Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
ISPRS Journal of Photogrammetry & Remote Sensing 61 (2006) 11–22
PHOTOGRAMMETRY & REMOTE SENSING
www.elsevier.com/locate/isprsjprs

A multi-resolution approach for filtering LiDAR altimetry data

J.L. Silván-Cárdenas [a,b,*], L. Wang [a]

[a] Texas State University-San Marcos, Department of Geography, 601 University Dr., ELA 139, San Marcos, TX 78666, US
[b] Centro de Investigación en Geografía y Geomática, Ing. "Jorge L. Tamayo", Contoy 137, Lomas de Padierna, Mex. D.F.

Received 12 February 2006; received in revised form 8 June 2006; accepted 14 June 2006
Available online 24 July 2006

Abstract

Discrimination of above-ground objects from terrain has proven to be surprisingly difficult to automate in computers, especially for large areas of varied terrain characteristics. Several methods have been developed for filtering the LiDAR data, of which three approaches are more prevalent: linear prediction, slope based and morphological filtering. A common ground to all these approaches is that the range of scales at which feature variations exist tends to be smaller than the range of scales at which terrain variations exist. In this paper, a model-based approach is described in which multiscale gradient of the surface variation is computed and used to adaptively erode the gridded LiDAR data within a multi-resolution, analysis–synthesis framework, namely the multiscale Hermite transform (MHT). The method was tested over nineteen datasets, including urban and forest areas. An average coefficient of agreement was computed over all datasets and compared with that obtained from other methods. Results showed that the proposed method was within the top three among nine methods tested.

Keywords: Laser scanning; Digital terrain model; Digital surface model; Scale–space representation; Multiscale Hermite transform

1. Introduction

High-spatial resolution digital terrain models (DTM) are of prime importance for many forestry and urban applications, including flood control, road design, forest management, as well as urban planning and management. Airborne Light Detection and Ranging (LiDAR) is now a widely used technology for the high-spatial resolution measurement of the earth surface. LiDAR systems deliver dense sets of three-dimensional points of ground elevation plus the height of above-ground features such as buildings, trees and cars. The terrain surface is generated through the interpolation of ground points (see e.g. Lloyd and Atkinson, 2002). Before the interpolation process can be run, non-ground points need to be eliminated from the entire LiDAR data, a process referred to as filtering (Kraus and Pfeifer, 1998; Lohmann et al., 2000; Vosselman, 2000).

The surface elevation data (z) delivered by the LiDAR system can be seen as a superposition of three components, namely terrain elevation (t), above-ground feature height (f) and error term (e). Otherwise stated,

$$z = t + f + e \tag{1}$$

where e accounts for vertical errors that may be due to several factors (e.g., measurement deviations due to instrument calibration, quantization errors due to finite-precision processing and/or storage, etc.). Thus, the

* Corresponding author. Texas State University-San Marcos, Department of Geography, 601 University Dr., ELA 139, San Marcos, TX 78666, US.
E-mail addresses: jlsilvan@txstate.edu (J.L. Silván-Cárdenas), lewang@txstate.edu (L. Wang).

0924-2716/$ - see front matter
doi:10.1016/j.isprsjprs.2006.06.002

filtering problem of LiDAR data implies its disaggregation on the above components. Evidently, even in the case of a negligible error component, the problem is ill-conditioned as there are more unknown variables than equations. Therefore, no unique solution exists, and different solutions will differ in the way unknown variables are further constrained. The general guidance for building *good* constraints should be that the different components have essentially different spatial structures. While discrimination of objects with slightly different spatial structures embedded in the LiDAR data is more or less easily carried out by human beings, it has proven to be surprisingly difficult to automate in computers. Therefore, developing efficient and effective methods for terrain filtering is currently an active topic of research.

In the past, several methods have been developed for filtering the LiDAR data within which linear prediction (Kraus and Pfeifer, 1998), slope/height-difference (Roggero, 2001; Shan and Sampath, 2005; Sithole, 2001; Vosselman, 2000) and morphologic filters (Kilian et al., 1996; Zhang et al., 2003) rank among the most popular approaches. A common ground to all these approaches is that the range of scales at which feature variations exist tend to be smaller than the range of scales at which terrain variations exist; however, explicit scale estimation has not been considered yet in the filtering problem.

In this work, we investigate how the scale–space theory (Witkin, 1984) — widely used in the fields of digital image processing and computer vision — can be tailored to extract the terrain profile from the LiDAR data. The scale–space theory was first introduced as a generalization of existing notions of Gaussian pyramids (Burt and Adelson, 1983). Moreover, it has been shown that both the scale–space representation and the Gaussian pyramid are embedded in the so-called multiscale Hermite transform (MHT) (Silván-Cárdenas and Escalante-Ramírez, 2006). The MHT provides a well-founded way of relating spatial structures across different scales and provides a rich representation of a signal in terms of local derivatives. It also shares several properties with the wavelet theory, a mathematical tool that has deserved the consideration of scientists in many fields, even for LiDAR segmentation (Thuy and Tokunaga, 2001, 2004).

The content of this paper is organized in five sections. In Section 2, we describe the data used and its preparation for testing the proposed method. In Section 3, we first introduce the MHT theory and its relation to the well-known scale–space theory for signal processing. Then, we present a surface model used to infer the relation of terrain elevations with measured surface elevations. We also define a multiscale erosion operator in the MHT domain. Finally, we describe a semiautomatic MHT-based filtering method. The method was tested over nineteen datasets, including urban and forest areas from data sets acquired. The results and conclusions are presented in Sections 4 and 5, respectively.

