A Multiscale Erosion Operator for Discriminating Ground Points in LiDAR Point Clouds

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. Digital terrain models (DTM) are basic products required for a number of applications and decision making processes. Nowadays, high spatial-resolution DTMs are primarily produced through airborne laser scanners (ALS). However, the ALS does not directly deliver DTMs but a dense cloud of 3-d points that embeds both terrain elevation and height of natural and human-made features. Such a point cloud is generally rasterized and referred to as the digital surface model (DSM). The discrimination of aboveground objects from terrain, also termed ground filtering, is a basic processing step that has proved especially difficult for large areas of complex terrain characteristics. This paper presents the development of a multiscale erosion operator for removing aboveground features in the DSM, thus producing a surface that is close to the DTM. Such an approximation was used to separate ground from non-ground points in the original point-cloud and the discrimination accuracy was assessed using publicly available data. Results indicated an improvement over a previously published method.

Keywords: Remote sensing, LiDAR, Ground filtering, multiscale Hermite transform.

1 Introduction

The small-footprint, discrete-return light detection and ranging system (LiDAR) has become one of the most important means to produce high-resolution DTM data. This has been due in part to a number of advantages over competing aerial photogrammetric techniques, such as independence of sun light, higher vertical accuracy, less missing data by occlusion, low redundancy and because it does not rely on existence of textured surfaces and discontinuities for a successful point matching [1]. The discrete return LiDAR system is based on the accurate measurement of the elapsed time between emitted and backscattered laser pulses. The emitted pulse is typically short-time and unimodal, whereas

^{*} The author thanks the ISPRS for providing the LiDAR data used in this study.

J.A. Carrasco-Ochoa et al. (Eds.): MCPR 2013, LNCS 7914, pp. 213–223, 2013. © Springer-Verlag Berlin Heidelberg 2013

the backscattered may spread over longer times exhibiting multiple modes called returns. Returns are associated to distinct layers the laser interacted with. For each return, the 3-d position is determined and its intensity recorded. Hence the raw LiDAR data consists of a dense cloud of 3-d points with associated return intensity. Each point provides the location where the laser hit the Earth's surface during the scanning process, whereas the intensity is a digital representation of the fraction of pulse energy reflected at that location.

The z-coordinate of points correspond to terrain elevation with respect to a horizontal datum, typically the mean see level, plus the height of non-terrain features in some instances. In order to produce a DTM by interpolation of ground points, a discrimination of ground from non-ground points must be carried out first. This discrimination process is referred to as ground filtering and is generally considered a preprocessing for generating not only the DTM, but also the height of non-ground components [2, 1]. Further classification of non-terrain points into meaningful features, such as trees, buildings, roads, and so on, is also common task that is needed for detailed analysis and quantification of landscape characteristics.

Methods to ground filtering can be grouped into two big categories, namely point-based and raster-based methods. Methods in the first category classify directly the point cloud whereas methods in the second category first rasterize the point cloud onto a regular grid through an interpolation method. Each approach has advantages and disadvantages. In general, rasterizing the data first allows to take advantage of digital image processing algorithms which run much faster than point-based operations, whereas point-based processing tends to be more accurate [2]. A recent review of methods and critical issues of the ground filtering problem is provided in [3]. Interestingly, the need for processing at multiple scales/resolutions has been increasingly recognized by several studies [4–6]. This has let to the adoption of multiresolution image decomposition techniques in raster-based approaches by some authors [4, 7]. These algorithms tend to be very efficient, less sensitive to parameter selection, less sensitive to point density and are able to remove large non-ground features, such as bridges and complex buildings.

The method presented in this paper falls in the raster-based category and is based on an wavelet-like transform termed the multiscale hermite transform (MHT), which is an overcomplete signal decomposition based on scaled and rotated Gaussian derivatives [8]. The original filtering method was introduced in [7] and, based on an extensive accuracy assessment, was ranked among the top three out of nine algorithms tested. Furthermore, its performance has been demonstrated in several real applications [9–11]. The rationale of the method is as follows. A DSM is first generated through a point-to-raster conversion of the point cloud. Then the DSM is decomposed using the MHT expansion with transform coefficients processed through the so-called erosion operator, so that when applying the inverse transform, the re-synthesized surface corresponds to an eroded version of the DSM. The resulting surface is then used to generate a ground/non-ground mask by comparing the original DSM with its eroded

version. The overall process is illustrated in Figure 1, but major improvement presented in this paper corresponds to the "Multiscale Erosion Operation" step. A more complete description of the overall algorithm is to appear in [12].

