MOVING POLYNOMIAL IN FILTERING OF AIRBORNE LASER SCANNING DATA – CORRECTNESS EVALUATION OF THE METHOD

Grzegorz Jóźków[1]

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. ALS data filtration is realized mainly automatically using specialist software for classification or filtration of raw data. Many researchers propose various methods of the filtration, nevertheless all algorithms do not give 100% of effectiveness. In this work the method of filtration to extract ground points 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 in the work 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. The results of filtration were compared with referenced data and the percentage values of filtering errors were calculated. Obtained results confirmed that moving polynomial method is effective – the values of filtering errors are nearly the same as values obtained using the best filtering algorithms. The correctness of presented method amount about 90%.

Keywords

Airborne laser scanning data, filtration, moving polynomial, hierarchical model, least squares method, robust estimation

1 INTRODUCTION

Airborne laser scanning is a technology used mainly for digital models building. In data processing, from acquisition to final product e. g. digital terrain model (DTM), the hardest part is an extraction of subsets, that belong to proper surfaces. An automatic elimination of points not belonging to the modelled surface or an automatic extraction of points belonging to the proper surface is called filtration. Development of new scanners delivers more data and the density of points keeps getting larger. Manual points classification is impossible – there is large number of points in the points cloud. All solutions go to 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. All automatic algorithms do not give a hundred percent of effectiveness, therefore manual check and correction of an automatic process are necessary. For large points cloud (often more than a hundred million points) even a few percent correctness increase in the automatic process brings benefit in a form of manual work decrease.

Many authors are interested in the problem of filtration of airborne laser scanning data. They propose various solutions based upon:

- linear prediction [11] [12] [7],
- adaptive TIN models [1] [2],
- mathematical morphology (slope adaptive filtration) [17][15],
- data clustering analysis [14] [10],
- surface energy minimization (active shape models or flakes) [8][9][3][4],
- wavelet domain [13][6].

In the study [16] overview of some filtering methods, their accuracy and restrictions can be found.

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.

[1] Grzegorz Jóźków, M.Sc., Wroclaw University of Environmental and Life Sciences, The Faculty of Environmental Engineering and Geodesy, Institute of Geodesy and Geoinformatics, Grunwaldzka 53, 50-357 Wroclaw, jozkow@kgf.ar.wroc.pl

In the context of formulated assumptions an attempt of filtration using moving polynomial [5] was in this work presented. Filtration agorithm using moving polynomial surface model was tested onto original airborne laser scanning data. Results of the automatic filtering were compared with referenced data and amount of bad automatic classified points was calculated. Filtering error was calculated as share of incorrect classified points in whole set of points. Percentage values of filtering errors provide the filtering correctness.

2 FILTRATION ALGORITHM

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, …

$a_{i,j}$ - 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 is matched 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:

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

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

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 and measured points

$w(v)$ - damping function

Choice of right damping function is the main issue in the robust estimation. In the work Kraus function (Fig. 1) as the best damping function [5] was used:

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

where σ - empirical chosen parameter (usually equal laser scanning RMS)

α, β - empirical chosen parameters to adjust power of weights modification

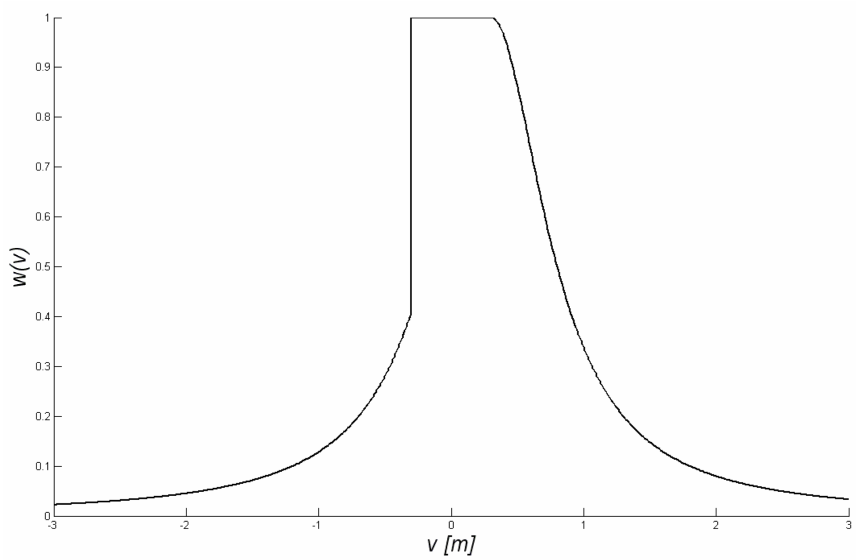

Fig. 1 *Kraus function (σ = 0.3, α = 2, β = 2)*

In the iteration process new polynomial parameters were calculated as:

$$X = (A^T \cdot U \cdot A)^{-1} \cdot A^T \cdot U \cdot H \tag{8}$$

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:

$$\left| v_i^{(k)} - v_i^{(k-1)} \right| \le \varepsilon \tag{9}$$

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. If the residues between interpolated from polynomial height *z* (2) and measured height were larger than some severity δ of filtration the point is eliminated from raw cloud as object point, otherwise classified as ground point.

