A FFT BASED METHOD OF FILTERING AIRBORNE LASER SCANNER DATA

U. Marmol, J. Jachimski

University of Science and Technology in Krakow, Poland Department of Photogrammetry and Remote Sensing Informatics –(entice, jachim@uci.agh.edu.pl)

Commission III, WG III/3

KEY WORDS: LIDAR, Algorithms, DEM/DTM, Laser scanning, Extraction

ABSTRACT:

Airborne laser altimetry is a relatively new method for the acquisition of information of terrain surface. A laser scanning system generates a 3-dimensional cloud of points with irregular spacing. The data consists of the mixture of terrain surface and non-surface points (buildings, vegetation). The separation of ground points from the other points located on top of buildings, vegetation or other objects above ground is one of the major problems. Algorithms and software used for the surface reconstruction have limitations that should be studied and overcome. Removing non-ground points from LIDAR data sets is still a challenging task. The paper presents a new method concerning data filtering for determining the ground surface, i.e. defining a digital terrain model DTM, as a subset of a measured data. The filtering algorithm is based on Fast Fourier Transformation (FFT) and is closely related to the digital filters used for signal processing theory. In analogy with electrical signal, terrain relief can be defined as superposition of finite number sine components with different amplitudes and frequencies. The article presents also the results of empirical studies on using the mentioned filter. Accuracy analysis has been made on the basis of numerical comparison of the filter output with a reference data set for the same site (obtained photogrammetrically).

1. INTRODUCTION

Airborne laser altimetry is a relatively new method for the acquisition of information of terrain surface.
For many applications with a need for high density measured data, high accuracy elevation models, laser altimetry offers unique technical capabilities, lower field-operation costs and reduced post-processing time and effort compared to traditional survey methods (Flood, 2001).
Numerous filter algorithm have been developed for example: iterative robust interpolation (Pfeifer et al., 2001), adaptive TIN model (Axelsson, 2000), morphological operator (Kilian et al., 1996; Vosselman, 2000).
The paper presents a new method concerning data filtering for determining the ground surface, i.e. defining a digital terrain model DTM, as a subset of a measured data. The filtering algorithm is based on FFT and is closely related to the digital filters used for signal processing theory.

2. A FFT BASED ALGORITHM OF FILTERING

2.1 Theory

Digital signal processing is a base of various scientific disciplines connected with digital photogrammetry, such as artificial intelligence and digital image processing. Discrete digital signal is represented by a sequence of numbers (samples) that vary with respect to some independent variables, for example the time.
However the term ″signal″ can be understood in a broader sense. If we introduce as an independent variable – the location, then the surface of terrain and terrain features, represented in a discrete form (DEM), may be considered as a discrete signal that records the elevation variation for data points in a x, y location.

This analogy allows to draw a conclusion that some of the fluctuations of the surfaces reflecting laser pulses are very similar to the noise present at a signal during the transmission process. In connection with this, the spectral analysis and digital filtering reveal some interesting results of the DTM analysis.
In this study the two major signal processing tools: spectral analysis based on FFT and digital filtering are discussed.

2.1.1 Fast Fourier Transformation: Spectral analysis describes relationship between space domain and frequency domain. Discrete spatial data can be described in the frequency domain by applying FFT. A function that is known as a spectrum is a result of this transformation:

$$X(\omega) = \sum_{n=0}^{N-1} z(n) \cdot e^{-j \cdot n \cdot \Delta x \cdot m \cdot \omega} = X_{real}(\omega) + j \cdot X_{imag}(\omega) \quad (1)$$

where

$$X_{real}(\omega) = \sum_{n=0}^{N-1} z(n) \cdot \cos(n \cdot \Delta x \cdot m \cdot \omega) \quad (2)$$

$$X_{imag}(\omega) = \sum_{n=0}^{N-1} z(n) \cdot \sin(n \cdot \Delta x \cdot m \cdot \omega) \quad (3)$$

where ω – angular frequency
Δx – sampling interval
$j = \sqrt{-1}$
z(n) – elevation value of measured points

According to this equation a surface can be treated as the superimposition of a finite number of sinusoidal components of the adequate amplitudes, frequencies and phases.
The function that emerged from FFT does not have a simple physical interpretation. Certain conversion of spectrum is necessary in order to determine physical parameters. An amplitude spectrum it is the absolute value of the FFT spectrum:

$$X_{amp}(\omega) = |X(\omega)| = \sqrt{X^2_{real}(\omega) + X^2_{imag}(\omega)} \qquad (4)$$

