MOVING POLYNOMIAL IN FILTERING OF AIRBORN LASER SCANNING DATA

Grzegorz Jóźków

Institute of Geodesy and Geoinformatics, Wrocław University of Environmental and Life Sciences,
jozkow@kgf.ar.wroc.pl

KEY WORDS: Laser scanning data, Filtration, Moving polynomial, Hierarchical method, Least squares method, Robust estimation, DTM

ABSTRACT:

The filtration of airborne laser scanning data is a form of an automatic elimination of scanning points not belonging to the modeled surface. The points that are not removed from the cloud of points become a base for the surface modeling i.e. DTM. In this work the subject of research is a terrain surface. Many researchers propose various methods of the filtration. Nevertheless a filtration constitutes still a big problem. In this work the method of filtration using moving polynomial was presented. The small rank polynomial surface was locally fitted to the measured data in the iteration process. Parameters of the surfaces were calculated based upon M-estimators of robust estimation method. In the estimation process the distance inverse function as the weighting function and the asymmetrical damping function were used. The filtration algorithm was realized using the hierarchical method. Described below moving polynomial method was implemented as an algorithm in MATLAB software environment and tested on the real airborne laser scanning data captured by Optech ALTM scanner and ScaLARS system. The results of filtration were compared with referenced data.

1. INTRODUCTION

Within the last years airborne laser scanning has become the leading technology of acquiring geometrical information about ground surface and objects existing on it. Development of new devices and processing technology delivers high resolution data faster and with better accuracy. The main problem in data processing is classification of points to proper surfaces. The subject of research in this work is the bare earth surface. Reflections from objects existing on the terrain (i.e.: roofs and walls of buildings, forests or other vegetation) become gross errors in the process of terrain surface modelling. Manual points classification is impossible – there is large number of points in the points cloud. All solutions go to automatic or semi-automatic classification of points belonging to the proper surfaces. Another way is the automatic elimination of points not belonging to the modeling surface. This elimination is called filtration. Many authors are interested in this problem and they propose various solutions based upon:

- linear prediction [Kraus, 2000; Kraus and Pfeifer, 2001; Briese et al., 2002],
- adaptive TIN models [Axelsson, 2000],
- mathematical morphology (slope adaptive filtration) [Sithole, 2001],
- data clustering analysis [Roggero, 2001],
- surface energy minimization (active shape models or flakes) [Elmqvist, 2002; Borkowski, 2004, 2005],
- wavelet domain [Borkowski and Keller, 2006].

Overview of some filtration methods, their accuracy and restrictions can be found in study “Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds “ [Sithole and Vosselman, 2004]. Based upon the analysis of the literature and the experiences of authors some assumption to the algorithms of filtration can be formulated:

- if it is possible the filtration should be carried out on the original data,
- modeled by the algorithm surface should fit well to the local terrain structures,
- additional information a-priori can be taken into account,
- algorithm should be as simple as it is possible, because there is a lot of laser scanning data.

In the context of formulated assumptions an attempt of filtration using moving polynomial [Borkowski and Jóźków, 2006] was in this work presented. The surface of low rank polynomial was locally fitted to the measured data. The unknown polynomial parameters were determined using robust estimation. In the next part of this work the numerical algorithm and the results of tests carried out on the original laser scanning data will be presented.

2. MOVING POLYNOMIAL FILTRATION

2.1 Moving Polynomial

In the $3D$ space every polynomial can be written as:

$$z(x,y) = \sum_{i,j} a_{i,j} \cdot x^i \cdot y^j \tag{1}$$

where x, y, z = polynomial coordinates
$i, j = 0, 1, 2, \dots$
$a_{i,j}$ = polynomial parameters

Only small rank polynomials have a good proprieties to approximate the terrain surface. Because of that second rank polynomial was used. This polynomial is called moving polynomial because every time it matches to the closest neighbourhood of measured point. Used polynomial model was described as:

$$z(x,y) = a_{00} + a_{10} \cdot x + a_{01} \cdot y + a_{11} \cdot x \cdot y + a_{20} \cdot x^2 + a_{02} \cdot y^2 \tag{2}$$

where x, y = coordinates of measured point
z = calculated from polynomial height
$a_{00}, a_{10}, a_{01}, a_{11}, a_{20}, a_{02}$ = polynomial parameters

