),(p,q)}$
15:     **end if**
16: **end for**

IV. EXPERIMENTS AND DISCUSSION

A. Experimental Data

We used ALSM data set and Digital Ortho Quarter Quads imagery acquired in New Orleans [4]. The specifications of ALSM data include field of view: $40°$, overlap: 30%, acquisition height: 8000 ft, sample spacing: 4 m, vertical accuracy: 0.5–1 ft, and horizontal accuracy: 3–6 ft. The image is projected to the Universal Transverse Mercator 15 North American Datum 83 coordinate. Each pixel is 1 m × 1 m. The ALSM data and imagery are tiled in quarter quadrangle (QQ). Each QQ covers an area of 43 km^2 and contains about 8.5 million data points, i.e., the average point density is about one point per 5 m^2.

B. Efficient Implementation of kNN-based Sparseness

To efficiently identify the nearest neighbors, we constrain our search to a window $\mathcal{W}$. The window size doubles if the number of points in this window is less than k. ALSM data is gridded into a Boolean matrix, where one and zero denote a data value and a void, respectively. Our algorithm is summarized in Algorithm 2. An auxiliary matrix D is used to save pair distances to avoid repeated computation.

C. Performance Analysis

1) Robustness Analysis: Fig. 3 shows two experimental results. Fig. 3(a) shows an area with a mixture of large water body and asphalt rooftop. Two regions are highlighted with boxes and numbered. Region 1 shows buildings with asphalt roofs of which a zoomed-in view is shown on the corner showing
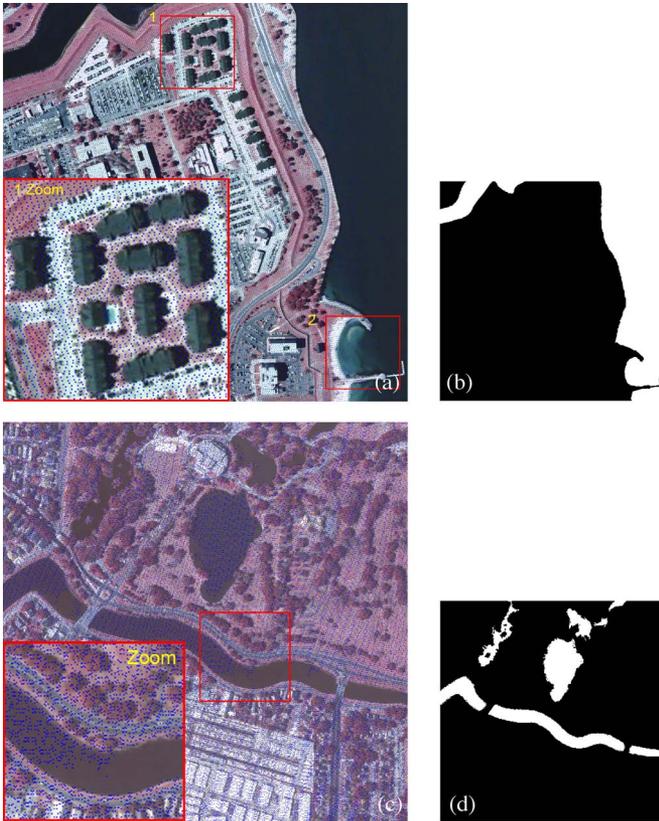
Fig. 3. Exemplar delineation results. (a) and (c) the superimposed ALSM data on images. (b) and (d) the segmentation results.

TABLE I
COMPUTATIONAL TIME (IN SECONDS). SP: TIME ON SPARSENESS; CL: TIME ON CLUSTERING; LONG.: LONGITUDE; LAT.: LATITUDE

	Long.	Lat.	Win-based		kNN-based		TIN-based	
			SP	CL	SP	CL	SP	CL
1	29.9819	-90.0944	8	239	24	273	443	273
2	29.9817	-90.0841	7	80	22	235	46	438
3	30.0264	-90.0942	7	122	45	43	49	83
4	30.0178	-90.1151	8	194	26	113	203	154
5	30.0176	-90.1048	7	118	27	76	46	87
6	30.0083	-90.0947	7	435	58	350	216	551
7	30.0261	-90.0838	7	206	42	313	198	252
8	30.0259	-90.0735	8	79	36	115	113	116
MEAN			7.4	184.1	35	189.8	164.3	244.3
STD			0.5	117.5	12.6	117	134.6	172.3