2. Data used

2.1. Data source

The International Society for Photogrammetry and Remote Sensing (ISPRS) Commission III/WG3 has tested a number of algorithms developed in the past (Sithole and Vosselman, 2003; Vosselman, 2002). The results of the test, as well as the data used, have been made available through the society's web site (www.commission3.isprs.org/wg3/). For this research, nineteen datasets over urban and forest areas, as well as their respective reference (ground-truth) data, were acquired from the ISPRS web site. The data was captured by an Optech ALTM scanner, and both first and last returns are available. These datasets are located along eight study sites over the Vaihingen test field and Stuttgart city center. The study cites have varied terrain characteristics and diverse feature content (e.g., open fields, vegetation, buildings, road, railroads, rivers, bridges, power lines,

Table 1
Characteristics of the study sites used in the ISPRS test

Site	Cell size	Ref. data	Characteristics
CSite1	1 m	samp11	Steep slopes, mixture of vegetation and buildings on hillside, buildings on hillside, data gaps
	1 m	samp12	
	2 m	*red1	
	4 m	*red2	
CSite2	1 m	samp21	Large buildings, irregularly shaped buildings, road with bridge and small tunnel, data gaps
	1 m	samp22	
	1 m	samp23	
	1 m	samp24	
CSite3	1 m	samp31	Densely packed buildings with vegetation between them, building with eccentric roof left (bottom corner), open space with mixture of low and high features, gaps
CSite4	1 m	samp41	Railway station with trains (low density of terrain points), data gaps
	1 m	samp42	
FSite5	2 m	samp51	Steep slopes with vegetation, quarry, vegetation on river bank, gaps
	2 m	samp52	
	2 m	samp53	
	2 m	samp54	
FSite6	2 m	samp61	Large buildings, road with embankment, gaps
FSite7	2 m	samp71	Bridge, underpass, road with embankments, gaps

The study site FSite8 is not included here. Names starting with 'C' correspond to urban regions, while names starting with 'F' correspond to rural regions.

water surface, among others). The sites represent four regions with urban characteristics and four regions with rural characteristics as described by Sithole and Vosselman (2003). These are listed in Table 1. The study site FSite8 was not used in this research because no ground-truth data was available.

2.2. Data preparation

The original point cloud, as well as the reference terrain points, were interpolated to a grid. Prior to interpolation, outliers were removed following two steps. In the first step, the elevation histogram was examined and elevation thresholds were set up to eliminate the lowest and highest tails from the distribution. Remaining outiers were searched using the minimum height difference of each point with respect to all its neighbors. A Delaunay triangulation was used to define the neighbors of each point. Points that were too high or too low, with respect to their neighbors, were removed from the dataset. The remaining points were then gridded within a 1- 2- or 4-meter resolution grid using the nearest-neighbor rule. The spatial resolution was set up based upon the point spacing of each dataset (Table 1). In addition, a gap mask was built. Each cell having its nearest point at a distance greater than twice the cell size was defined as a gap cell. Although gaps were filled with the nearest-neighbor interpolation, they were not taken into account for the accuracy assessment.

3. Methodology

3.1. The multiscale Hermite transform

The implementation of the MHT is based on the single-scale Hermite transform, whereas the single-scale transform decomposes a signal into a set of decimated (i.e., filtered and sub-sampled) versions of the original signal. For the purpose of this research, the signal is a rectangular grid of elevation values measured by the LiDAR sensor and is referenced in a two-dimensional space, $z(x, y)$. The signal is analyzed by convolution with a bank of Gaussian derivative filters.[1] Here, we summarize the theory for two-dimensional discrete signals and refer the reader to the original for more details concerning the decomposition of continuous signals (see Silván-Cárdenas and Escalante-Ramírez, 2006, and reference cited therein).

[1] Here, the term *filtering* is used in the context of signal processing: the application of an operator (filter) that removes frequency components from the signal.

Let $z(x, y)$ denote a two-dimensional discrete signal defined on grid $\mathcal{G}$. The discrete Hermite transform decomposes the signal as

$$z(x,y) = \sum_{n,m=0}^{N} \sum_{p,q} z_{n,m}(p,q) b_n^*(2p-x) b_m^*(2q-y) \qquad (2)$$

with the coefficient of the expansion given by

$$z_{n,m}(p,q) = \sum_{(x,y)\in\mathcal{G}} z(x,y) b_n(x-2p) b_m(y-2q) \qquad (3)$$

for $n,m=0,\ldots,N$, where

$$b_n(x) = 2^{-N}\sqrt{C_N^n} \sum_{j=0}^{n} (-1)^{n-j} C_n^j C_{N-n}^{x+N/2+j-n} \qquad (4)$$

and $b_n^*(x)=2b_n(-x)=(-1)^n 2b_n(x)$, for $x=-N/2, \ldots, N/2$ and $n=0, \ldots, N$, are the analysis and synthesis functions, respectively. These functions are normalized finite differences of binomial kernels (where, $C_n^m = n!/(n-m)!\,m!$, for $m=0, \ldots, n$, denotes the binomial coefficients). Table 2 shows the values obtained for the finite differences (i.e, the sum term of Eq. (4)) and its norm (i.e., the inverse of the remaining factor), for $N=6$ and $N=8$, respectively. These sequences are hereafter referred to as binomial filters.

The binomial filters are discrete approximations of the Gaussian-derivative filters. More specifically, $b_n(x)$ tends to $g_n(x/\sqrt{N/2})/\sqrt{N/2}$ as N goes to infinite, where

$$g_n(x) = \frac{1}{\sqrt{\pi 2^n n!}} \frac{d^n}{dx^n} e^{-x^2}. \qquad (5)$$

Thus, the coefficient $z_{n,m}$ approximates the *normalized* partial derivative of order n with respect to x and order m with respect to y, of the Gaussian-smoothed version of z at each sampling site, where the standard deviation $\sqrt{N/4}$ of the Gaussian function determines the degree of smoothness. Thus, $z_{0,0}$ denotes the samples of the smoothed signal, $z_{1,0}$, the samples of its first-order partial derivative with respect to x, $z_{0,1}$, the samples of its first-order partial derivative with respect to y, and so forth.

By definition, the first resolution level in the multiscale setting corresponds to the single-scale transform with $N=8$, say $z_{n,m}^{(1)}$. Subsequent resolution levels are iteratively computed from the low-pass coefficient, with $N=6$, as

$$z_{n,m}^{(k+1)}(p,q) = \sum_{x,y} z_{0,0}^{(k)}(x,y) c_n(x-2p) c_m(y-2q) \qquad (6)$$