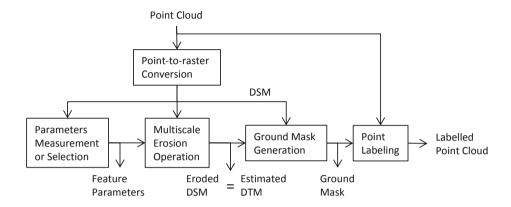


Fig. 1. Flow diagram of ground filtering method. Adapted from [12]

The rest of the paper is organized as follows. Section 2 provides a summary of the theoretical background around the MHT. Then, in Section 3 the multiscale erosion operator is defined and illustrated. Next, in Section 4 some filtering tests are described and results of the accuracy assessment presented. Conclusions drawn from the study are presented in Section 5.

2 Background

The MHT can be defined for one, two and higher dimensions and for both continuous and discrete signals, here we summarize the theory for two-dimensional, discrete signals, and refer the interested reader to the original source [8] for other forms. For simplicity, exposition starts with the single-scale case and then extend the result to multiple scales.

2.1 Single-Scale Discrete Hermite Transform

The discrete Hermite transform (DHT) of a signal z(x, y) defined on a discrete domain can be expressed as a convolution of the input signal with a bank of 1-d filters along each dimension and followed by a subsampling with a rate factor of two. This is expressed in mathematical terms as follows:

$$z_{n,m}(p,q) = \sum_{x,y} z(x,y)b_n(x-2p)b_m(y-2q)$$
 (1)

where

$$b_n(x) = 2^{-N} \sqrt{C_N^n} \sum_{j=0}^n (-1)^{n-j} C_n^j C_{N-n}^{x+N/2+j-n}$$
(2)

for $x = -N/2, \ldots, N/2$ and by $b_n(x) = 0$, for all other values of x, represent the bank of 1-d filters for $n = 0, \ldots, N$. The set of such filters is referred to as the binomial family. Since the binomial family is the discrete counterpart of Gaussian derivatives, a transform coefficient $z_{n,m}(p,q)$ approximates (up to a normalizing factor) the partial derivative of order n with respect to x and of order x x and order x and of order x and of order x and of order x and order x and of orde

Conversely, the full set of transform coefficients allows recovering the input signal through the inversion procedure summarized in the following equation:

$$z(x,y) = \sum_{n,m=0}^{N} \sum_{p,q} z_{n,m}(p,q)\tilde{b}_n(2p-x)\tilde{b}_m(2q-y)$$
 (3)

where the interpolation filters $\tilde{b}(x) = 2b(-x)$, for n = 0, 2, ..., N, are binomial filters reflected around the origin.

In general, the DHT expansion compacts most of the signal information in the first few coefficients. Hence, a near perfect reconstruction can be obtained with a truncated expansion. Further compaction of the information can be achieved through local orientation of the coordinate axis along the strongest signal variation. More specifically, the expansion coefficients can be defined with respect to a coordinate system (u, v), that has been rotated by an angle θ with respect to the original coordinate system (x, y). Since different orientation angles are chosen for different sampling points, this operation is referred to a local spatial rotation. Figure 2 shows an example of original versus rotated coefficients.

2.2 Local Spatial Rotation

Let $z_{n,m}^{(\theta)}$ denote the rotated coefficients at a generic sampling location, where location is purposely omitted in the notation for simplicity. This can be computed as linear combinations of the original (un-rotated) coefficients through the following mathematical relation:

$$\frac{z_{n-m,m}^{(\theta)}}{\sqrt{C_n^m}} = \sum_{k=0}^n a_{m,k}^{(\theta,n)} \frac{z_{k,n-k}}{\sqrt{C_n^k}} \tag{4}$$

The recovery of the original coefficients from the rotated ones is performed through rotation by a negative angle. The angle functions $a_{m,k}^{(\theta,n)}$ correspond to the generalized binomial filters (GBF), a family of discrete sequences with parameters n and θ , which are given by

$$a_{m,k}^{(\theta,n)} = s^k c^{-k} \Delta^m \{ C_{n-m}^{k-m} c^{2k-m} s^{n-2k+m} \}$$
 (5)