TABLE II
ACCURACY, PRECISION, AND RECALL OF OUR METHOD

Test Case	Win-based Spa.			kNN-based Spar.			TIN-based Spar.		
	$\mathcal{A}$	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{A}$	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{A}$	$\mathcal{P}$	$\mathcal{R}$
1	99.2	97.6	94.7	99.0	95.3	95.5	99.2	97.4	95.2
2	99.9	96.1	91.1	99.9	91.2	89.9	99.9	95.9	90.5
3	99.6	99.8	99.6	99.7	99.8	99.8	95.1	99.9	93.6
4	99.6	95.7	87.3	99.4	100	91.5	99.4	100	76.6
5	98.9	94.0	85.9	99.2	98.9	86.0	98.5	86.4	85.2
6	97.6	98.5	85.7	97.4	94.2	88.3	98.1	96.8	90.7
7	98.2	97.4	99.4	97.5	99.3	96.4	99.0	99.0	99.2
8	98.7	98.9	94.9	98.1	99.1	92.2	98.4	99.1	93.6
MEAN	99.0	97.3	92.3	98.8	97.2	92.5	98.4	96.8	90.6
STD	0.8	1.9	5.7	1.0	3.2	4.5	1.5	4.4	7.0
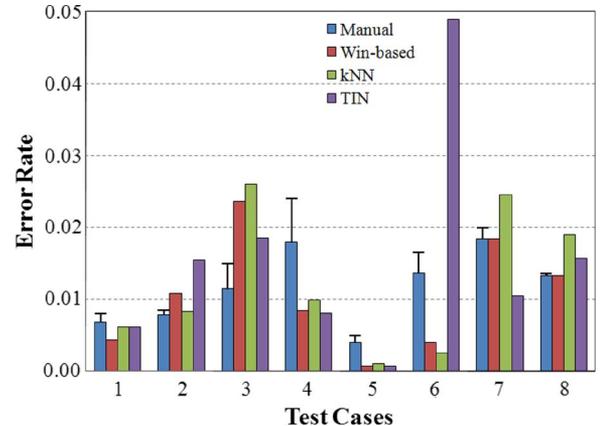
Fig. 4. Error rates of our methods and manual tracing with variance.

that very few ALSM data were collected. Region 2 shows high variation in color for the water body. Both regions are correctly classified, as shown in Fig. 3(b). Fig. 3(c) shows data discrepancy, as highlighted in the box, and the zoomed-in view is displayed on the corner. The density of ALSM data points in the canal implies nonwater region, whereas the canal is clearly shown in the photographic image. The result in Fig. 3(d) shows correct classification.

2) Efficiency Analysis: Our method was implemented with MATLAB 2007a in a PC with Intel Core 2 Duo 3-GHz CPU and 2-GB memory. Table I lists the computational time of eight test cases. The average time of five repetitions is reported, and the times for sparseness computation and the clustering are reported separately. The window-based metric is most efficient, whereas the TIN-based metric takes the longest time. Large variance exists in the time of TIN-based metric, which is due to the degree of the graph. Computational time for clustering is consistent.

3) Accuracy Analysis: We prepared three sets of reference images by manually tracing on images by three specialists independently, and the final reference images were aggregated via averaging and thresholding at two. The absolute difference between two delineations is normalized by the water-body size. In our quantitative evaluation, human operators are mostly consistent, and the variance is less than 1%. Our method was evaluated using the accuracy $(\mathcal{A})$, precision $(\mathcal{P})$, and recall $(\mathcal{R})$: $\mathcal{P} = tp/tp + fp$, and $\mathcal{R} = tp/tp + fn$. Table II presents the results for three sparseness metrics. The overall accuracy is above 98%. However, the kNN-based method achieved slightly

better results with the highest average recall value and the smallest standard deviation.

Fig. 4 shows the error rates in comparison with manual tracing. Due to time lag in data acquisition, data inconsistency is inevitable. In addition, shadows in imagery cause significant change of color. These factors contribute to misclassification, which are the cases 2, 3, 7, and 8 in Fig. 4. Our automatic method achieved competitive accuracy in comparison with manual tracing.

