

# A Segmentation Method for Tree Crown Detection and Modelling from LiDAR Measurements

José Luis Silván-Cárdenas\*

Centro de Investigación en Geografía y Geomática “Ing. Jorge L. Tamayo” A.C.  
Contoy 137, Lomas de Padierna, Tlalpan, Mexico D.F. 14240  
jlsilvan@centrogeo.org.mx  
<http://www.centrogeo.org.mx>

**Abstract.** A watershed segmentation algorithm is proposed for automatic extraction of tree crowns from LiDAR data to support 3-d modelling of forest stands. A relatively sparse LiDAR point cloud was converted to a surface elevation in raster format and a canopy height model (CHM) extracted. Then, the segmentation method was applied on the CHM and results combined with the original point cloud to generate models of individual tree crowns. The method was tested in 200 circular plots ( $400\text{ m}^2$ ) located over 50 sites of a conservation area in Mexico City. The segmentation method exhibited a moderate to perfect detection rate on 66% of plots tested. One major factor for a poor detection was identified as the relatively low sampling rate of LiDAR data with respect to crown sizes.

**Keywords:** Remote sensing, LiDAR, Watershed segmentation, tree crown detection, 3-d modelling.

## 1 Introduction

Three-dimensional (3-d) forest models are becoming a useful resource for a number of applications including forest management, conservation, inventorying, biophysical modelling, wild fire dynamics, habitat analysis, entertainment, tourism, etc. Despite its importance, databases supporting 3-d models of forests are currently limited in coverage due to lack of effective and efficient methods for automatic extraction of tree crowns.

Airborne light detection and ranging (LiDAR) is a technology used in many countries for ecological applications, forest inventorying, and for single tree-based studies. The basic for modern LiDAR-based forest measurements rely on the acquisition of three surfaces, namely canopy height model (CHM), digital terrain model (DTM) and digital surface model (DSM) [6]. The watershed algorithm from markers is the typical procedure applied to CHMs for extracting individual tree crown of woody vegetation [3,7]. This approach have proved effective even

---

\* The author thanks INEGI and INIFAP for providing the data used in this study.

for leaf-off canopies [2], as long as the point density of the LiDAR data is in the order of several returns per square meters.

The present study sought to develop a simple method for extraction and modelling of vegetation canopies from relatively sparse point clouds (less than one return per square meter) of LiDAR altimetry data. A marker controlled segmentation method was developed and tested for extracting tree crowns and for building simple 3-d models of detected trees. The study built upon prior work on ground filtering [9] and building detection methods [10]. The rest of the paper presents the study site (Section 2), the method description (Section 3), the results for tree-crown detection and modelling (Section 4), and major conclusions (Section 5).

## 2 Study Site and Data Used

The study site corresponds to the so-called conservation ground located in the south of Mexico City (Suelo de Conservación de la Zona Metropolitana de la Ciudad de México). This area is perhaps the single most important element of the water cycle in the region and provides a number of environmental services to the inhabitant of Mexico City, including carbon sequestration, biodiversity conservation, and recreational space. Despite its importance, this area has been threaten by a number of factors that are consequences of an accelerated urban growth, resulting in a significant reduction of its coverage in the recent years [5]. Continuous monitoring and quantification of the amount of vegetation are required for an effective management of the area. Furthermore, digital databases of accurate tree models would be of great value for land managers and environmental authorities that are in charge of preserving this important area. The study reported here aimed at providing the scientific basis for a new methodology to build such a database based on remote sensing, pattern recognition algorithms and geographical information systems supporting 3-d layers.