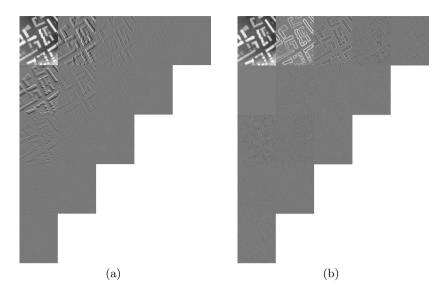


Fig. 2. Example of DHT coefficients up to fourth order (a) and corresponding rotated coefficients (b). In both cases, the order of derivation with respect to x- and y-coordinates grows from left to right and from top to bottom, respectively.

for $m,k=0,\ldots,n$, and $c=\cos\theta$ and $s=\sin\theta$. The first few GBFs can be expressed using the matrix notation $\mathbf{A}^{(\theta,n)}=[a_{m,k}^{(\theta,n)}]_{m,k=0,\ldots,n}$, so that

$$\begin{pmatrix} s & c \\ c & -s \end{pmatrix}, \begin{pmatrix} s^2 & 2sc & c^2 \\ sc & c^2 - s^2 - sc \\ c^2 & -2sc & s^2 \end{pmatrix}, \begin{pmatrix} s^3 & 3s^2c & 3sc^2 & c^3 \\ s^2c - s^3 + 2sc^2 - 2s^2c + c^3 - sc^2 \\ sc^2 - 2s^2c + c^3 & s^3 - 2sc^2 & s^2c \\ c^3 & -3sc^2 & 3s^2c & -s^3 \end{pmatrix}$$

correspond to the matrices for n = 1, 2 and 3, respectively.

In all the examples presented here, the rotation was set to the orientation of the local gradient, that is:

$$\theta = \arctan \frac{z_{0,1}}{z_{1,0}} \tag{6}$$

This selection of the angle makes the rotated coefficients $z_{0,1}^{(\theta)}=0$, and

$$z_{1,0}^{(\theta)} = \sqrt{z_{1,0}^2 + z_{0,1}^2} \tag{7}$$

The latter corresponds (up to a constant factor) to the magnitude of the local gradient. Other angles may be computed that involve higher order coefficients, however the gradient angle generally achieves higher energy compaction of the signal along the first coordinate. This is the case if the input signal embeds strongly oriented features, such as the edges of buildings in a DSM (see Figure 2).

2.3 Multiscale Discrete Hermite Transform

The multiscale discrete Hermite transform (MDHT) is implemented by recursively replacing the zero-order (low-pass) coefficient of previous level, $z_{0,0}^{(k-1)}(x,y)$, by its DHT expansion times a scaling factor, thus yielding a wavelet-like pyramidal decomposition.

Mathematically, the MDHT with K+1 pyramid levels is expressed through:

$$z_{n,m}^{(0)}(p,q) = \sum_{x,y} z(x,y)b_n(x-2p)b_m(y-2q), \quad \text{for } k = 0$$
 (8)

and

$$z_{n,m}^{(k)}(p,q) = \sum_{x,y} z_{0,0}^{(k-1)}(x,y)c_n(x-2p)c_m(y-2q), \quad \text{for } k = 1, \dots, K$$
 (9)

for n, m = 0, ..., N and $c_n(x) = (\sqrt{3}/2)^n b_n(x)$. Conversely, the reconstruction of the original signal is carried out through successive reconstruction of low-pass coefficients from the coefficients of the upper pyramid levels. This is expressed as:

$$z_{0,0}^{(k-1)}(p,q) = \sum_{n,m=0}^{6} \sum_{i,j} z_{n,m}^{(k)}(i,j)\tilde{c}_n(2i-p)\tilde{c}_m(2j-q), \quad \text{for } k = K,\dots,1$$
 (10)

and

$$z(x,y) = \sum_{n,m=0}^{8} \sum_{p,q} z_{n,m}^{(0)}(p,q)\tilde{b}_n(2p-x)\tilde{b}_m(2q-y)$$
(11)

with
$$\tilde{c}_n(x) = (\sqrt{3}/2)^{-n} \tilde{b}_n(x)$$
.