2.3 Hierarchical filtering

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 [7]. Hierarchical filtration is executed in two main parts. First part is the reduction of number of points that are not suerly terrain points.

The second part is the polynomial interpolation in non-removed points. In this work first part was carried out in maximally 5 passes. In every next pass the size of subarea is getting smaller. The steps of one pass of the first part was as follow:

- 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.

After all passes proceed the second part of hierarchical filtration. It was carried out in two steps:

- polynomials interpolation using points not removed in firs part of hierarchical filtration,
- comparison of interpolated polynomials surfaces and measured points – points classification as objects‘ points or as ground points.

Numeric tests proved that more than five passes of first part are not usefull and cause in longer time of computing. Sometimes even one pass is sufficent.

3 Empirical tests

3.1 Testing data

The algorithm was tested on 15 examples of original airborne laser scanning data [18]. Points were captured with an Optech ALTM scanner and both pulses (first and last) were recorded. In each set points have referenced flags: 0 - terrain point, 1 - non-terrain point. Referenced data was generated by manual filtering or classification [16]. Referenced data (flags) helps in the evaluation of accuracy of presented algorithm. Some parameters that characterize chosen samples were as follow:

- quantity of points: 7492÷52119 points in sample,
- point spacing: 1÷1.5 m (for 9 samples) and 2÷3.5 m (for 6 samples),
- point density: 0.67 points per square meter (for 9 examples) and 0.18 points per square meter (for 6 examples).

The work [16] describe detailed (type of terrain, type of existing objects, etc.) all of 15 samples. Few samples of raw points clouds (height coded by colour) are presented below (Fig. 2, Fig. 3).