Table 2
Binomial filters for (a) $N=6$ and (b) $N=8$

n\x	0	1	2	3	4	5	6	Norm
0	1	6	15	20	15	6	1	$2^6/\sqrt{1}$
1	1	4	5	0	−5	−4	−1	$2^6/\sqrt{6}$
2	1	2	−1	−4	−1	2	1	$2^6/\sqrt{15}$
3	1	0	−3	0	3	0	−1	$2^6/\sqrt{20}$
4	1	−2	−1	4	−1	−2	1	$2^6/\sqrt{15}$
5	1	−4	5	0	−5	4	−1	$2^6/\sqrt{6}$
6	1	−6	15	−20	15	−6	1	$2^6/\sqrt{1}$

n\x	0	1	2	3	4	5	6	7	8	Norm
0	1	8	28	56	70	56	28	8	1	$2^8/\sqrt{1}$
1	1	6	14	14	0	−14	−14	−6	−1	$2^8/\sqrt{8}$
2	1	4	4	−4	−10	−4	4	4	1	$2^8/\sqrt{28}$
3	1	2	−2	−6	0	6	2	−2	−1	$2^8/\sqrt{56}$
4	1	0	−4	0	6	0	−4	0	1	$2^8/\sqrt{70}$
5	1	−2	−2	6	0	−6	2	2	−1	$2^8/\sqrt{56}$
6	1	−4	4	4	−10	4	4	−4	1	$2^8/\sqrt{28}$
7	1	−6	14	−14	0	14	−14	6	−1	$2^8/\sqrt{8}$
8	1	−8	28	−56	70	−56	28	−8	1	$2^8/\sqrt{1}$

Each row gives the filter order (first column), the de-normalized filter coefficients (center columns), and its norm (right column).

with $c_n(x) = (\sqrt{3}/2)^n b_n(x)$, for $n,m=0, \ldots, 6$, $k=1, \ldots, K$. The computed coefficients approximate the Gaussian derivatives of the smoothed signal with scale parameter $N_k=2\cdot 4^k$ at a sub-sampling rate of 2^k. Thus, the ratio between the smoothing degree and the sub-sampling rate remains constant across the resolution levels.

In the inversion process, the low-pass coefficients are re-synthesized from coarser resolution levels as

$$z_{0,0}^{(k)}(p,q) = \sum_{n,m=0}^{6} \sum_{i,j} z_{n,m}^{(k+1)}(i,j) c_n^*(2i-p) c_n^*(2j-q) \quad (7)$$

for $k=K-1, K-2, \ldots, 1$ with $c_n^*(x) = (\sqrt{3}/2)^{-n} b_n^*(x)$.

Fig. 1(a) illustrates the scheme of computation for the MHT at three resolution (or scale) levels, where only the coefficients up to the second order for each level are shown. The first coefficient corresponds to the low-pass (smoothed) version of the signal, whereas signal details are encoded in the higher-order (derivative) coefficients. The re-synthesis of the signal from the coefficients are illustrated in Fig. 1(b). The set of residuals obtained at each level (middle) comprises the Laplacian pyramid (Burt and Adelson, 1983), whereas the set of partially synthesized signal at each level comprises the Gaussian pyramid (right). If zeroes replace residuals up to a given level of the pyramid, the re-synthesized signal is a Gaussian-smoothed version of the original signal. Moreover, any Gaussian-smoothed version of the original signal can be reconstructed by the proper weighting of the coefficients (Silván-Cárdenas and Escalante-Ramírez, 2006). The set of all Gaussian-smoothed signals is termed the scale–space representation of the input signal (Witkin, 1984).

3.2. Surface model

Linear operators such as those employed in the MHT are good tools to detect elevation changes that occur at different scales. While changes in elevation can be due to natural terrain relief or variation in feature height, change in elevation due to transition between ground and non-ground surface elevations are more important from the filtering problem point of view. The transition between ground and non-ground is certainly complex for most real datasets; however, we argue that most of them may be expressed as linear combinations of shifted, scaled, and rotated versions of a basic template. In the context of the MHT, the ideal template function is the error function (erf), that is introduced in probability theory.