Parameters $a_{i,j}$ were computed separately in each measured scanning point using least squares method with the assumption:

$$\sum_{i=1}^{n} p_i \cdot v_i^2 \rightarrow \min \tag{3}$$

where n = quantity of points belonging to the local neighbourhood of measured point
p_i = weight of point from local neighbourhood
$v_i = z(x_i, y_i) - h_i$ - residuum of polynomial surface and measured height h_i

Weights of points depended on the distance between interpolated point and points from the local neighbourhood:

$$p_i = \left(\frac{c}{d_i}\right)^r \tag{4}$$

where c, r = empirical chosen parameters to adjust influence of points which are more distant from interpolated point
d_i = distance between interpolated point and point from local neighbourhood

Solution of unknown polynomial parameters using least squares method was written in matrix notation as:

$$X = (A^T \cdot P \cdot A)^{-1} \cdot A^T \cdot P \cdot H \tag{5}$$

where $X = \begin{bmatrix} a_{00} & a_{10} & a_{01} & a_{11} & a_{20} & a_{02} \end{bmatrix}^T$ - polynomial parameters

$$A = \begin{bmatrix} 1 & x_1 & y_1 & x_1 \cdot y_1 & x_1^2 & y_1^2 \\ 1 & x_2 & y_2 & x_2 \cdot y_2 & x_2^2 & y_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & x_n \cdot y_n & x_n^2 & y_n^2 \end{bmatrix}$$

$P = diag\{p_1 \quad p_2 \quad \dots \quad p_n\}$ - weights matrix

$H = \begin{bmatrix} h_1 & h_2 & \dots & h_n \end{bmatrix}^T$ - measured heights of points from local neighbourhood

In this way different local polynomials were determined in each measured point. This polynomial approximate terrain surface in this point.

2.2 Robust Estimation

Using least squares method polynomial parameters are determined from all points, that means from points that are reflected not only from bare earth, but from objects too. In order to avoid this situation robust estimation is necessary. Points that are not reflected from terrain are regarded as gross errors. In the robust estimation weights of points that are gross errors were decreased in the iteration process:

$$u_i^{(k)} = p_i \cdot w(v_i^{(k-1)}) \tag{6}$$

where $u_i^{(k)}$ = new weights used in step *(k)* of iteration
$v_i^{(k-1)}$ = residues between approximated polynomial surface in previous *(k-1)* step of iteration and measured points
$w(v)$ = damping function

Choice of right damping function is the main issue in the robust estimation. In the work three functions were tested:

$$w(v) = \begin{cases} 1, & |v| \le \sigma \\ \dfrac{1}{1 + (\alpha \cdot |v - \sigma|)^{\beta}}, & |v| > \sigma \end{cases} \tag{7}$$

$$w(v) = \begin{cases} 1, & |v| \le \sigma \\ e^{-(v-\sigma)^2/\sigma^2}, & |v| > \sigma \end{cases} \tag{8}$$

$$w(v) = \begin{cases} 1, & |v| \le \sigma \\ \dfrac{\sigma}{|v|}, & |v| > \sigma \end{cases} \tag{9}$$

where σ = empirical chosen parameter (usually equal laser scanning RMS)
α, β = empirical chosen parameters to adjust power of weights modification

Weights of points that are nearby polynomial surface were not modified. Third damping function (9) is symmetrical, but it seems that asymmetrical functions (7, 8) will be better. Figure 1 presents all tested damping functions.