It should be noted that the replaced low-pass coefficients are not part of the MDHT expansion, but only used for the computation of the next coarser level. In contrast, since the coarsest low-pass coefficient is not replaced, it has to be considered part of the MDHT expansion. Also, the length of binomial filters of the first pyramid level is N=8, whereas that for higher levels is N=6. These filter length values ensure that the low-pass coefficients comprise a scale-space representation in normalized spatial coordinates for a discrete scale sequence. Such a representation is equivalent to filtering the original signal with a sequence of low-pass binomial filters of length $N_k=8\times 4^k$, for $k=0,\ldots,K$, and then resample the output by a factor 2^k . Because of this, the MDHT can also be interpreted as an expansion of the Gaussian pyramid in terms of its derivatives. Furthermore, the MDHT coefficients can also be rotated using the same formulas as stated previously. In this case, the notation is extended to $z_{n,m}^{(\theta,k)}$, which accounts for the rotation angle.

3 The Multiscale Erosion Operator

The multiscale erosion operator is a processing of the coefficients that allows to remove aboveground features in the transform domain. This can be seen as a local spatial shifting of the signal, a process that is performed only around detected ground/non-ground transition pixels, so that the portion of the signal in the vicinity of a ground/non-ground transition is replaced by a portion of the signal located along neighbor ground points.

The local spatial shifting is based on a property of the binomial filters stating that the members of a binomial filter family can reconstruct the members of a binomial filter family at a shifted location and decreased filter length. Otherwise stated:

$$\frac{b_n(x - M/2; N - M)}{\sqrt{C_{N-M}^n}} = \sum_{m=0}^M (-1)^m C_M^m \frac{b_{n+m}(x; N)}{\sqrt{C_N^{n+m}}}$$
(12)

for M < N, where shifting occurs both in scale (N) and space (x). Since the DHT is a linear transform, a similar relation can be obtained for the DHT coefficients in 1-d. For 2-d and higher-dimensional signals, the spatial shifting operation can be combined with the rotation operation, thus resulting in a directional spatial shifting. In particular, for 2-d signals the spatial shifting along a direction defined by an angle θ is expressed through:

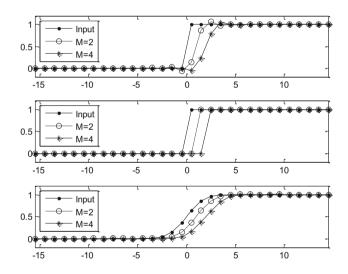


Fig. 3. These plots illustrate the effect of the shifting parameter (M) in reconstructed surface profiles (circles and stars). The input profile (dots) corresponds to a sharp step for the first and second case and to a smooth step for third case. The DHT of the input was computed with N=8 and the transform coefficients were processed with a scale-space shifting of Eq. (13) for M=2 and M=4. Then, the profiles were reconstructed from processed coefficients with the inverse DHT using the same filter length (N=8) for the first and third cases, but a decrease filter length (N-M) for the second case.

$$\frac{z_{n,l}^{(\theta,N-M)}(p-M/2,q)}{\sqrt{C_{N-M}^n}} = \sum_{m=0}^M (-1)^m C_M^m \frac{z_{n+m,l}^{(\theta,N)}(p,q)}{\sqrt{C_N^{m+m}}}$$
(13)

for n = 0, ..., N - M and l = 0, ..., N, where now the filter length is included in the notation of the coefficients to make explicit the shifting in scale as well¹. Figure 3 illustrates the effect of the shifting parameter M on a reconstructed profile from its shifted transform coefficients. The shifting in scale can be ignored only if the signal is smooth enough (bottom panel). However, for sharp transitions the inverse transform requires to account for the scale shifting (middle panel), or otherwise signal distortions may occur (upper panel) and this is accentuated with increasing values of M.

Of special interest is the case for N=8, M=2 and n=l=0, for which the above equation reduces to:

$$z_{0,0}^{(\theta,6)}(p-1,q) = z_{0,0}^{(\theta,8)}(p,q) - \frac{z_{1,0}^{(\theta,8)}(p,q)}{\sqrt{2}} + \frac{z_{2,0}^{(\theta,8)}(p,q)}{\sqrt{28}}$$
(14)

which would correspond to the erosion operator developed in [7] if the scale shifting and the third term in the left-hand side were neglected. In that study, such an erosion operator was derived from an explicit model of a 1-d surface profile in the continuous domain, and then extrapolated for the 2-d discrete case. In contrast, the formula of Equation (13) represents a more general way as it allows shifting the higher order coefficients as well, instead of simply setting them to zeroes as in the previous work.