The phase spectrum is calculated using following equation:

$$X_{\phi}(\omega) = arctg\left(\frac{X_{imag}(\omega)}{X_{real}(\omega)}\right) \qquad (5)$$

The frequency structure of tested surface is represented by the amplitude spectrum and phase spectrum. Using the spectrum one can unambiguously reconstruct a topographical surface. The spectrum is thus an alternative and equivalent way of a surface representation.
The power spectrum is defined as the squared magnitude of the FFT of the data divided by the number of measured points. A periodogram it is an estimation of the spectral density function.
It is important to know the characteristic of such curve in order to be able to distinguish between portion representing true elevation of surface and portion representing noise (so-called cut-off frequency ω_c).

2.1.2 Digital filtering: The digital filters, which do not pass a given frequency components of topographical surface, can be applied when the cut-off frequency is known.
According to definition digital filter is a discrete system that converts the input data in a certain way, by changes in the spectrum.
In a numerical form the filter is described by an *P[ ]* operator, which converts *z(n)* input data in $z_F(n)$ output data, called a filter response (Ionescu, 1996):

$$z_F(n) = P[z(n)] \qquad (6)$$

Digital filter can be presented in the spatial domain as a discrete convolution sum of the measured data and impulse response *h(n)*:

$$z_F(n) = h(n) * z(n) \qquad (7)$$

Digital filter may be also described in frequency domain as a so-called frequency response i.e. Fourier transform of impulse response:

$$H(\omega) = \sum_{n=0}^{N-1} h(n) \cdot e^{-j \cdot n \cdot \Delta x \cdot m \cdot \omega} \qquad (8)$$

The frequency response of a low-pass filter is presented in the Fig. 1.

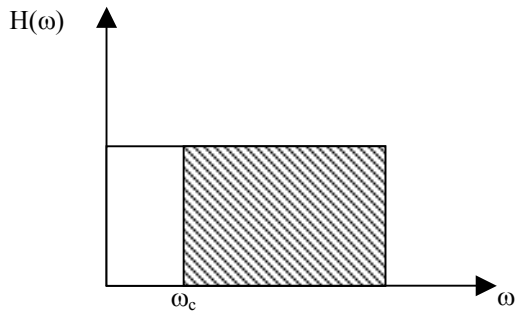

Figure 1. The frequency response of low-pass filter (cut–off frequency ω_c)

2.2 Filtering procedure

The principal assumption is based on the idea, that the low frequencies are responsible for the run of topographical surface, while the high frequencies are connected to the field objects, located on the ground. It is necessary, therefore, to design a low-pass filter, which let lower frequencies pass, and blocks higher frequencies (Marmol, 2002). The designed method it is an iterative process, which consists of the following phases:

Interpolation

The theory of digital transformation of signals is based on the assumption, that a considered flow of the initial data it is a set of data distributed in equal distances. To fulfil that condition it is necessary to interpolate points located in regular matrices out of the cloud of points registered during the laser scanning. As it is well known, the laser data comprise not only ground points, but also the points located on various field objects. An interpolation process can contribute to the significant falsification of the topographical surface run, and can make the filtration process difficult. To reduce the errors the Nearest Neighbour interpolation was applied to the original data.

The trend analysis

The subsequent research phase was aimed on determination of the trend of analysed surface. The removal of trend must necessarily be executed before applying the FFT, to omit an influence of trend during estimation of spectrum.
The trend analysis consist in a least square approximation of a set of points with the use of a polynomial function. The degree of fitting the data by the trend is evaluated by determination coefficient. (Kokesz, 1984).
The significance of empirical data approximation was verified with the use of F-Snedecor test with the significance level α=0.01 (Marmol, 2003).
The influence of the polynomial degree on a quality of fitting of empirical data by the trend was also tested.

Designing the digital filters

In the research a FIR filters (Finite Impulse Response) are used. The simple method to design a FIR filter is the window method. The procedure consists of following stages:

1. Determination of required, ideal frequency response $H(\omega)$ (see Fig. 1) – determination of cut – off frequency ω_c
2. Calculation of infinite impulse response $h(n)$ via inverse transformation of ideal frequency response
3. Truncating impulse response using a window function to receive finite and causal impulse response. The size of window is equal to the filter order.