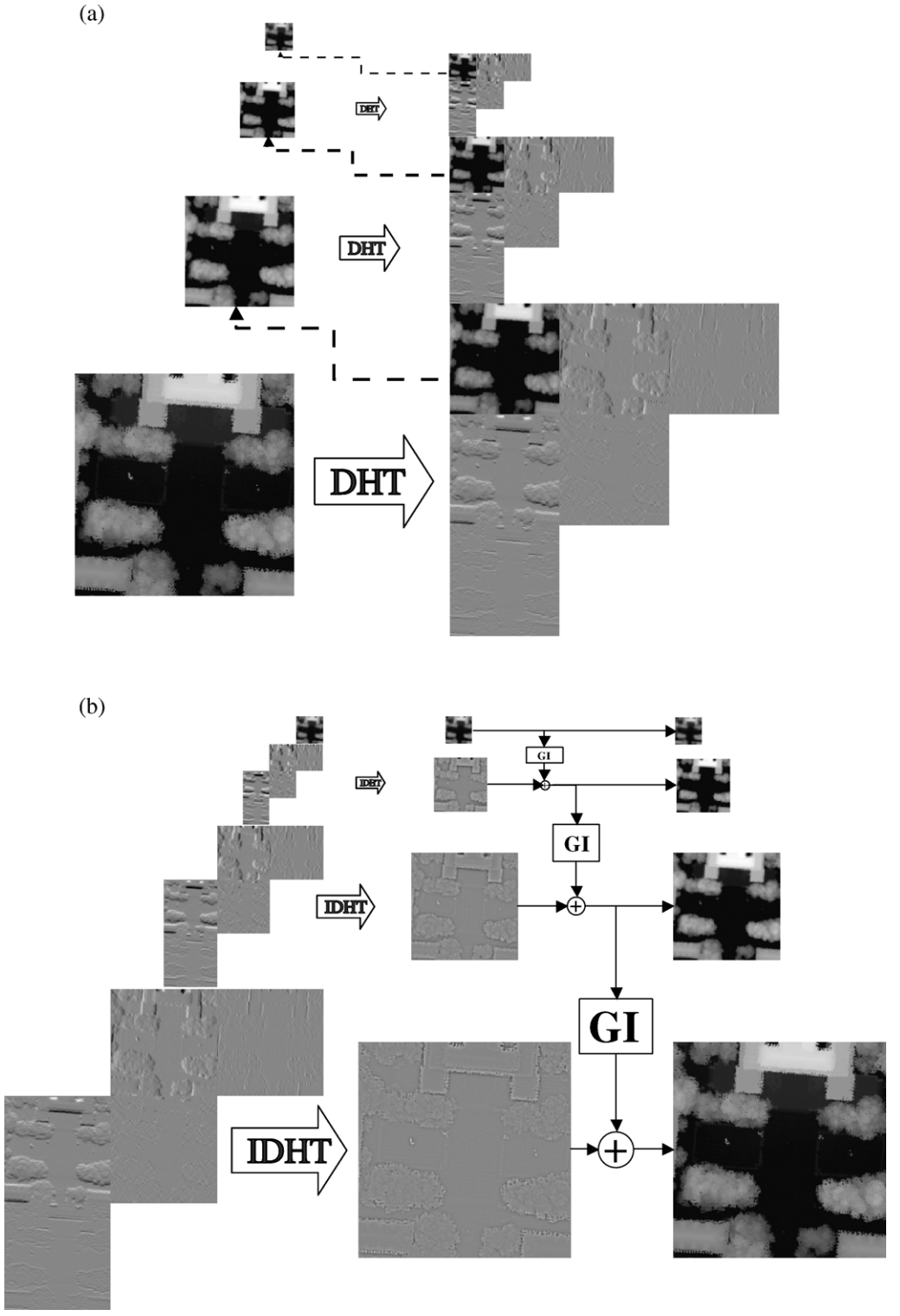

Fig. 1. (a) Computational diagram of the MHT showing three resolution levels, where DHT stands for discrete Hermite transform. The array of sub-images at each level represents the coefficients up to the second order, where $z_{n,m}$ is located in row $n+1$ and column $m+1$ within the array. (b) Computational diagram of the inverse MHT, where IDHT stands for inverse discrete Hermite transform and GI stands for Gaussian interpolation.

Consider the following model for one-dimensional surface profiles:

$$z(x) = a_0 + \sum_{i=1}^{M} a_i \mathrm{erf}\left(\frac{x - x_i}{\sqrt{4 s_i}}\right) \tag{8}$$

where a_0, a_i, x_i, and s_i, for $i=1, \ldots, M$, are the model parameters. These parameters give the model flexibility to represent both terrain and features of varied shapes and sizes, provided that enough number of terms are used.

For example, Fig. 2(b) shows a simulated profile using 22 erf terms. The contribution of the i-th erf term to the overall profile has the shape of a smooth step (Fig. 2(a)) with maximum elevation difference of $a_i/2$, center of transition located at x_i, and smoothness determined by the scale parameter s_i. This smoothness characteristic of surfaces may be hard to quantify from real datasets, especially because we are more used to measuring distances and slopes. The estimation of the scale parameter has been recognized as an important issue in other fields, such as computer vision and image processing, where many approaches have been proposed. From our simple model, the maximum contribution of the i-th term to the slope, say m_i, is reached at x_i and the scale parameter is related to this as

$$s_i = \frac{a_i^2}{\pi m_i^2} \tag{9}$$

The advantage of using Eq. (8) as a model for both terrain and features is that its shape remains the same in the scale–space representation up to a shift of the scale parameters, i.e., s_i becomes $s+s_i$ for all i, where s is the scale parameter of the scale–space representation. In fact, its MHT representation in the continuous domain can be written explicitly as

$$z_n^{(k)}(\xi) = \begin{cases} a_0 + \sum_i a_i \mathrm{erf}\left(\dfrac{\xi - 2^{-k} x_i}{\sqrt{1 + 4^{1-k} s_i}}\right), & n = 0 \\ \sum_i \dfrac{a_i}{\sqrt{(1 + 4^{1-k} s_i)^n n/2}} g_{n-1}\left(\dfrac{\xi - 2^{-k} x_i}{\sqrt{1 + 4^{1-k} s_i}}\right), & n > 0 \end{cases}$$

This representation is particularly useful for estimating the model parameters. First, the maximum contribution of each term to the slope of the profile is given by the coefficient of first-order at transition locations, i.e.,

$$m_i = z_1^{(k)} \left(\frac{1 + 4^{1-k} s_i}{2 s_i}\right)^{1/2} \tag{10}$$

where $z_1^{(k)}$ defines the surface gradient at x_i. Second, there is an approximated linear relation between the ratio of the first- and third-order coefficients and the underlying scale at each transition location, i.e.,

$$s_j = 4^k \left(a + b \frac{z_1^{(k)}}{z_3^{(k)}}\right) \tag{11}$$

with $a=-1/4$ and $b = -1/\sqrt{96}$. In practice, because of the discrete approximation of the MHT, these parameters are no longer constants but functions of the resolution level k. Table 3 shows the least square (LS) estimate of these values for up to six resolution levels when the discrete MHT of an erf term is used. Third, a transition point x_j can be detected through the well-known zero-crossing test, i.e.,

$$\begin{cases} z_2^{(k)}(2^{-k} x_j) = 0 \\ z_3^{(k)}(2^{-k} x_j) < 0 \end{cases} \tag{12}$$

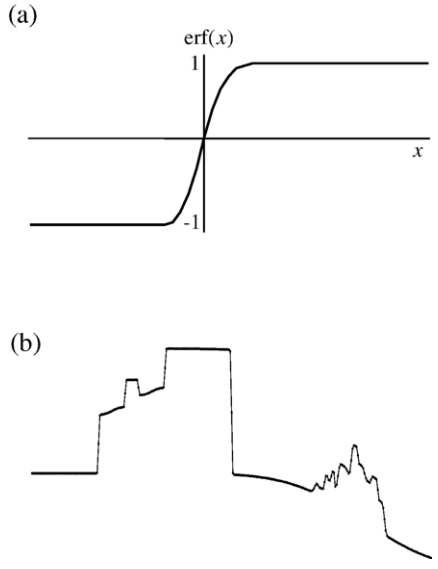

Fig. 2. (a) Error function plot and (b) simulated surface profile using the model of Eq. (8).

3.3. Multiscale erosion

Notice that the model of Eq. (8) is highly non-linear on the parameters and there may be no unique set of parameters for a given surface. Moreover, its generalization of two dimensions is not straightforward. However, it is still useful for inferring how the coefficients of the terrain component might be related to the surface.

Eq. (1) can be translated into the MHT representation as $z_n^{(k)} = t_n^{(k)} + f_n^{(k)}$, where the contribution of the error component has been neglected as it is not of major

Table 3
LS-estimate of constants in Eq. (11) for the discrete MHT of an erf term

k	1	2	3	4	5	6	∞
a	−0.1708	−0.1745	−0.1835	−0.1966	−0.2076	−0.2131	−0.2500
b	−0.0827	−0.0762	−0.0771	−0.0800	−0.0836	−0.0857	−0.1021

The exact values obtained from the continuous case are given in the right-most column.

concern here. We hypothesize that the coefficients of the ground component can be estimated from the observed elevations. In particular, we consider a model of the form [2]

$$\hat{t}_n^{(k)} = \begin{cases} z_n^{(k)} & \text{if terrain} \\ \sum_{i=n}^{N} c_i z_i^{(k)} & \text{if non−terrain and } n = 0 \\ 0 & \text{if non−terrain and } n > 0 \end{cases} \quad (13)$$

for terrain extraction, where c_i are constants to be determined. For reasons that will become clear later, the second case in Eq. (13) is termed the *erosion* operator. The third case assumes a maximally flat terrain surface underneath features and their vicinity. This model relies on an *a priori* ground detection around transition points.