Additionally, in the original formulation, the terrain elevation underneath non-ground features was assumed flat. This was convenient because the erosion operator is essentially a local spatial shifting. In order to account for higher order variations of terrain elevations at those sites, a truncated Taylor expansion is also proposed in this study. Specifically, the following approximation was used to estimate terrain elevations at transition points:

$$z_{n,m}^{(\theta,N)}(p,q) \approx z_{n,m}^{(\theta,N)}(p-M/2,q) + c_1 z_{n+1,m}^{(\theta,N)}(p-M/2,q)$$
 (15)

with the constant $c_1 = (M/4)\sqrt{n+3}$, which accounts for the relation between the transform coefficients and the signal derivatives. Hence, the new erosion operator is implemented by plugging the local spatial shifting of Equation (13) into Equation (15).

In all tests presented below the shifting parameter was set to M=2 and the Equation (15) applied to the coefficients $z_{0,0}^{(\theta,k)}$ and $z_{1,0}^{(\theta,k)}$ at detected ground/nonground transitions, while all other coefficients are set to zero. Such a processing was inserted in the computation of the MDHT, so that erosion is carried on the coarser layers. The processed coefficients are then used to synthesize a surface

¹ It should be noted that the scale rather than a scale index is used here.

that approximates the DTM, which in turn was used to classify the original point cloud into ground and non-ground points. Further details on both the detection of ground/non-ground transition points, as well as on the insertion of the erosion operator in the MDHT, are provided in [7].

4 Ground Filtering Tests

Both the previously developed and the new erosion operators were applied to the ISPRS datasets for two sites (www.commission3.isprs.org/wg3/). The first site corresponds to an urban area (CSite2) exhibiting large and irregular shaped buildings, as well as a road with bridge and small tunnel. The second site corresponds to a rural area (FSite5) with vegetation on steep slopes, quarry, vegetation on river banks and gaps. The original point cloud included elevation and intensity from first and last returns; however, only first return was used in the tests as it was generally the cleanest measurement.

Parameter	CSite2	Fsite5
Maximum Feature Width [m]	100	50
Maximum Terrain Elevation Difference [m]	30	90
Maximum Terrain Slope [Deg]	25	55
Maximum Elevation Tolerance [m]	0.25	0.25
Cell Size [m]	1	2

Table 1. Selected filter parameters for each test site

The parameters used for the filtering of each dataset are provided in Table 1. Cell sizes for point-to-raster conversions were roughly equivalent to the original average point spacing. Further details on the interpolation of the data are available in the reference [7]. Other parameters were interactively measured from visualizations of the DSM as suggested in [12]. Nonetheless, although a very steep slope of 88 degrees was measured for the forested area, a value of 55 degrees was used instead in order to avoid leaving most non-ground features unfiltered.

The accuracy assessment of the filtering results was carried out for up to four test insets per site, for which classification was available. Ground points from the reference data were interpolated to generate a reference DTM that served to compute the root mean square error (RMSE) of the estimated DTM (eroded DSM) (Table 2(a)). The latter was also used to classify the original point cloud and the overall classification accuracy (OCA) was computed for each sample set (Table 2(b)).

In general, the new method showed lower RMSE for both sites, thus indicating a superior method for estimating the DTM. In terms of classification accuracy, the new method was still better than the old one in the average, yet there were a few cases were the old method remained comparable or even superior to the new method. In either case, results for the urban site showed significant commission

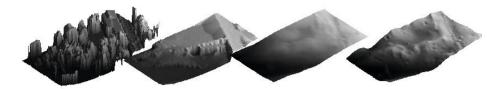


Fig. 4. This figure illustrates the surfaces derived for one of the test samples (samp24). From left to right: Input DSM, reference DTM, synthesized DTM with the old erosion method and synthesized DTM with the new erosion method.