The impulse responses of a FIR filter are also the filter coefficients.
In the first phase a one-dimensional filter was elaborated, and then it was transformed to the 2-dimensional filter for the 3-dimensional space. The determined filter coefficients $h(n1,n2)$ were used to filter the analysed topographical surface

FFT testing and determination of the cut-off frequency

Using FFT the amplitude spectrum and power spectrum is determined. From those graphs a concrete information characterizing the topographical surface in the frequency domain can be obtained. The example of periodogram is presented on Fig 2.

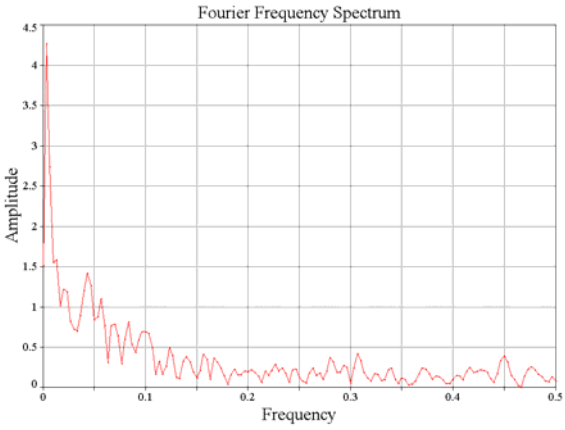

Figure 2. The example of a periodogram

The cut-off frequency can be determined by the graph analysis. That parameter posses the following properties (Hassan, 1988):

- At a low frequencies, below the cut-off frequency, the periodogram values come out mainly from the true elevations of the topographical surface
- At the frequencies close to the cut-off frequency the influence of noise (field objects) and the true terrain elevations is identical.
- At the frequencies higher than the cut-off frequency, the periodogram values come out mainly from the noise. In this area of the periodogram mainly the small values appear, and the curve shape is rather irregular.

Determination of the filter order using the geostatistical theory

The innovative value of geostatistics, invented by George Matheron in Fontainebleau (France) in 1960-ies, consists in treating an analysed parameter, in our case it is the elevation, as a ″regionalized variable″ (Matheron, 1962, 1963). The values of that ″regionalized variable″ are a function of locations of measured points. The structure of variability is described in a synthetic form by the so-called semivariogram, which determines dependence between the average variance of the differences of elevation values (semivariance) and distances between the surveyed points.
For a regular grid of measured points the semivariance values $\gamma(d)$ are given by equation:

$$\gamma(d) = \frac{\sum_{i=1}^{n}[z(x_i) - z(x_i + d)]^2}{2n} \qquad (9)$$

where $z(x_i)$, $z(x_i+d)$ – the elevations values in the points separated by the distance d
n – the number pairs of surveyed points separated by distance d

The semivariogram is a plot of semivariance as a function of distance between measured points (Fig. 3).

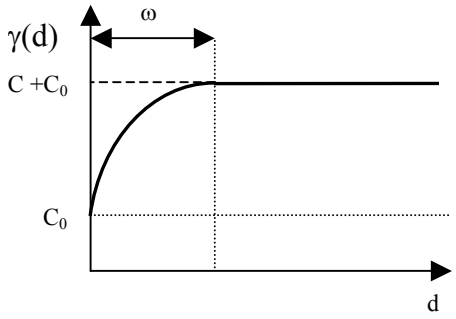

Figure 3. A hypothetical semivariogram. C_0 - a local variability in the scale of a single sample, so-called the nugget effect; C - the sill value; ω - the influence range; d – distance (lag); $\gamma(d)$ - semivariance

To enable a quantitative assessment of variability of elevations on the tested surfaces, it is necessary to estimate the empirical semivariograms by the simple analytical functions, which can be treated as the geostatistical variability models. In geostatistics occur a number of semivariogram models, among others: linear, spherical, logarithmic, Gaussian, exponential, etc (David, 1998). The analysis of theoretical semivariograms provides the so-called range of influence ω (see Fig. 3). The range can be used as a measure of spatial dependency. This range it is a greatest distance from the analysed point to the surrounding points, which have real influence on the value of the interpolated parameter assigned to the analysed point. In a more general sense, range of influence is used to predict optimal filter order for use in filtering.

.