A simple detection method can be carried out by thresholding the terrain gradient, which is usually smaller than the feature gradient. Combining Eqs. (9) and (10) one can find a bound for the multiscale terrain gradient $G^{(k)}=|z_1^{(k)}|$ in terms of its maximum slope and elevation difference, i.e., $G^{(k)} \le T^{(k)}$, where

$$T^{(k)} = \frac{2^k m_{\max}}{\sqrt{2 + 2\pi(2^k m_{\max}/\Delta_{\max})^2}} \quad (14)$$

is the multiscale threshold given in terms of the maximum terrain slope $m_{\max}$ and the maximum terrain elevation difference $\Delta_{\max}$.

For two-dimensional signals, the multiscale gradient is written as

$$G^{(k)} = \sqrt{z_{1,0}^{(k)2} + z_{0,1}^{(k)2}}. \quad (15)$$

Unfortunately, the generalization of the erosion operator to two dimensions is not straightforward. A simple generalization that considers the coefficients up to the first-order is considered here:

$$\hat{t}_{0,0}^{(k)} = c_0 z_{0,0}^{(k)} + c_1 G^{(k)} \quad (16)$$

In this simple case, the parameters c_0 and c_1 can be determined by considering the one-dimensional case of an erf term. This results in $c_0=1$ and $c_1 = -1/\sqrt{2} = -0.7071$ as the best approximation.

3.4. The MHT-based filtering algorithm

The erosion operator of Eq. (16) and the multiscale threshold of Eq. (14) are applied to adaptively remove above-ground features as the MHT is computed. When the gradient is greater than the corresponding threshold, the low-pass coefficient is eroded by subtracting the gradient, and all the other coefficients are set to zero. Otherwise stated,

$$\hat{t}_{n,m}^{(k)} = \begin{cases} z_{n,m}^{(k)} & \text{if } G^{(k)} \le T^{(k)} \\ z_{0,0}^{(k)} - G^{(k)}/\sqrt{2} & \text{if } G^{(k)} > T^{(k)} \text{ and } n,m = 0 \\ 0 & \text{if } G^{(k)} > T^{(k)} \text{ and } n,m > 0 \end{cases} \quad (17)$$

As a rule of thumb, the maximum resolution level K of the MHT must be chosen in terms of the largest feature size to be removed (say 2^K).

The inverse transform is then computed from the processed coefficients $\hat{t}_{n,m}$ to obtain an approximation of the terrain surface, $\hat{t}$. In a final step, the original grid is compared against the approximated ground surface to filter the ground points from the LiDAR dataset. More specifically, a label image L is generated with the following rule

$$L = \begin{cases} \text{ground}, & \text{if } z < \hat{t} + \epsilon \\ \text{non−ground}, & \text{if } z \ge \hat{t} + \epsilon \end{cases} \quad (18)$$

where ϵ is tolerance value. A good choice for this tolerance would be the quantile of the approximation errors that ensure omission (Type I) and commission (Type II) errors are equally probable. In practice, however, the true terrain surface is unknown and thus the errors cannot be derived. Here, we determined experimentally that $\epsilon = 0.1$ m was a good choice for most of the tested cases.

4. Results

The validation of the erosion model (Eq. (16)) was carried out by observing the average LS-estimate of c_0 and c_1 over all the reference sets. In this case, the

[2] Here, $\hat{t}_n$ denotes an estimate of t_n.

Table 4
Examples of LS-estimate of c_0 and c_1, and measurement (M) vs. estimate (T) of the maximum gradient

Dataset	k	c_0	c_1	R	C	P	N	D	A	T	M
sample11	1	1.0	−1.0	0.38	0.99	0.0	772	100	56	2.09	2.55
	2	1.0	−1.1	0.55	0.99	0.0	193			4.17	3.51
	3	1.0	−0.5	0.79	0.99	0.0	28			8.21	5.13
sample24	1	1.0	−0.6	0.63	0.26	0.0	85	17	30	0.81	2.14
	2	1.0	−2.1	0.77	0.19	0.4	22			1.59	2.85
	3	1.0	1.1	1.09	0.15	0.9	4			2.94	3.34
sample53	1	1.0	−0.1	0.26	0.99	0.0	385	33	88	12.52	6.30
	2	1.0	−0.1	0.21	1.00	0.0	99			12.99	9.42
	3	1.0	0.0	0.17	1.00	0.0	25			13.12	11.77

multiscale gradient threshold was set to the maximum terrain gradient M, which was measured directly from the MHT of the bare ground surface. At each level of the multi-resolution decomposition, the root mean square error (R) and the correlation coefficient (C) between the eroded ($\hat{t}_{0,0}^{(k)}$) and true ($t_{0,0}^{(k)}$) transform coefficients were computed. Only the sites that had a gradient value above the maximum terrain gradient were assumed to be non-ground and thus used for the LS estimation. In addition, the multiscale threshold of Eq. (14) was computed using measurements of the maximum aspect (A) and the maximum height difference (D), which were made from interactive 2-D and 3-D visualizations of the (unfiltered) reference datasets.

Results of the LS-estimate for a few sample sets are shown in Table 4, where the probability of no-correlation (P) and the number of points (N) used in the fitting procedure are provided as well. The average LS-estimate computed over all the datasets agreed with our initial approximation of $c_0=1$ (avg.=1.00) and $c_1=-0.7071$ (avg.=−0.70). The few cases in which the correlation coefficient was not statistically significant (say, when $P>0.1$) are attributed to the relatively small sample size used in the fitting procedure (say, $N<30$), which of course decreases with the resolution. We also noticed a considerably larger variability of c_1 (st. dev.=0.70) in comparison to that of c_0 (st. dev.= 0.01). The corresponding values of M, as well as the values of the theoretical gradient threshold T, are also provided in Table 4 for comparison purposes. Even though there was a significant difference between the observations and the model-based thresholds in several cases, certain agreement was still observed. We believe that discrepancies are mainly attributed to the fact that parameter measurements were carried out through interactive visual inspection, rather than by an optimization method.