### 2.1 Field Campaigns

Ground reference data was acquired during a field campaign between February and March of 2010 by people of INIFAP (Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias). Fifty forest stands in the conservation ground were sampled at both tree- and plot-level. The sampling unit used was jointly designed by SEMARNAT (Secretaría del Medio Ambiente y Recursos Naturales) and CONAFOR (Comisión Nacional Forestal) for the national forest/soil inventory [4]. Each conglomerate is composed by four secondary sampling units defined by circular plots arranged in an “upside down Y” shape. Each circular plot has an area of 400 m<sup>2</sup>. The central plot identifies the conglomerate and the other three plots are located at 45.14 m away from the center, at 0, +120 and -120 degrees with respect to the north axis. The location of the central plot was recorded with a GPS of submeter accuracy and, for each circular plot, the scientific and common names of trees were recorded, the diameter at breast height (DBH) and tree height

were measured for each tree with  $DBH \geq 7.5$  cm. Additional qualitative and quantitative information was also recorded in the field (such as number of tree rings, tree age, etc.), but not used in this study.

## 2.2 LiDAR Data

The LiDAR data was acquired with the ALS50-II sensor flown by INEGI (Instituto Nacional de Estadística, Geografía e Informática) between November and December of 2007 over the entire valley. The data had an average horizontal distance of 2.0 m, minimum point density of 0.433 points/m<sup>2</sup> and vertical root mean square error of 7.3 cm. The original point cloud had four pulse returns with intensity value attached, but the latter was not used in this study. The original raw data was provided by INEGI in LAS format. For this study, subsets of 25 ha were extracted around each conglomerate and rasterized at 1-m cell size. All processing were carried out in MATLAB R2010b (The Mathworks Inc.) and results exported to ESRI's shapefile format for visualization in ArcScene and statistical analysis ArcGIS (see Figure 1 for an illustration of major products).