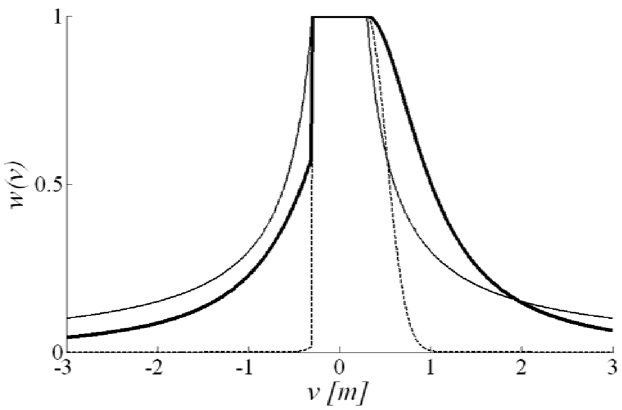

Figure 1. Damping functions ($\sigma = 0.3$): first (7) ($\alpha = 2, \beta = 2$) – wide line, second (8) – dot line, third (9) – thin line

In the iteration process new polynomial parameters were calculated as:

$$X = (A^T \cdot U \cdot A)^{-1} \cdot A^T \cdot U \cdot H \tag{10}$$

where $U = diag\{u_1 \quad u_2 \quad \dots \quad u_n\}$ - new weights matrix

Iteration process ended when parameters computed in step *(k)* were nearly the same as parameters computed in step *(k-1)*. This condition was fulfilled when all differences between residues of the same points calculated in the steps *(k)* and *(k-1)* were insignificant:

$$\left| v_i^{(k)} - v_i^{(k-1)} \right| \le \varepsilon \tag{11}$$

where ε = severity of iteration process

When differences between residues were less than ε, weights were no more modified and the polynomial parameters did not change anymore. When the iteration process was finished the local polynomials were determined in each measured point. In this way the local polynomials approximate the terrain. The last step is the comparison of polynomials surfaces and measured points. Some severity δ of filtration must be chosen:

$$\left| v_i^{(k)} \right| > \delta \tag{12}$$

where δ = severity of filtration

When the difference between measured height and interpolated from polynomial height z (2) is less than δ, point is classified as terrain point, otherwise as object point.

2.3 Hierarchical Model

In some cases even robust estimation fails. When there is more points reflected from objects than from bare earth (i.e. forest), the local polynomial parameters are determined from non-terrain points. To make the filtration more accurate the hierarchical filtration was carried out. The hierarchical model is applied in other filters too [Briese et al., 2002]. In this approach first step is the reduction of number of points that are not terrain points. In the work the hierarchical filtration was carried out in five stages:

- partition whole area to smaller sub-areas and choice for each sub-area one representative point (point with smallest height),
- polynomial interpolation using only representative points, interpolated in each point polynomial surface is the trend of terrain (trend approximate terrain without local structures),
- removing points that are not included in the cache of trend, cache of terrain trend is chosen few meters below and above trend and includes all local terrain structures,
- polynomials interpolation using points not removed in previous stage,
- comparison of interpolated polynomials surfaces and measured points.

It is possible to use hierarchical model in more stages where every time the sub-areas getting smaller, but more steps cause more computation in filtration algorithm.

2.4 Implementation Issues

Presented method is carried out on the original data. Calculation to the regular grid is unnecessary. In order to avoid numerical problems all coordinates of points can be standardized. The main issue of the hierarchical filtration is removing as much as it is possible objects' points. In order to carry out good approximation of terrain trend the sub-areas should be not too big, but too small sub-area does not remove much points. The density of laser scanning data must be taken into account too. If density is about 1 point per square meter the size of sub-area will be chosen as square of 10 m side. There is a problem with choice of representative point. Point of the smallest height may constitute multipath reflection, therefore if it is possible multipath reflections should be removed before the filtration. Last two stages of hierarchical filtration was carried out along the scheme presented on figure 2.