V. CONCLUSION

We have presented a novel spatially constrained model-driven clustering method that automatically detects and

delineates water bodies in urban area using ALSM data and photographic imagery and three sparseness metrics. Experiments were conducted using data and imagery acquired in the city of New Orleans. Our method demonstrated superior robustness in both identifying and correctly segmenting water bodies regardless of the low density of ALSM sample and data discrepancy. The robustness of our algorithm stems from the integration of multimodal information and employment of water-body models. With respect to efficiency, window-based metric takes significantly less time and, on average, the TIN-based metric takes the longest time. Large variance exists in the time expenses of TIN-based metric because many cluster centroids are close to the boundary of the water body. Using manual segmentation results from the three specialists, our automatic method achieved improved and very competitive accuracy ($> 98\%$) consistently. The kNN-based method performed slightly better with the highest average recall value and the smallest standard deviation.

ACKNOWLEDGMENT

REFERENCES

[1] H. Liu and L. Wang, "Mapping detention basins and deriving their spatial attributes from airborne LiDAR data for hydrological applications," *Hydrol. Process.*, vol. 22, no. 13, pp. 2358–2369, Jun. 2008.

[2] G. Schumann, P. Matgen, M. Cutler, A. Black, L. Hoffmann, and L. Pfister, "Comparison of remotely sensed water stages from LiDAR, topographic contours and SRTM," *ISPRS J. Photogramm. Remote Sens.*, vol. 63, no. 3, pp. 283–296, May 2008.

[3] S. K. Lodha, D. M. Fitzpatrick, and D. P. Helmbold, "Aerial LiDAR data classification using AdaBoost," in *Proc. 6th Int. Conf. 3-D Digital Imaging Model.*, Montreal, QC, Canada, Aug. 2007, pp. 435–442.

[4] R. Cunningham, D. Gisclair, and J. Craig, The Louisiana statewide LiDAR project, access in July 2009. [Online]. Available: http://atlas. lsu.edu/central/la_lidar_project.pdf

[5] A. T. Hudak, M. A. Lefsky, W. B. Cohen, and M. Berterretche, "Integration of LiDAR and Landsat ETM data for estimating and mapping forest canopy height," *Remote Sens. Environ.*, vol. 82, no. 2/3, pp. 397–416, Oct. 2002.

[6] J. T. Mundt, D. R. Streutker, and N. F. Glenn, "Mapping sagebrush distribution using fusion of hyperspectral and LiDAR classifications," *Photogramm. Eng. Remote Sens.*, vol. 27, no. 1, pp. 47–54, Jan. 2006.

[7] R. Harris, *Satellite Remote Sensing; An Introduction*. New York: Routledge & Kegan Paul Books Ltd., 1987.

[8] A. Brzank, J. Gopfert, and P. Lohmann, "Aspects of LiDAR processing in coastal areas," in *Proc. Int. Arch. Photogramm. Remote Sens.*, Hanover, Germany, 2005, vol. XXXVI PartI/W3.

[9] A. Brzank and C. Heipke, "Classification of LiDAR data into water and land points in coastal areas," in *Proc. Symp. ISPRS Comm. III—Photogrammetric Computer Vision*, Sep. 2006, pp. 197–202.

[10] A. Antonarakis, K. Richards, and J. Brasington, "Object-based land cover classification using airborne LiDAR," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2988–2998, Jun. 2008.

[11] P. Matgen, G. Schumann, J. B. Henry, L. Hoffmann, and L. Pfister, "Integration of SAR-derived inundation areas, high precision topographic data and a river flow model toward real-time flood management," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 9, no. 3, pp. 247–263, Aug. 2007.

[12] C. C. Took and S. Sanei, "Exploiting sparsity, sparseness and super-Gaussianity in underdetermined blind identification of temporomandibular joint sounds," *J. Comput.*, vol. 2, no. 6, pp. 65–71, Aug. 2007.

[13] S. Institute, *SAS/STAT 9.1 User's Guide*, 1st ed. Cary, NC: SAS Publishing, 2004.

[14] G. J. McLachlan, "Mahalanobis distance," *Resonance*, vol. 4, no. 6, pp. 20–26, Jun. 1999.