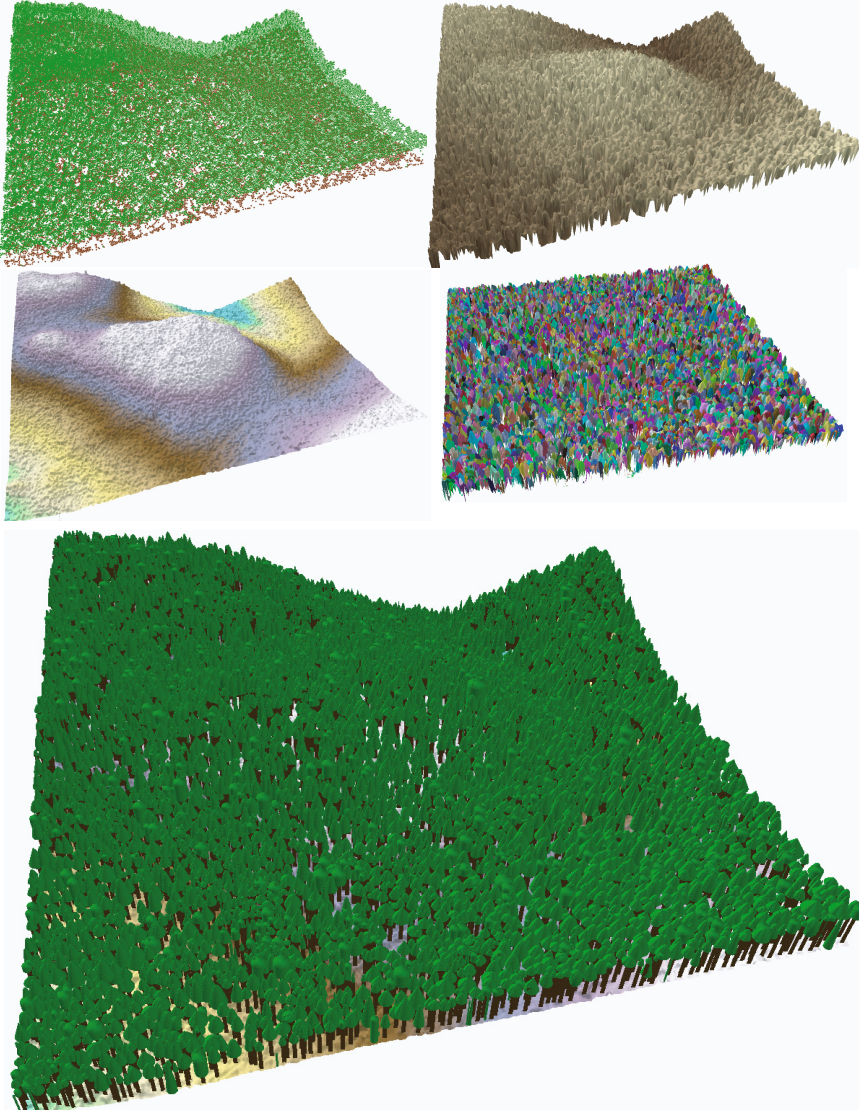
## 3 Methods

### 3.1 Ground Filtering

In absence of non-vegetation features, the CHM is computed by simply subtracting the DTM from the DSM. Otherwise, such difference is sometimes referred to as digital height model (DHM). In turn, the DSM is the grided version of the original point cloud, where each cell is given a value from either the maximum elevation of points within it, or an interpolated value from neighboring points. On the other hand, the DTM is the the interpolated surface from ground points only. Hence, the key processing step here is the so-called filtering of ground points from the LiDAR point cloud.

The ground filtering method used here was introduced in [9], and is based on a wavelet-like decomposition of the DSM surface, the so-called multiscale hermite transform. A recent summary of the theory behind this technique is included in [10]. The product of the ground filtering method is a binary mask indicating ground/non-ground pixels in the DSM. The ground mask is then used to select ground points from which the DTM is interpolated using and interpolation method, such as the cubic convolution interpolation algorithm [8].

In this study, most sites were located in non-urbanized areas. Yet a few sites included building features, which demanded a special treatment. In such cases, the building detection method introduced in [10] was applied to the DHM. The main product of the building detection method is another binary mask of buildings in the DHM which, when combined with the ground mask led to a vegetation mask. The later was applied to the DHM to produce the so-called CHM.



**Fig. 1.** This figure illustrates major products derived from LiDAR data. Top left: classified point cloud (green = vegetation point, brown = ground point); top right: digital surface model (DSM); middle left: digital terrain model (DTM); middle right: segmented crown height model (CHM); bottom: 3-d forest model.

### 3.2 Tree Crown Detection

The basic idea behind the crown detection method is to build regions from the CHM using the classic watershed by flooding approach [1]. Since, the coarse shape of crowns is generally convex, they tend to form watersheds in the CHM with the  $z$ -coordinate reversed. The fine details present in the CHM may, however, cause over segmentation and, therefore, a smoothing is necessary that reduces such an effect. Gaussian filtering has proved an effective procedure for suppressing irrelevant local maxima in treetop detection [3,2]. A binomial filter of order 2, which approximates a Gaussian with  $\sigma = 0.7$ , was used here. In addition to the smoothing, a watershed merging mechanisms was implemented so that barriers are built only when meeting watersheds have an elevation difference greater than a tolerance value  $\Delta z$ , otherwise watershed are merged together. The proposed algorithm can be summarized as follows:

1. Label each minimum with a distinct label and initialize a set  $\mathcal{S}$  with labelled pixels, so that for any  $s \in \mathcal{S}$ , the variables  $\mathbf{elev}(s)$  and  $\mathbf{lab}(s)$  indicate the elevation and the label of  $s$ , respectively.
2. Let  $\mathcal{U}$  be the set of unlabelled pixels with all its lower neighbors in  $\mathcal{S}$ , otherwise stated  $\mathcal{U} = \{u | \forall s \in \mathcal{S} \cap \mathcal{N}_u, \mathbf{elev}(s) < \mathbf{elev}(u)\}$ . If  $\mathcal{U}$  is empty, the procedure ends and  $\mathbf{lab}(\mathcal{S})$  gives the final segmentation. Otherwise, go to Step 3.
3. Label members of  $\mathcal{U}$  as follows. For each  $u \in \mathcal{U}$ :
  - 3.a. Let  $\mathbf{depth}_u(s) \leftarrow \mathbf{elev}(u) - \mathbf{elev}(s)$  for all  $s \in \mathcal{S} \cap \mathcal{N}_u$ ,
  - 3.b. Pick the neighbor from shallowest watershed as  $s^{shallow} = \min_s \mathbf{depth}_u(s)$
  - 3.c. Add  $u$  to  $\mathcal{S}$  with:
 