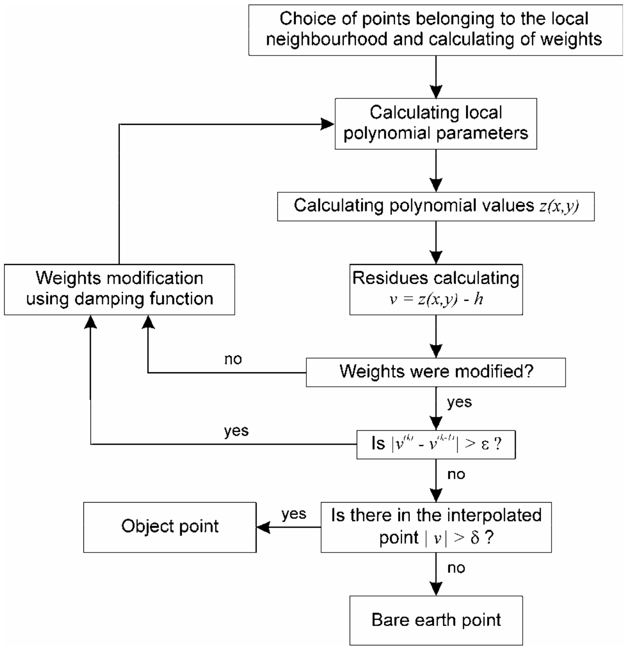

Figure 2. Scheme of moving polynomial filtration algorithm

Local neighbourhood for each point was chosen as a square. Circle neighbourhood seems to be better, but the mathematical description of square is simpler. The choice of right size of local neighbourhood is the big problem. The local terrain structures and the laser data density are important. Polynomial parameters can be estimated using only 6 measured points, but then robust estimation is impossible (residues equal zero). The bigger neighbourhood the more points to evaluate polynomial parameters, but the more computation in algorithm too. When the parameter *r* (4) is smaller the points that are more distant from the interpolated point have much influence to determine polynomial parameters. The best damping function seems to be first (7) damping function, because free choice of parameters α and β. The cache of terrain trend was chosen as 3 m above (buildings' roofs were surely removed) and below terrain trend.

2.5 Testing Data

The algorithm was tested on two examples. First testing data comes from web site http://www.itc.nl/isprswgIII-3/filtertest/Reference.zip (file samp12.txt). Points were captured with an Optech ALTM scanner and both pulse (first and last) were recorded. In the file there is 52119 scanning points with referenced flags: 0 - terrain point, 1 - non-terrain point. Referenced data was generated by manual filtering or classification [Sithole and Vosselman, 2004]. Referenced data (flags) helps in the evaluation of algorithm accuracy. Density of the cloud points is about 1 point per square meter. Data presents part of a city with high buildings, cars in the street (small objects) and some city vegetation. Second testing data comes from airborne laser scanning of Widawa river valley. These points were captured by ScaLARS system (sampled reflection of continuous wave) [Borkowski et al., 2006]. The testing file is part of one strip and contain 127175 points, but there is not referenced data. Density of measured points in this cloud of points is about 1.5 points per square meter. The second example presents part of village Szewce near to Wrocław city (Poland). This example shows quite flat terrain with several small and one big building. There is a road along the village with high trees along the road. Some middle height vegetation (stocked bushes) can be found on the example too. Measured points of both tests data are presented on figure 3 (second testing data) and 4 (first testing data). Heights are greyscale coded.