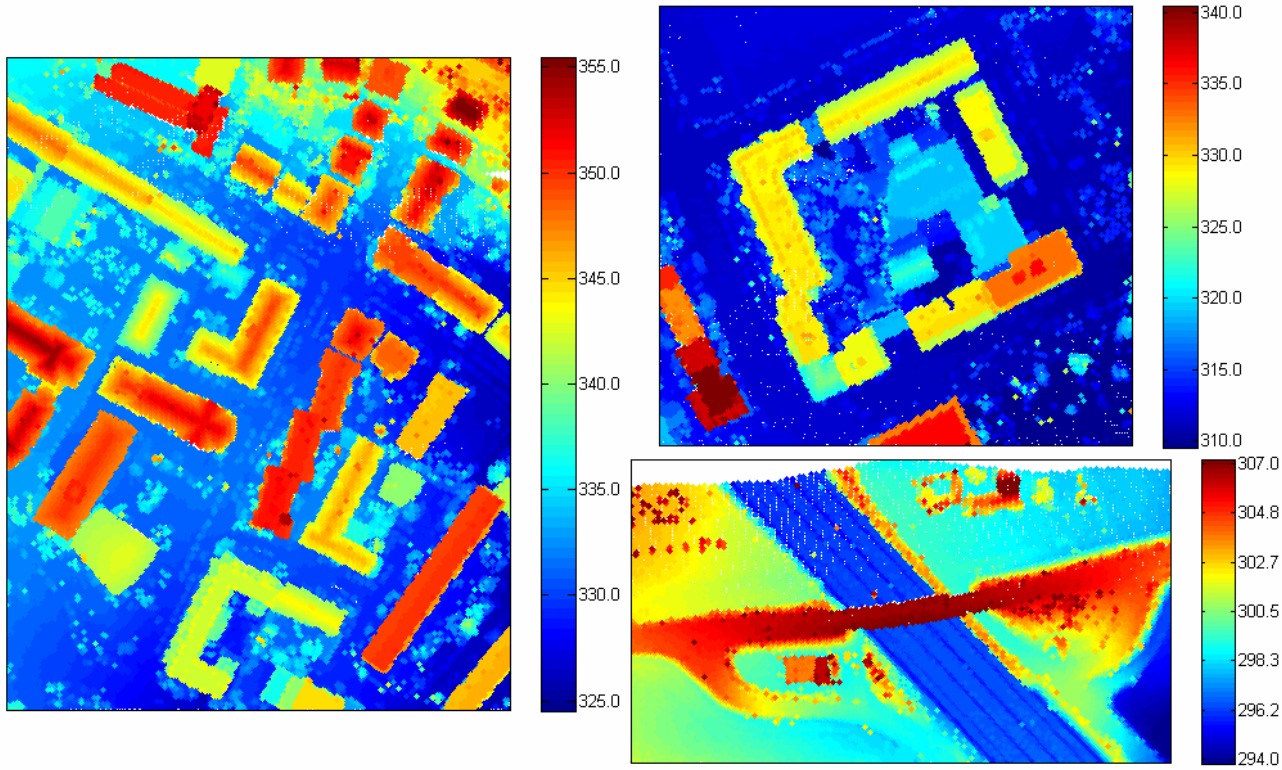

Fig. 2 *Measured data (samp12, samp31, samp71)*

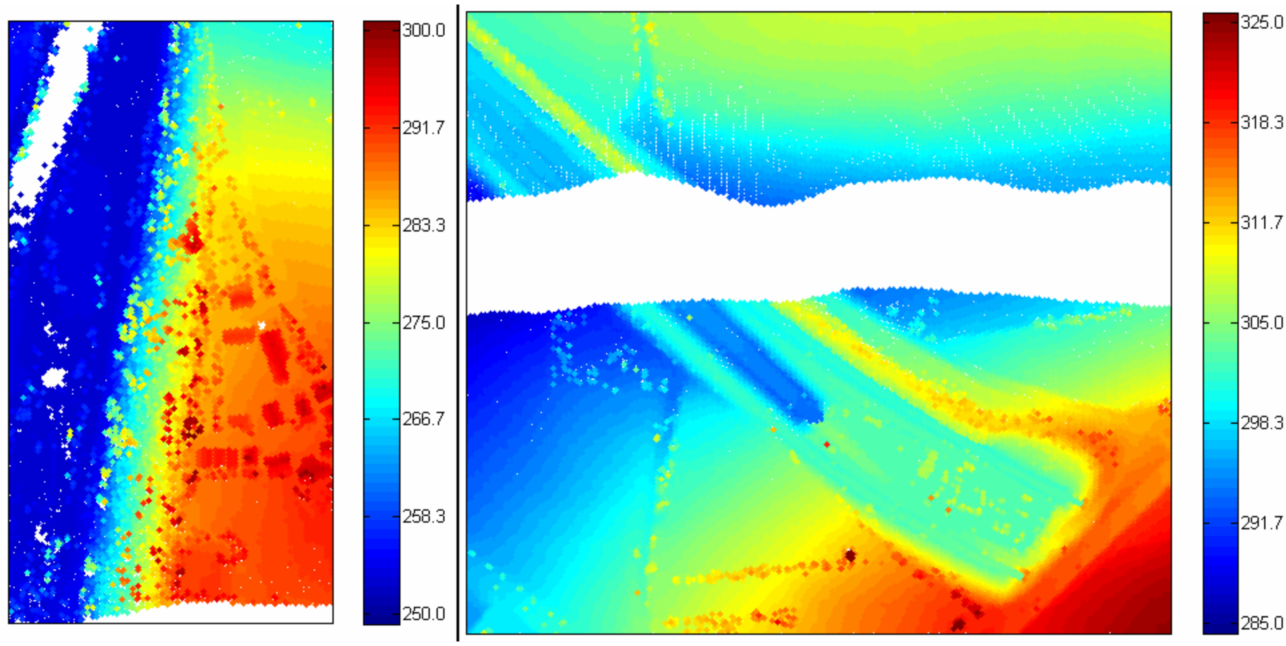

Fig. 3 *Measured data (samp51, samp61)*

3.2 Correctness evaluation

The procedure of evaluation of filtering correctness was executed based upon comparison of referenced data with set of points that were results of automatic filtering. Referenced data was two subsets detached based upon flags in testing samples:

- correctly classified terrain points – set P,
- correctly classified objects' points – set Q.

In the result of automatic filtering carried out onto measured data there were two parallel subsets obtained:

- points classified as terrain points – set R,
- points classified as objects' points – set S.