$$\mathbf{lab}(u) \leftarrow \mathbf{lab}(s^{shallow}), \text{ and}$$

$$\mathbf{elev}(u) \leftarrow \mathbf{elev}(s^{shallow}).$$
4. Merge contiguous shallow watersheds as follows. For each  $u \in \mathcal{U}$ :
  - 4.a. Let  $\mathcal{S}_u = \{s | s \in \mathcal{S} \cap \mathcal{N}_u, \mathbf{depth}_u(s) \leq \Delta z\}$ .
  - 4.b. If  $|\mathcal{S}_u| > 1$ ,
    - 4.b.i: Pick the neighbor from deepest watershed as  $s^{deep} = \max_s \mathbf{depth}_u(s)$
    - 4.b.ii: For all  $s \in \mathcal{S}$  and  $q \in \mathcal{S}_u$  so that  $q \neq p$  and  $\mathbf{lab}(s) = \mathbf{lab}(q)$  set:
 
$$\mathbf{lab}(s) \leftarrow \mathbf{lab}(s^{deep}) \text{ and}$$

$$\mathbf{elev}(s) \leftarrow \mathbf{elev}(s^{deep}).$$
5. Repeat from Step 2.

### 3.3 Tree Modelling

**Parameter Estimation.** The segmentation image was used to select LiDAR points for each tree crown detected (an example of the segmented CHM is shown in Figure 1, middle-right). Since the original  $z$ -coordinate value of crown points accounted for both terrain elevation and tree height, a bare terrain elevation value ( $t$ ) was attached to form 4-D points  $(x, y, z, t)$ , with the  $t$ -value corresponding to the sampled value of the DTM at  $(x, y)$ .

In order to guide the construction of a 3-d model (like the one in Figure 1-bottom), a number of tree parameters were estimated from the selected crown

points. These parameters included, tree location  $(x_0, y_0)$ , base elevation  $(z_0)$ , tree height  $(h)$ , crown depth  $(c)$  and crown radius  $(r)$ . The tree location was calculated as the average  $xy$ -coordinate of crown points. The base elevation of the tree was estimated as the average of terrain elevation values  $(t)$ . The tree height was calculated by subtracting the base terrain elevation from the maximum  $z$ -coordinate value of crown points. The crown depth was calculated by subtracting the average  $z$ -coordinate of crown points in the 1st quartile from the maximum  $z$ -value. Conversely, the crown radius considered points in the 4th quartile only. Although each segment was meant to represent one tree, they were required to enclose a minimum number of points  $(p \geq 3)$ , to represent a crown diameter under a threshold  $(r \leq 10 \text{ m})$ , and to represent a non-flat crown  $(r/c \leq 3/2)$ . Segments that did not satisfy these constraints were neglected.

**Table 1.** Parametric equations of tree crown models expressed in terms of crown radius  $(r)$ , crown depth  $(c)$ , and tree height  $(h)$ , while tree location  $(x_0, y_0, z_0)$  is assumed at the origin

Model name	$x/r$	$y/r$	$z' = (z - h)/c$
cylinder	$\cos \theta$	$\sin \theta$	$-1 \leq z' \leq 0$
sphere	$\cos \phi \cos \theta$	$\cos \phi \sin \theta$	$-0.5 + 0.5 \sin \phi$
cone	$\rho \cos \theta$	$\rho \sin \theta$	$-\rho$
ellipsoid	$\cos \phi \cos \theta$	$\cos \phi \sin \theta$	$-0.5 + 0.5 \sin \phi$
paraboloid	$\rho \cos \theta$	$\rho \sin \theta$	$-\rho^2$
$z$ -paraboloid	$\rho \cos \theta$	$\rho \sin \theta$	$-\sqrt{\rho}$
hyperboloid	$\sinh \phi \cos \theta$	$\sinh \phi \sin \theta$	$(1 + \sqrt{2})(1 - \cosh \phi)$

for  $\theta \in [-\pi, \pi]$ ,  $\phi \in [-\pi/2, \pi/2]$ , and  $\rho \in [0, 1]$