3. THE EMPIRICAL RESEARCH

3.1 The test surfaces

A laser data, concerning two test fields (Vaihingen/Enz, and Stuttgart), provided by the OEEPE project, were used (Pfeifer et al., 2001). The topographical surfaces of those two test fields are very rough, of diversified shape and field objects.
In the both cases the first and last laser impulse reflection was registered together with the returned laser impulse intensity.

The test site nr 1 (Stuttgart)

The highly urbanized area with a steep slopes, mixture of vegetation and buildings on hillside. The distance between the laser measured points ranges from 1 to 1.5m.

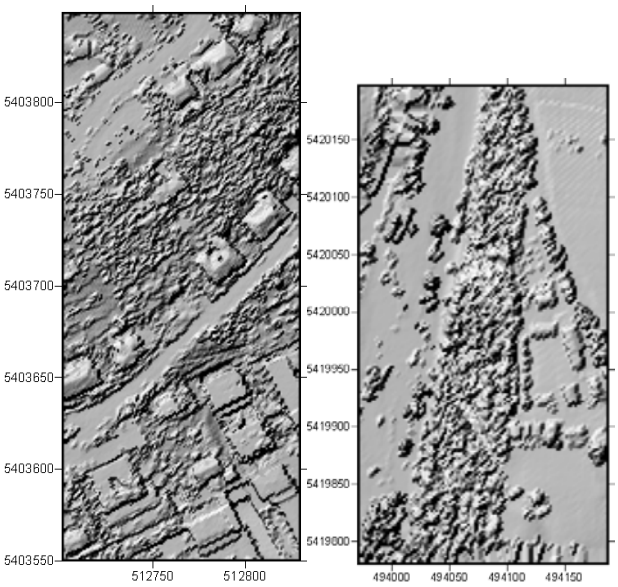

Figure 4. The test site 1 and 2

The test site nr 2 (Vaihingen)

This test field is located in a rural terrain, of steep slopes and lush vegetation. The distance between the laser measured points ranges from 2 to 3.5m.

4. RESULTS

4.1 The accuracy assessment

The set of a reference data

The reference data was generated by the hand-controlled filtering of the cloud of laser points. In the filtration process the knowledge about the tested area, and aerial photographs were used. After the hand-classification all the classified points were properly assigned to two groups: bare earth, and object.

The quantitative assessment of data

Considering the possibility of the easy comparison of presented filtration algorithm with the algorithms used in the ISPRS tests (see paragraph 4.2), it was decided to use the same method of the accuracy assessment.
In the accuracy analysis there were considered two types of errors. The Type I classification error appeared, when the reflecting laser point was assumed to be a terrain feature point instead to be properly classified as a ground point. In opposite, the Type II classification error was noticed when a laser-scanned point was wrongly assigned to the ground points, instead to the feature points.

			Filtered	
			Bare Earth	Object
Site 1	Reference	Bare Earth	14195	6880
		Object	**558**	14386
Site 2		Bare Earth	11973	626
		Object	**198**	3494

Table 5. Results of filtering for site 1 and 2. Bold – number of Type II errors. Underlining – number of Type I errors.

For the groups of errors of Type I and Type II, there were calculated the following statistical functions of deviations of the estimated values from the true values: the mean, RMSE, and magnitude of the maximum and minimum error (Tab. 6).

		%	Min. [m]	Max. [m]	Mean [m]	Std Dev. [m]
Site 1	Type I	32.6	-4.6	5.9	-0.1	0.7
	Type II	**3.7**	-11.5	17.8	-1.3	2.8
Site 2	Type I	5.0	-1.28	1.53	-0.01	0.36
	Type II	**5.4**	-5.14	1.68	-0.63	0.92

Table 6. Results of filtering for site 1 and 2.

4.2 Comparison with the selected filtration algorithms

The directives of the WG III/3 ″3D reconstruction from airborne laser scanner and InSAR data″ gave in 2002 a push to the research project concerning comparison of the existing methods of automatic laser data filtering (Sithole, Vosselman, 2003). The main objective of that research project it was determination of functionality of the filtering algorithms in a certain field conditions (mainly the configuration of topographical surface, and different types of the terrain features).
The presented algorithm was compared with the eight methods reported to the ISPRS test. To get the reliable comparison factors, the same accuracy measures were used. The graphical visualization of the comparison results is shown on the Fig. 7 and Fig. 8.