Table 2. Root mean square error in meters of estimated DTM (a) and overall accuracy of ground/non-ground classification (b) for each test site and method

(a)					
Sample	CSite2	Fsite5			
Set	Old New	Old New			
	4.12 4.07				
	$6.85\ 4.26$				
	$10.41\ 8.86$				
average	$6.92\ 5.74$	6.28 3.84			

(b)							
Sample	CSit	e2	Fsit	e5			
Set	Old	New	Old	New			
samp*1	70	68	77	75			
samp*2		81	69	86			
samp*3	78	84	62	82			
samp*4	65	74	85	79			
average	73	77	73	81			

errors along the eastern side due to the edge effect of the multiscale filtering and the terrain elevation gradient towards the eastern direction, yet this was less significant than in the tests reported in previous work. The reason for this was a change from symmetric reflection to an anti-symmetric reflection of the boundary cells. According to this condition, the extension of a sloppy terrain near the edge will maintain its slope rather than change it with a symmetric reflection. For the forested site, the major problem was the omission errors along the southern side, where a sharp terrain shape was present.

5 Conclusions

This paper presented a filtering method based on a multiscale signal decomposition termed the multiscale Hermite transform, which has been formulated in the context of the scale-space theory for signal processing [8]. As the overall filtering method had been extensively tested and applied elsewhere [7, 9–11], the main focus of this work was on an extension of the erosion operator used in that method. Such an extension used a truncated Taylor expansion and a local spatial shifting operator in the transform domain for predicting terrain elevation underneath non-ground features, such as buildings and trees. When compared with the original method, it was found that synthesized DTM was more accurate than for the previous method, which also led to a better discrimination of ground points from the original point cloud. Although the synthesized DTM relied on

coefficients of up to first order only, there were involved coefficients of up to fourth order through the scale-space shifting operation, which was a major reason for the reconstruction of more details. Future work should develop practical ways for parameter selection and investigate the effects of incorporating even higher order coefficients.

References

- Pfeifer, N., Mandlburger, G.: 11. In: LiDAR Data Filtering and DTM Generation, pp. 307–333. CRC Press, Boca Raton (2009)
- 2. Axelson, P.: Processing of laser scanner data algorithms and applications. ISPRS Journal of Photogrametry and Remote Sensing 54(2-3), 138–147 (1999)
- 3. Meng, X., Currit, N., Zhao, K.: Ground filtering algorithms for airborne lidar data: A review of critical issues. Remote Sensing 2(3), 833–860 (2010)
- Thuy, T., Tokunaga, M.: Filtering airborne laser scanner data: A wavelet-based clustering method. Photogrammetric Engineering and Remote Sensing 70(11), 1267–1274 (2004)
- Chen, Q., Gong, P., Baldocchi, D., Xie, G.: Filtering airborne laser scanning data with morphological methods. Photogrammetric Engineering & Remote Sensing 73(2), 175–185 (2007)
- Evans, J.S., Hudak, A.T.: A multiscale curvature algorithm for classifying discrete return lidar in forested environments. IEEE Transactions on Geoscience and Remote Sensing 45(4), 1029–1038 (2007)
- Silván-Cárdenas, J., Wang, L.: A multi-resolution approach for filtering LiDAR altimetry data. ISPRS Journal of Photogrammetry and Remote Sensing 61(1), 11–22 (2006)
- 8. Silván-Cárdenas, J., Escalante-Ramírez, B.: The multiscale Hermite transform for local orientation analysis. IEEE Transactions on Image Processing 15(5), 1236–1253 (2006)
- Silván-Cárdenas, J., Wang, C.W.L., Rogerson, P., Feng, T., Kamphaus, B.: Assessing fine-spatial-resolution remote sensing for small-area population estimation. International Journal of Remote Sensing 31(21), 5605–5634 (2010)
- 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)
- Silván-Cárdenas, J.L.: A segmentation method for tree crown detection and modelling from lidar measurements. In: Carrasco-Ochoa, J.A., Martínez-Trinidad, J.F., Olvera López, J.A., Boyer, K.L. (eds.) MCPR 2012. LNCS, vol. 7329, pp. 65–74. Springer, Heidelberg (2012)
- Silván-Cárdenas, J., Wang, L.: A multiscale approach for ground filtering from LiDAR altimetry measurements. In: Scale Issues in Remote Sensing. John Wiley and Sons (to appear, 2013)