**Tree Crown Model Selection.** Once computed, tree parameters were used to define a parametric model of the crown. The crown shape models considered in this study are represented by the parametric equations provided in Table 1 (see also Figure 2). The model selection was made by minimizing, over all possible models, the total error. In turn, the total error was defined as the sum of absolute distance between crown points on the convex hull and the model surface. Hence, the key for an optimal model selection was the computation of the distance between a point and a surface. Distances to all constituent surface parts, including bottom face (not for sphere), and top face (only for cylinder), were computed and the minimum distance taken. For the non-flat part, the distance was computed as follows.

Let  $\mathbf{p}_i$  be the  $i$ -th crown point, let  $f(\mathbf{p}) = 0$  denote the model surface, and let  $\mathbf{n}_i$  be the normal vector at the closest point to  $\mathbf{p}_i$  in the surface. Since the minimum distance is defined along the normal direction then

$$f(\mathbf{p}_i - \lambda_i \mathbf{n}_i) = 0 \quad (1)$$

where  $|\lambda_i|$  is the minimum distance to the surface  $f$  from the point  $\mathbf{p}_i$ . Considering the convexity of surfaces, the sign of  $\lambda_i$  depends on whether the point falls



**Fig. 2.** 3-d view of tree crown models in Table 1. From left to right: sphere, ellipsoid, cylinder, paraboloid, hyperboloid, cone, and  $z$ -paraboloid.

above or below the surface . Hence, the approach adopted consisted in solving Equation (1) for each off-ground point  $\mathbf{p}_i$ , and for each segment.

In practice, errors from competing models may be too close leading to several equally acceptable model. In such cases, a tie-breaking scheme needs to be implemented. The scheme adopted in this study consisted in considering all the models with an error in the lowest quartile and picking among them the model that defines the most common crown shape. The priority order was set as follows: paraboloid, hyperboloid, ellipsoid, sphere, cone,  $z$ -paraboloid, cylinder. This order was adopted based on the observed crown shape from common species in the area (primarily *Pinus montezumae*, *Alnus jorullensis*, *Pinus patula*, *Abies religiosa*, *Clethra mexicana*, *Schinus molle* and *Pinus hartwegii*). Thus, for instance, if both hyperboloid and cone models yielded errors in the lowest quartile, then the former model was selected as it had higher priority.

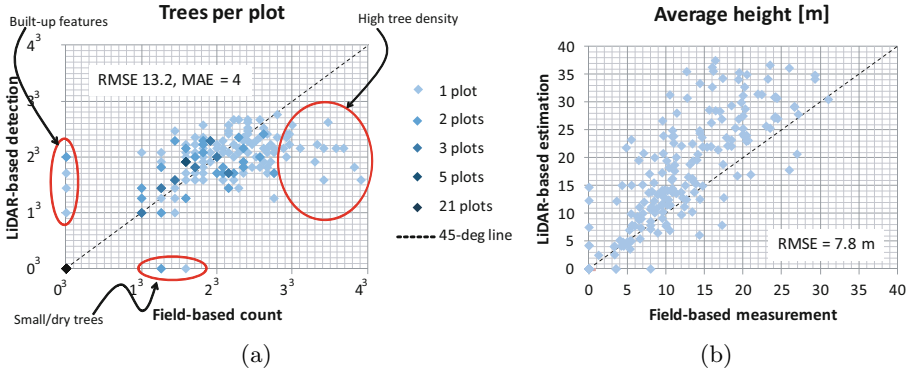
## 4 Results

### 4.1 Detection Rate

The accuracy achieved by the segmentation (with  $\Delta_z = 0.5$  m) was assessed by comparing the number of trees counted in the field against the number of detected trees from the LiDAR data. Sites were first partitioned according to one of three levels of detection rate (percent correct): perfect detection (100% of trees detected), moderate/high detection (more than half and less than twice but not perfect) and low/very low detection (less than half or more than twice detected). Results showed that 38 sites (19%) had a perfect detection rate, 68 sites (34%) had a low/very low detection rate, and the remaining 94 sites (47%) showed moderate/high detection rate.