The Cohen's kappa coefficient of agreement, as derived from the error matrix (Congalton, 1991), was used to rank the proposed method. Fig. 3(a) shows a bar plot of the average kappa value over all the datasets for each method, including the MHT-based method introduced in the previous section. The kappa values for the other methods were computed from the error matrices reported by Sithole and Vosselman (2003). The kappa statistic gives a measure of agreement between the classification (terrain vs. non-terrain) produced by the method and the true classification (obtained with the reference data). In this sense, Type I (classify ground

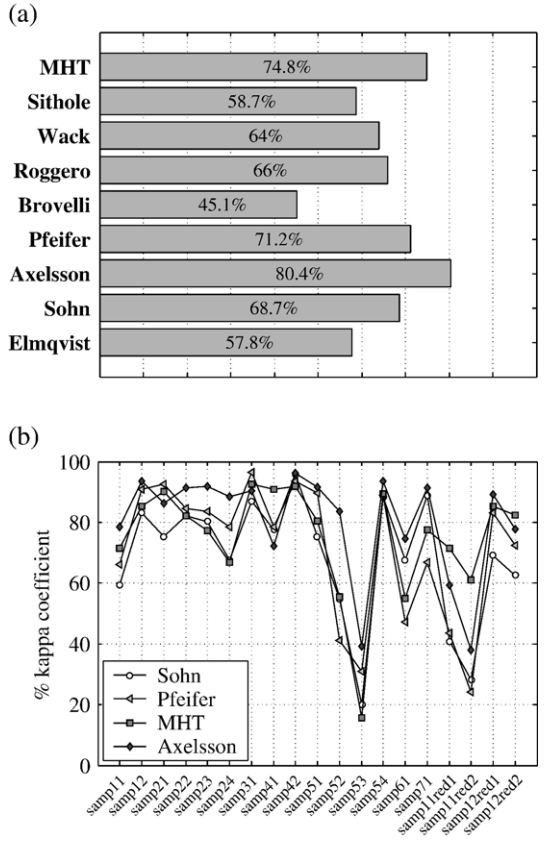

Fig. 3. (a) Average kappa value for several methods and (b) kappa value by sample set for the top-four methods.

points as non-ground) and Type II errors (classify non-ground points as ground) are considered equally important. This, of course, may not be the case in some applications, yet it provides an objective means to compare different methods.

Even though the parameter selection for the MHT-based method was not carried out through an optimization process, but rather guided by visual inspection of the original unfiltered data, the results of the average kappa value showed that, in general, the MHT method performs very well (Fig. 3(a)), although below Axelsson's method (Axelsson, 1999) yet similar to Pfeifer's method (Pfeifer et al., 2001). Furthermore, from the plots shown in Fig. 3(b) we can see that the MHT-based method is more robust than other methods when data is at low resolution (e.g., samp11red1, samp11red2. See Fig. 4(b) and (c)). This method could also be potentially useful in cases of large buildings (e.g., samp22, samp41.

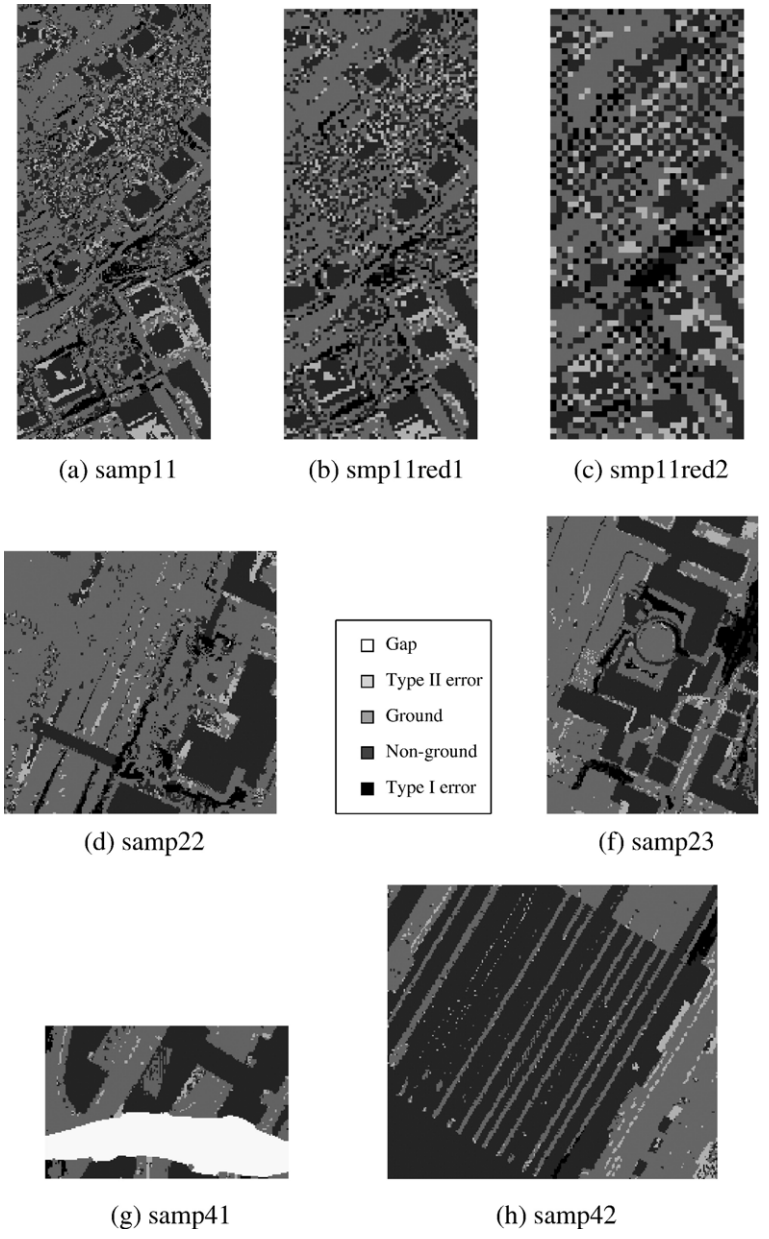

Fig. 4. Error distribution from several data sets (see the text).

See Fig. 4(g) and (d)), and when there is a low density of terrain points (e.g., samp42. See Fig. 4(h)).

However, the major limitations of the method seems to be the need for a proper selection of the maximum slope and maximum height difference required for the thresholds. The fact that it performed particularly poorly (below the average) with the dataset samp53 may be attributable not only to the wrong selection of these parameters (notice the large discrepancies between the measured and estimated maximum terrain gradient given in Table 4), but also to the poor separability achieved with the slope (see the histograms of Fig. 6(d) and the discussion presented at the end of the Conclusions section).