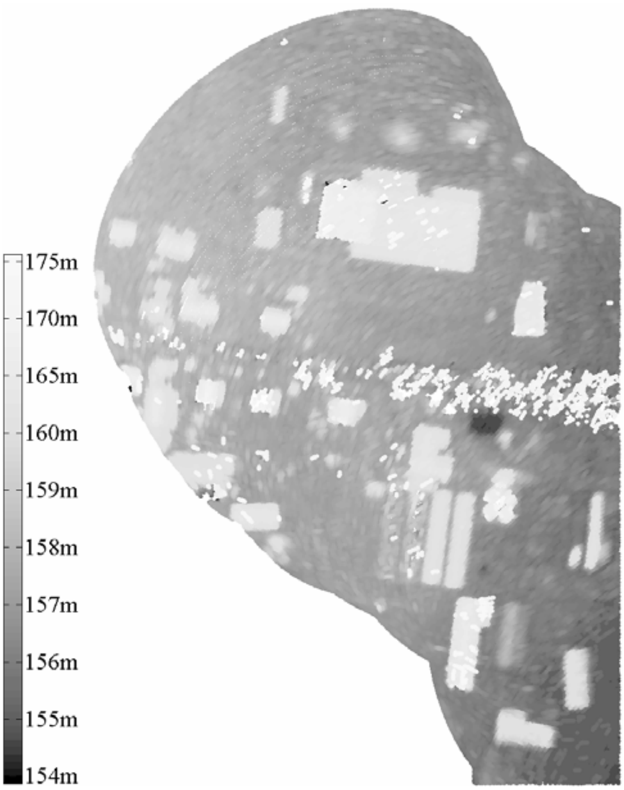

Figure 3. Measured points of second testing data

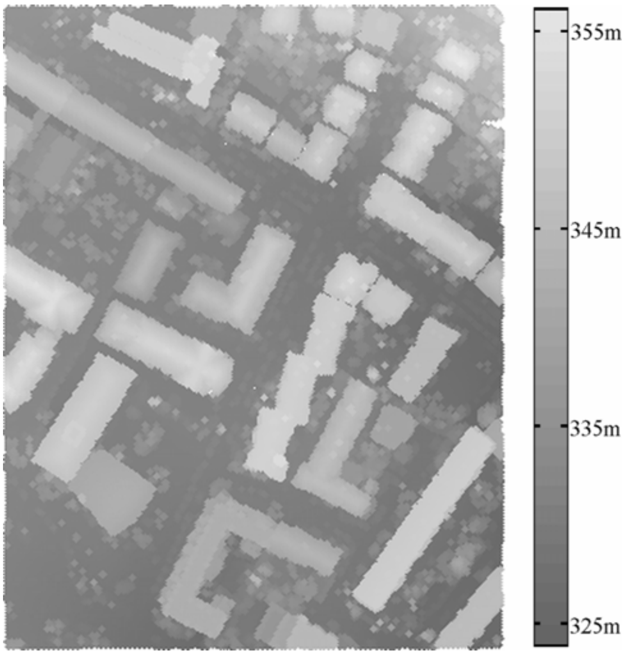

Figure 4. Measured points of first testing data

2.6 Tests Results

In order to determine best parameters of weight function (4) and best damping function there were several tests carried out on the first example. The tests showed that the best results gave the first damping function (7) with parameters: $\alpha = 3$, $\beta = 3$, $\sigma = 0.3$. Size of local neighbourhood was chosen as a square of 5 m side. Bigger side size does not results in better accuracy of filtration. The most important is parameter *r* of weight function (4). Bests results gave the small $r = 0.1$ or $r = 0.5$. Difference between these values causes in proportion of type 1 to type 2 errors. Points of first example not removed in the filtration process were compared with the referenced data. Upon this comparison, the number of points correctly classified as bare earth, object or type 1 and 2 of errors was calculated. Throughout these quantities the percentage values of filtration errors were determined. Table 1 presents results of first testing data filtration for two values *r* of weight function (4). Figure 5 shows filtration errors.

			$r=0.1$	$r=0.5$
Total (points)		e	52119	
Correct classified bare earth		a	24455	24851
Type 1 errors (bare earth as object)		b	2236	1840
Type 2 errors (object as bare earth)		c	401	803
Correct classified object		d	25027	24625
Reference	Bare earth	a+b	26691	
	Object	c+d	25428	
Filtered	Bare earth	a+c	24856	25654
	Object	b+d	27263	26465
Percentage of type 1 error		b/(a+b)	8.38%	6.89%
Percentage of type 2 error		c/(c+d)	1.58%	3.16%
Percentage of total error		(b+c)/e	5.06%	5.07%
Ratio type 1 to type 2 errors		b/c	5.58	2.29