A scatter plot of LiDAR-based detection against field-based measurements is shown in Figure 3(a). The overall ( $n = 200$ ) root mean square error (RMSE) was 13.2, whereas the median absolute error (MAE) was 4. The later was less sensitive to outliers, which occurred due to three major reasons: 1) due to the presence of non-vegetation features, such as power lines, towers, bridges and flying birds, 2) due to high tree density, and 3) presence of small trees. These three causes of errors were found to be closely related. As a matter of fact, most outliers corresponded to plots where the average DBH was under 24 cm.

The DBH is a good indicator of the size of the tree, both horizontal and vertical. The tree density (number of trees per unit of area) is also related to the



**Fig. 3.** (a) The scatterplot shows the number of trees detected through segmentation of the CHM from LiDAR data (vertical axis) against actual trees counted in the field (horizontal axis) per plot of  $400 \text{ m}^2$ . Both vertical and horizontal axes were transformed by a cubic root for a better distribution. The marker color indicates the number of plots with specific values. Circled set of points identify outliers caused by distinct sources of errors. (b) This scatterplot shows the LiDAR-based estimations of tree height (vertical axis) against field-based measurement (horizontal axis).

DBH, as density generally decreases with the increase of tree size. Also, urban features were present only in areas dominated by small vegetation. Then, if only plots with average DBH over 24 cm were considered ( $n = 109$ ), the RMSE and MAE decreased to 4.8 and 3 respectively. It should be noted that the threshold of 24 cm of average DBH is in close relation with the average point distance of LiDAR data, which was in this case 2 m. This is reasonable since trees with a crown diameter smaller than the average point spacing will be unlikely to be detected. Then, it is also reasonable to think that an increase in the point density would allow a more accurate detection of smaller trees.

## 4.2 3-D Modelling

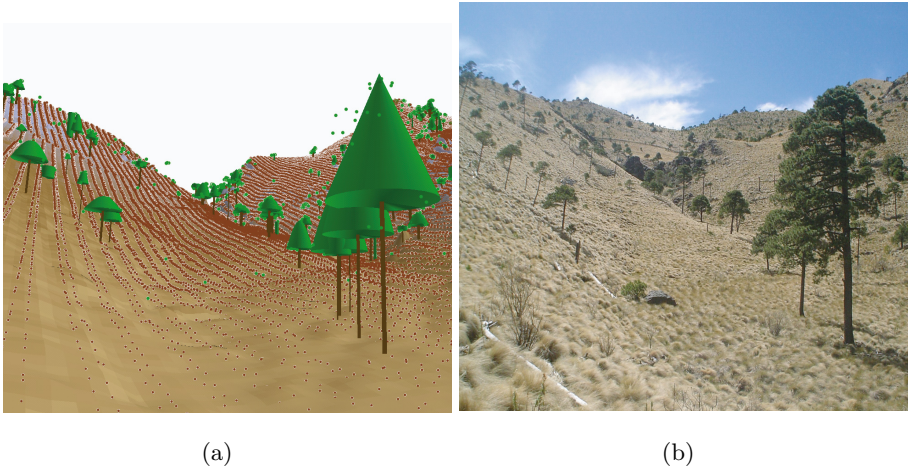
The key parameters for 3-d modelling are tree height, crown diameter and crown shape. However, only tree height was quantitatively assessed using field data. Although the comparison on a per tree basis would have been more valuable, it was not possible since the exact location of trees was not recorded in the field. Hence, a comparison of the average on a per plot basis was performed instead.