Another factor affecting the performance of the proposed method is the *boundary problem* of the convolution filters (i.e, the problem of computing the convolution at cells near the spatial boundaries). The boundary problem was addressed here by imposing the *symmetry condition*. According to this condition, the cell values along the boundaries are reflected symmetrically beyond the spatial extent. This solution may lead to misclassification of cells around the boundary when strong deviation from the symmetry assumption occurs. The errors can even propagate toward the central portion of the grid in cases when too many resolution levels (or scales) are used, which are in turn required for removing very large objects. Fig. 5 shows an example of this problem for CSite2. Here, the boundary problem results in the misclassification of the ground along the right-hand side edge. Notice that the method can detect the large buildings falling in samp22 and samp23; however, the kappa statistic obtained for these sample sets are relatively low due to the boundary effect. The error distribution for these sample sets are shown in Fig. 4(d) and (f), respectively.

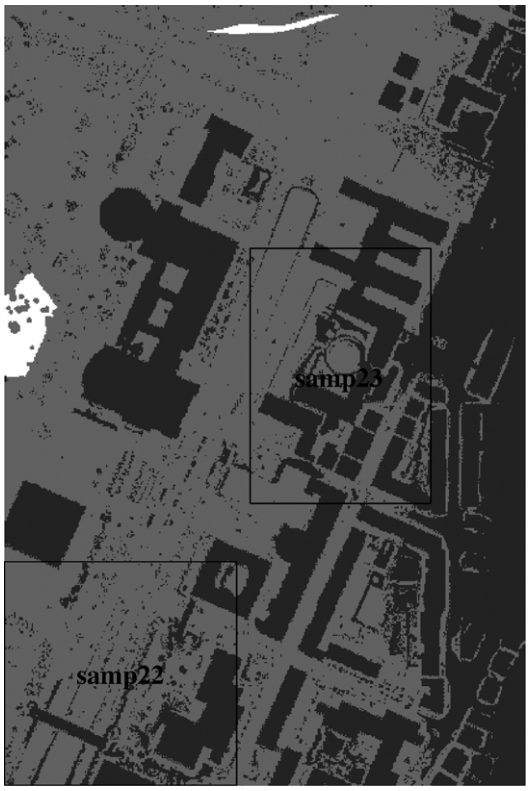

Fig. 5. Filtering result showing the boundary effect in CSite2.

5. Conclusions

Many approaches to LiDAR filtering are based on models of the ground in terms of local properties such as height difference and slope (e.g., Shan and Sampath, 2005; Vosselman, 2000). The properties are measured through spatially local operators, either morphologic or structural. As the term *local* is indeed scale-dependent, it is then necessary to integrate measurements at different scales. This already had been noticed in the context of mathematical morphology-based methods (e.g., Kilian et al., 1996; Zhang et al., 2003, Zhang and Whitman, 2005). However, an analysis–synthesis framework had not been implemented for filtering ground points from LiDAR data. In this paper, we have proposed an approach based on the scale–space representation of the gridded elevation values. Such a representation is enriched by the use of the so-called multiscale Hermite transform (MHT) which employs derivatives of the smoothed version of the signal. It is interesting to note that most of the approaches reviewed incorporate height differences and local slope measures. Others use the second derivatives (Axelsson, 1999) and the Laplace operator (Maas, 1999) to find textural variations that can be used in a more detailed classification of the LiDAR data (e.g. ground, vegetation and buildings). We believe that the MHT representation can be potentially useful in such cases, as it is comprised of multiscale derivatives. The advantage of using the MHT representation became evident when a model-based characterization of the LiDAR data was adopted. Specifically, the MHT allows estimating the parameters using the first few coefficients.

The method was tested with nineteen datasets covering a wide spectrum of terrain shapes and diverse feature types, both in urban and rural regions. The datasets had already been used in a comparative study conducted by the ISPRS Commission III/WG3 (Vosselman, 2002). Thus, this research has built upon previous

results. The results obtained herein indicate that the MHT representation was a viable tool to address the filtering problem. Even though several assumptions and simplifications were made in the model, the results showed that this technique could compete with the most popular methods.

Although the results reported herein look promising, it has to be recognized that there are still a number of issues that should be addressed in future research. For example, a more rigorous assessment of the vertical errors committed by the approximation used is a mandatory issue for real applications. Another interesting issue is how higher-order coefficients may be incorporated in the erosion model and/or in the prediction of the coefficients that are set to zero in the decomposition process.

Currently, we are investigating how the scale parameter could be used to obtain better separability between transitions of terrain–feature and terrain–terrain types. We believe that the scale parameter may be a determinant in the separability of above-ground features from terrain. For example, in the case of Fig. 2(b), only one term had a large scale (in the order of 100 times larger than the rest of terms) and it might be the only one interpreted as terrain by simply inspecting the profile. Indeed, the average feature scale is consistently smaller than the average terrain scale, whereas the average feature slope is not necessarily smaller (or greater) than the average terrain slope. This belief is supported by the histograms of the estimated scales and slopes (Eqs. (10) and (11)) of both terrain and feature components for all the sample sets. Fig. 6 shows examples of such histograms for the datasets samp11 and samp53. In all cases, neither slope nor scale parameters are enough to separate terrain elevations from feature height; nevertheless, the average scale for feature height was consistently smaller than the average scale for terrain variations. Furthermore, the scale histograms seemed