Table 1. Results of first testing data filtration

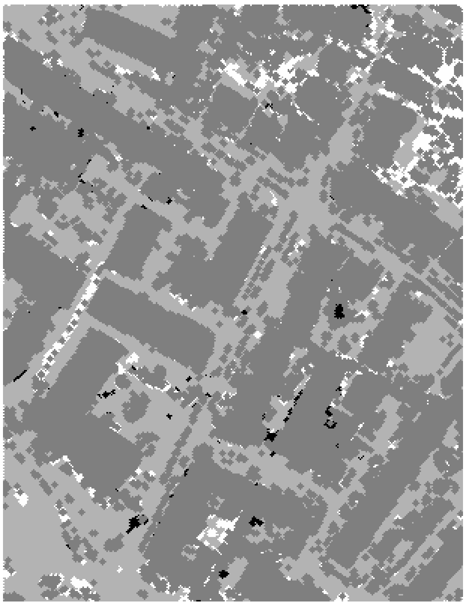

Figure 5. Filtration errors of first example, $r = 0.1$, (bright grey – correct classified bare earth points, dark grey – correct classified objects' points, white – type 1 error points, black – type 2 error points)

Second example had no referenced data so the errors could not be determined. After filtration process carried out on the second testing data, 90178 points were classified as bare earth points and 36997 points as objects' points.Figure 6 shows only points that were classified as terrains points.

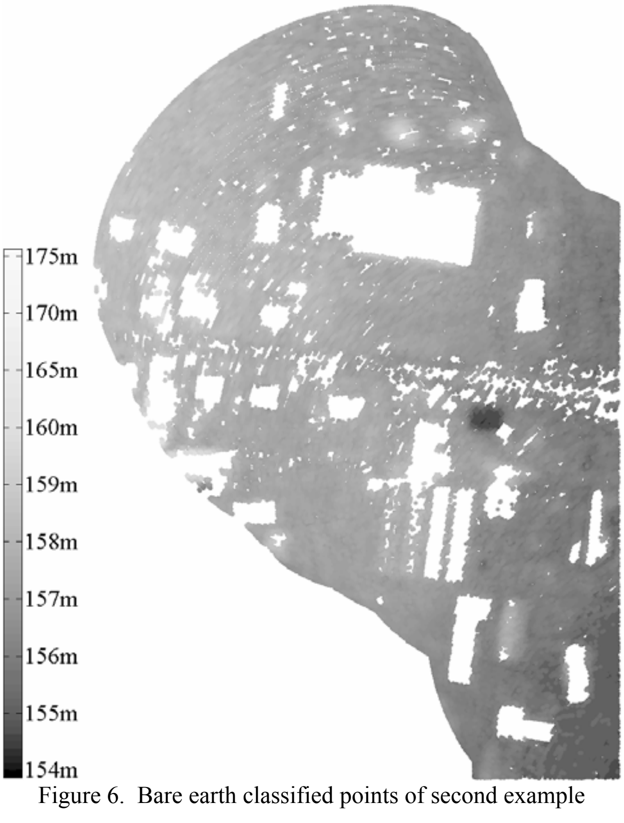

Figure 6. Bare earth classified points of second example

2.7 Discussion

In the first example almost all non-bare earth points were correctly removed as objects' points. Points that were reflections from buildings, cars, high vegetation were well classified as object points. Percentage value of type 2 errors is small for smaller r value. Problems appear only with points that are reflection from low objects and are close to other high objects i.e. small vegetation in close neighbourhood of buildings. Percentage value of type 1 errors in this example is bigger than type 2 errors, but in the process of DTM interpolation from filtered data, loss of some redundant data is less important than incorrect (type 2 errors) data. The second example had no reference data. Only the visual analysing of points that were classified as non bare earth points can be carried out. There are some objects' points that were classified as bare earth points (type 2 errors). In some places points of low parts of buildings' walls were not removed. Points belonging to the low stocked vegetation were classified as bare earth points. On the edge of points cloud several high objects' points left after filtration process. These points can be eliminated adding to the data points of next neighbouring strip. In the second example all buildings' roofs and trees along the road seem to be well classified as objects' points. Nevertheless the filtration was carried out correctly and the percentage value of total error calculated on the level about 5% is small.