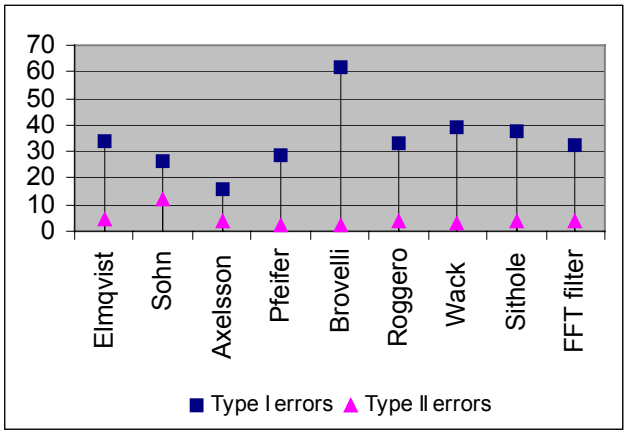

Figure 7. Comparison of the reliability factors for the Site 1. The errors Type-I and Type-II are shown in % for each of the considered filtration algorithms

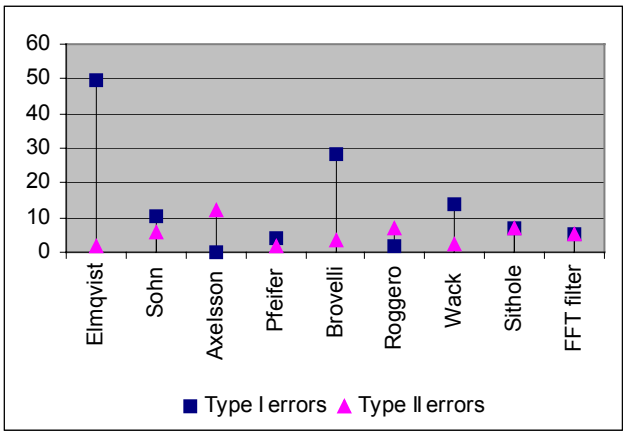

Figure 8. Comparison of the reliability factors for the site 2. The errors Type-I and Type-II are shown in % for each of the considered filtration algorithms

4.3 The algorithm modification – intensity of the first and last reflection

In our algorithm the filtration errors appeared in the first place on the forest fields.Therefore a special effort was made to analyze some additional data collected during the laser scanning. The intensities of the first and the last returned impulses were compared for the area of the selected small test fields of the site 2. Also the spatial distance between those two impulses was calculated. It was noticed that for over 45% that distance was smaller than 0.5 m. (see Fig. 9).

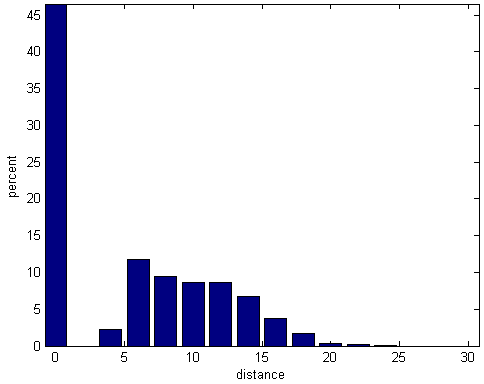

Figure 9. Histogram of first-last impulse distances for forest area.

Unfortunately, there was not discovered any precise correlation between the distance of the first-last impulse and the category of reflecting surface (bare earth or object). Only for the first-last impulse distance exceeding 15 m (Fig. 10)., we can with high probability assume, that the last impulse was really reflected by the bare earth.

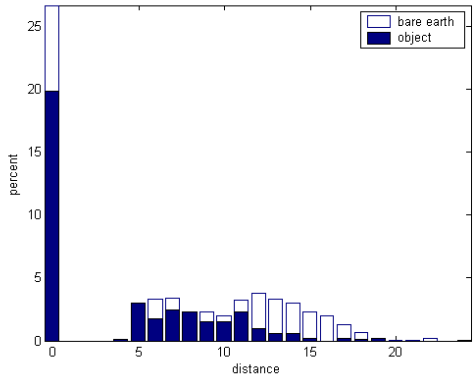

Figure 10. The relationship between the firs-last impulse distance, and the type of surface (bare earth, object) for forest area

In the next step of our research, there was analyzed the relationship between the intensity of impulse reflection and the type of reflecting surface. It was expected that the surface type (object – trees or bare earth) can be recognized by the intensity of reflected impulse. In the investigated case the intensity value of returned pulses ranged from zero to 190 relative units (see Fig. 11). Only for the intensity value greater than 190 relative units (Fig. 11) the points reflecting the last impulse can be qualified, with a high probability, as a bare earth.