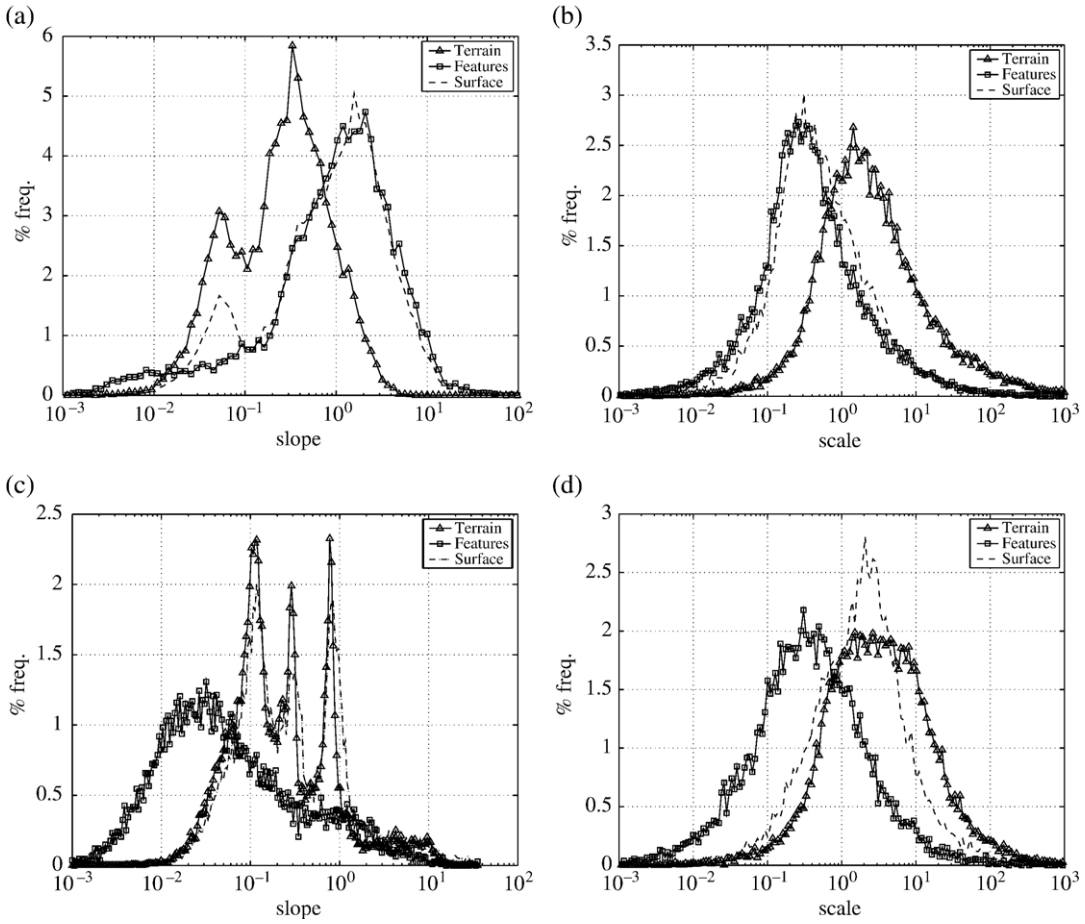

Fig. 6. (a),(c) slope and (b),(d) scale histograms for terrain, feature and surface obtained from (a),(b) samp11 and (c),(d) samp53.

uni-modal in all the cases, whereas the slope histograms appeared multi-modal in some cases (e.g., Fig. 6(d)).

Acknowledgements

The first author thanks CentroGeo for the support received while working in this research. The study was also supported by a grant to L. Wang from United States Department of Agriculture via Sul Ross State University (CSREES Award No. 2004-38899-02181), and a grant to L. Wang from the United States Geological Survey via Texasview Remote Sensing consortium in 2006. The authors also thank the anonymous reviewers for their suggestions that helped to improve the content of this paper.

References

Axelsson, P., 1999. Processing of laser scanner data — algorithms and applications. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2–3), 138–147.

Burt, P.J., Adelson, E.H., 1983. The laplacian pyramid as a compact image code. IEEE Transactions on Communications 31 (4), 532–540.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37, 35–46.

Kilian, J., Haala, N., Englich, M., 1996. Capture and evaluation of airborne laser scanner data. International Archives of Photogrammetry and Remote Sensing 31 (part B3), 383–388.

Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53 (4), 193–203.

Lloyd, C.D., Atkinson, P.M., 2002. Deriving DSMs from LiDAR data with kriging. International Journal of Remote Sensing 23 (12), 2519–2524.

Lohmann, P., Koch, A., Schaeffer, M., 2000. Approaches to the filtering of laser scanner data. International Archives of Photogrammetry and Remote Sensing 33 (Part B3), 534–541.

Maas, H.G., 1999. The potential of height texture measures for the segmentation of airborne laser scanner data. Proc. Fourth International Airborne Remote Sensing Conference, Ottawa, Canada, pp. 154–161.

Pfeifer, N., Stadler, P., Briese, C., 2001. Derivation of digital terrain models in the SCOP++ environment. Proc. OEEPE Workshop on Airborne Laser Scanning and Interferometric SAR for Detailed Digital Elevation Models, Stockholm, Sweden.

Roggero, M., 2001. Airborne laser scanner-clustering in raw data. International Archives of Photogrammetry and Remote Sensing 34 (Part B3), 227–232.

Shan, J., Sampath, A., 2005. Urban DEM generation from raw lidar data: a labeling algorithm and its performance. Photogrammetric Engineering and Remote Sensing 71 (2), 217–226.

Silván-Cárdenas, J.L., Escalante-Ramírez, B., 2006. The multiscale hermite transform for local orientation analysis. IEEE Transactions on Image Processing 15 (5), 1236–1253.

Sithole, G., 2001. Filtering of laser altimeter data using a slope adaptive filter. International Archives of Photogrammetry and Remote Sensing 34 (Part 3), 203–210.

Sithole, G., Vosselman, G., 2003. ISPRS comparison of filters. Delft University of Technology, Tech. Rep. Commission III/WG3, available online: http://enterprise.lr.tudelft.nl/frs/isprs/filtertest/Report05082003.pdf.

Thuy, V., Tokunaga, M., 2001. Wavelet and scale–space theory in segmentation of airborne laser scanner data. Proc. 22th Asian Conference on Remote Sensing, Bangkok, pp. 97–103.

Thuy, T., Tokunaga, M., 2004. Filtering airborne laser scanner data: a wavelet-based clustering method. Photogrammetric Engineering and Remote Sensing 70 (11), 1267–1274.

Vosselman, G., 2000. Slope based filtering of laser altimetry data. International Archives of Photogrammetry and Remote Sensing 33 (Part B3), 935–942.

Vosselman, G., 2002. ISPRS test on extracting DEMs from point clouds: a comparison of existing automatic filters. http://www.geo.tudelft.nl/frs/isprs/filtertest, www document.

Witkin, A., 1984. Scale–space filtering: a new approach to multiscale description. Image Understanding 3, 79–95.

Zhang, K., Whitman, D., 2005. Comparison of three algorithms for filtering airborne lidar data. Photogrametric Engineering and Remote Sensing 71 (3), 313–324.

Zhang, K., Chen, S., Whitman, D., Shyu, M., Yan, J., Zhang, C., 2003. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. IEEE Transactions on Geoscience and Remote Sensing 41 (4), 872–882.