3. CONCLUSIONS

In the work an algorithm of hierarchical classification of points belonging to the terrain was presented. The algorithm is based upon approximation of measured data using moving polynomial surface. Parameters of polynomials were determined based upon M-estimators of robust estimation method. Tests The carried out tests showed that best results gave the first (7) damping function and the proportion of type 1 to type 2 errors can be modified throughout changing parameter r of weight function (4). Total error of filtration was evaluated on the level about 5%. The results of filtration of data with local terrain structures as dykes or trenches could be worse. But it is possibility to modify an hierarchical approach to more steps or add to the algorithm additional information as few fixed points i.e. on the edge of dyke. These points are surely bare earth points and in the process of estimation polynomial parameters weight of these points will be never modified. An implementation to algorithm additional information is the subject of next research. Described in the paper algorithm is simple, based upon original data and throughout free choice of parameters of weight and damping functions approximate well the local terrain structures.

4. ACKNOWLEGDEMENTS

Interpolation procedures have been implemented based on MATLAB ® (licence no.: 101979), within the processing time grant awarded by Wroclaw Centre for Networking and Supercomputing.
The work financed from funds on science in 2005-2007 as a research project number 4T12E0172.

References

[Axelsson, 2000] Axelsson P., 2000. DEM generation from laser scanner data using adaptive TIN models. International

Archives of Photogrammetry and Remote Sensing, Vol. XXXIII-1/B4 110-117.

[Borkowski, 2004] Borkowski A., 2004. Modellierung von Oberflächen mit Diskontinuitäten. Deutsche Geodätische Kommission, Reihe C, Heft Nr 575.

[Borkowski, 2005] Borkowski A., 2005. Filtracja danych lotniczego skaningu laserowego z wykorzystaniem metody aktywnych powierzchni.

[Borkowski et al., 2006] Borkowski A., Gołuch P., Wehr A. Lotniczy skaning laserowy doliny rzeki Widawy. Archiwum Fotogrametrii, Kartografii I Teledetekcji, Vol. XVI, 2006.

[Borkowski and Jóźków, 2006] Borkowski A., Jóźków G. 2006. Wykorzystanie wielomianowych powierzchni ruchomych w procesie filtracji danych pochodzących z lotniczego skaningu laserowego. Archiwum Fotogrametrii, Kartografii I Teledetekcji, Vol. XVI, 2006.

[Borkowski and Keller, 2006] Borkowski A., Keller W. 2006. An Attempt to ALS-data filtering in wavelet domain. 8th Bilateral Geodetic Meeting Poland-Italy. Wrocław, 22-24 June.

[Briese et al., 2002] Briese C., Pfeifer, N., Dorninger P., 2002. Applications of the robust interpolation for DTM determination. Symposium ISPRS Commision III, Photogrammetric Computer Vision, Graz, 9 - 13 September 2002. International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV / 3A, 55 – 61.

[Elmqvist, 2002] Elmqvist, M., 2002. Ground surface estimation from airborne laser scanner data using active shape models. ISPRS, Commission III, Symposium Photogrammetric Computer Vision, September 9-13, Graz, 114–118.

[Kraus, 2000] Kraus K., 2000. Photogrammetrie. Band 3. Topographische Informatonssysteme. Dümmler, Köln.

[Kraus and Pfeifer, 2001] Kraus K., Pfeifer N. 2001. Advanced DTM generating from LIDAR data. International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV-3/W4, Annopolis, Maryland, 22-24 October, 23–30.

[Roggero, 2001] Roggero, M., 2001. Airborne laser scanning: Clustering in row data. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV-3/W4, Annopolis, Maryland, 22-24 October, 227–232.

[Sithole, 2001] Sithole, G., 2001. Filtering of laser altimetry data using a slope adaptive filter. International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV-3/W4, Annopolis, Maryland, 22-24 October, 203–210.

[Sithole and Vosselman, 2004] Sithole G., Vosselman G., 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, Vo. 59, 85-101.