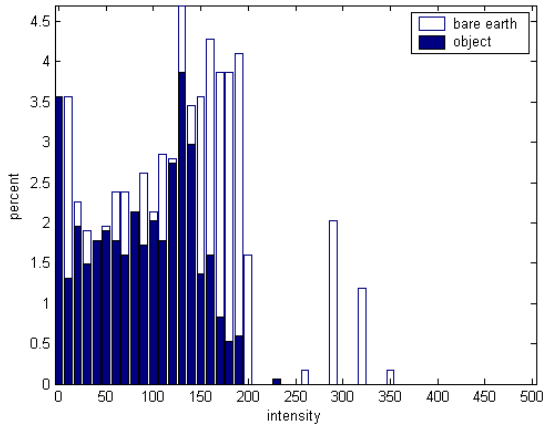

Figure 11. The relationship between the last impulse intensity for the bare-earth and object reflecting surfaces for forest area.

5. CONCLUSION

In this paper a FFT based method of filtering of airborne laser scanner data has been presented. The method is directly derived from the signal processing theory. It is so, because we introduce the laser points location as an independent variable, and we treat the terrain and the terrain features represented in a discrete form (DEM), as a discrete signal that records the elevation variations for data points in a x- y location.

The results from presented algorithm were compared with both reference data and with filtering results of eight methods reported to ISPRS test. Our experimental results reveal that quality of derived DTM is quite high. This algorithm allows for separation of the urbanized areas with high reliability. However filtering of forest areas has been the most difficult problem.

The simultaneous recording of location and intensity of first and last pulse can offer the possibility to improve reliability of filtering, also for forest areas. Different surfaces reflect pulses differently and therefore it may be possible to use this information in classification. As far as the reflection-intensity value is concerned only the points of the values greater than 190 relative units can be qualified, with a high probability, as the earth.

References:

Axelsson P., 2000. DEM generation from laser scanner data using adaptive TIN models. In: International Archives of Photogrammetry and Remote Sensing, Amsterdam, the Netherlands, Vol. XXXIII, Part B4, pp. 110-117.

David M., 1977. Geostatistical ore reserve estimation. Elsevier. Amsterdam, pp.102-112.

Flood M., 2001. Laser altimetry: from science to commercial lidar mapping. Photogrammetric Engineering and Remote Sensing, 68(9), pp. 1209-1217.

Hassan M., 1988. Filtering digital profile observations. Photogrammetric Engineering & Remote Sensing, 54(10), pp. 1391-1394.

Ionescu I., 1996. On the technique for terrain roughness determination. In: International Archives of Photogrammetry and Remote Sensing, Vienna, Austria, Vol. XXXI, Part B3, pp. 343-348.

Kilian J., Haala N., Englich M., 1996. Capture and evaluation of airborne laser altimeter data. In: International Archives of Photogrammetry and Remote Sensing, Vienna, Austria. Vol. XXXI, Part B3, pp.383-388.

Kokesz Z., 1984. Geostatystyczna charakterystyka złóż siarki na potrzeby rozpoznawania i szacowania zasobów. Praca doktorska. Krakow, Poland.

Marmol U., 2002. Analiza częstotliwościowa jako metoda filtrowania profili powierzchni topograficznej. Ogólnopolskie Sympozjum Naukowe „Fotogrametria i Teledetekcja w społeczeństwie informacyjnym." Białobrzegi k / Warszawy, Poland.

Marmol U., 2003. Modelowanie reprezentacji powierzchni topograficznej z wykorzystaniem metody geostatystycznej. Geodezja 9, Krakow, Poland.

Matheron G., 1962, 1963. Traité de géostatistique appliquée, Editions Technip, Paris, France.

Pfeifer N., Stadler P., Briese Ch., 2001. Derivation of digital terrain models in SCOP++ environment. Proceedings of OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Digital Elevation Models, Stockholm, Sweden.

Sithole G., Vosselman G., 2003. Report: ISPRS comparison of filters. http://www.geo.tudelft.nl/frs/isprs/filtertest

Vosselman G., 2000. Slope based filtering of laser altimetry data. In: International Archives of Photogrammetry and Remote Sensing Vol. XXXIII, Amsterdam, the Netherlands, Vol. XXXIII, Part B3, pp. 935- 942.