Figure 3(b) shows the scatterplot of LiDAR-based estimations plotted against ground-based measurements on a per plot basis. The expected deviation from ground-based values is given by  $\text{RMSE} = 7.8 \text{ m}$ . The scatterplot shows a bias towards overestimation of aboveground height of trees. This was somehow expected since the detection method was shown to fail for small trees, so that small trees did not contribute to the average height, thus resulting in an overestimation of average tree height per plot. The same overestimation bias was observed



when plots with an average DBH under 24 cm were excluded. In fact, the over-estimation bias was more pronounced for the latter case (data not shown) with a RMSE=9.2 m, thus suggesting a substantial mixture of small and big trees in the remaining plots.

Visual inspection of estimated crown diameters and heights on a per tree basis showed a good agreement in general with major errors due to segmentation errors. Figure 4 shows an example of a simulated scene (a) and a photograph of the ground view (b). The front tree in this site was detected as two trees due to separation of branches in the back, thus affecting the estimated crown diameter, depth and shape.



**Fig. 4.** (a) 3-d view of a simulated scene showing the DTM, tree models and ground/non-ground LiDAR points, (b) photo taken during the field campaign of February 2010

## 5 Conclusions

This study proposed and tested a tree crown detection from gridded LiDAR data and presented a simple method for generating 3-d models of forest stands based on the automatic crown detection method. The proposed crown detection method consisted of a watershed by flooding algorithm, which can be efficiently computed in raster format. Accuracy assessment results indicated that the crown detection method is accurate for detecting relatively large trees, with the break point between detectable and non detectable trees determined primarily by the sampling rate of the LiDAR data (average point spacing or point density) and, to some extent, by the tree density. Results also showed that detection can greatly impact the estimated average height as small trees are not accurately detected. Yet, it was recognized that an assessment at individual tree level could yield a more objective result on the estimated tree height parameter.

## References

1. Beucher, S., Lantuéjoul, C.: Use of watersheds in contour detection (1979), <http://cmm.enscm.fr/~beucher/publi/watershed.pdf>
2. Brandtberg, T., Warner, T., Landenberger, R., McGraw, J.: Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in north america. *Remote Sensing of Environment* 85(3), 290–303 (2003)
3. Chen, Q., Baldocchi, D., Gong, P., Kelly, M.: Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogrammetric Engineering and Remote Sensing* 72(8), 923–932 (2006)
4. CONAFOR-SEMARNAT: Manual y procedimientos para el muestreo de campo para el inventario nacional forestal y de suelos (S/F), [http://148.223.105.188:2222/gif/snif\\_portal/secciones/inventarionacional/documentos/ManualMuestreoCampo.pdf](http://148.223.105.188:2222/gif/snif_portal/secciones/inventarionacional/documentos/ManualMuestreoCampo.pdf)
5. Gutiérrez, M.T., González, J., Zamorano, J.: La Cuenca de México y sus cambios demográfico-espaciales. In: *Temas Selectos de Geografía*, p. 155. Instituto de Geografía, UNAM, México (2005)
6. Hyypä, J., Hyypä, H., Yu, X., Kaartinen, H., Kukko, A., Holopainen, M.: Topographic Laser Ranging and Scanning. In: *Forest Inventory Using Small-Footprint Airborne LiDAR*, ch. 12, pp. 335–370. CRC Press (2009)
7. Hyypä, J., Kelle, O., Lehtinen, M., Inkinen, M.: A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing* 39(5), 969–975 (2001)
8. Keys, R.: Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29(6), 1153–1160 (1981)
9. Silván-Cárdenas, J.L., Wang, L.: A multi-resolution approach for filtering LiDAR altimetry data. *ISPRS Journal of Photogrammetry and Remote Sensing* 61(1), 11–22 (2006)
10. Silván-Cárdenas, J.L., Wang, L.: Extraction of Buildings Footprint from LiDAR Altimetry Data with the Hermite Transform. In: Martínez-Trinidad, J.F., Carrasco-Ochoa, J.A., Ben-Youssef Brants, C., Hancock, E.R. (eds.) *MCPR 2011*. LNCS, vol. 6718, pp. 314–321. Springer, Heidelberg (2011)