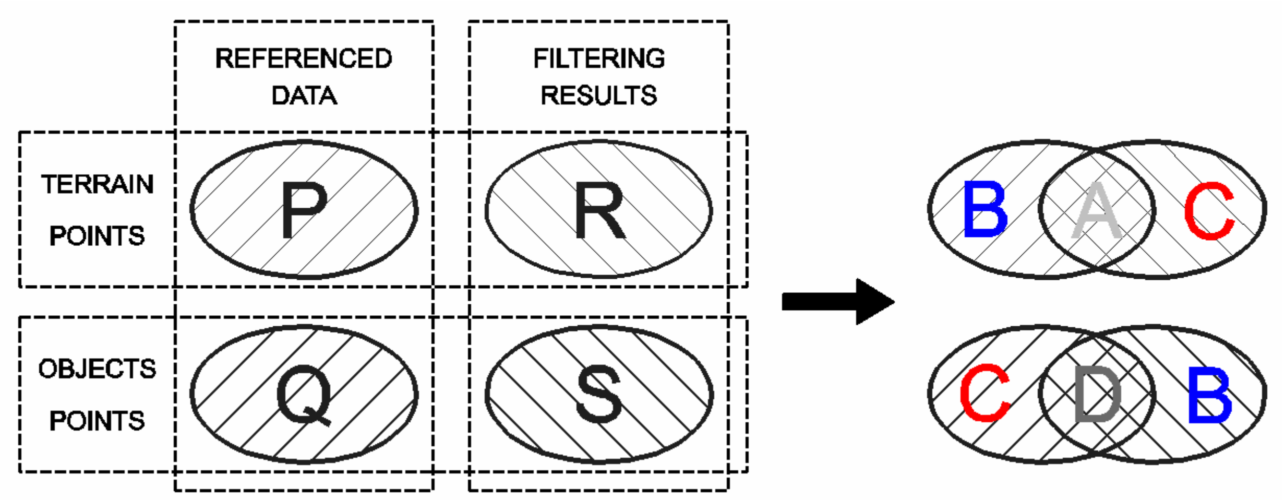

Fig. 4 *Comparison of referenced data with data after filtering*

The verification of correctness of moving polynomial method filtering depended on the comparison (Fig. 4) in pairs (P-R and Q-S) sets described above. In the result of sets operation (overlap or difference) executed onto both sets pairs there were other four sets obtained:

- set A – overlap of P and R sets ($P \cap R$) – terrain points that algorithm classified correctly (Fig. 5, Fig. 6, Fig.8 – bright gray colour of points),
- set B – difference between P and R sets ($P \setminus R$) – terrain points that algorithm classified incorrectly as objects' points (filtering error type I), (Fig. 5, Fig. 6, Fig. 8 – blue colour of points),

- set C - difference between R and P sets (R \ P) – objects' points that algorithm classified incorrectly as terrain points (filtering error type II), (Fig. 5, Fig. 6, Fig. 8 – red colour of points),

- set D – overlap of Q and S sets (Q ∩ S) – objects' points that algorithm classified correctly (Fig. 5, Fig. 6, Fig. 8 – dark gray colour of points).

Counting the quantities a , b , c , d of subsets A, B, C, D and calculating participation of incorrectly classified points in total number of points there is possibility to evaluate percentage errors of filtering. There were calculated three kinds of percentage errors.

Percentage error type I – participation of points that were errors type I in the set of real terrain points:

$$\sigma_1 = \frac{b}{a+b} \tag{11}$$

Percentage error type II – participation of points that were errors type II in the set of real objects' points:

$$\sigma_2 = \frac{c}{c+d} \tag{12}$$

Total percentage error – participation of all incorrectly classified points in the whole set of measured data:

$$\sigma = \frac{b+c}{a+b+c+d} \tag{13}$$

Percentage effectiveness of filtering could be estimated analogically as calculation of share of well classified points in set of measured data. Percentage effectiveness is the difference between value of 100% and percentage error of filtering.

3.3 Filtering results

Numeric tests of described in this work filtering algorithm were carried out onto all 15 testing samples Results of comparison of automatic filtering with referenced data (quantities of A, B, C, D sets) and percentage values of filtering errors are presented below (Tab. 1). For few testing samples the plain distribution of points belonging to the sets A, B, C, D were also presented on the pictures below (Fig. 5, Fig. 6).

Tab. 1 *Filtering results*

Sample	Points	a	b	c	d	σ_1 [%]	σ_2 [%]	σ [%]
samp11	38010	18737	3052	1433	14791	14.0	8.8	11.8
samp12	52119	52119	1688	390	25038	6.3	1.5	4.0
samp21	12960	10009	76	95	2780	0.8	3.3	1.3
samp22	32706	21417	1087	832	9370	4.8	8.2	5.9
samp23	25095	12485	738	696	11176	5.6	5.9	5.7
samp24	7492	5107	327	159	1899	6.0	7.7	6.5
samp31	28862	15411	145	218	13088	0.9	1.6	1.3
samp41	11231	5361	241	171	5458	4.3	3.0	3.7
samp42	42470	12075	368	622	29405	3.0	2.1	2.3
samp51	17845	13626	324	139	3756	2.3	3.6	2.6
samp52	22474	19160	952	415	1947	6.4	17.4	7.5
samp53	34378	31317	1672	364	1025	5.1	26.2	5.9
samp54	8608	3770	213	236	4389	7.8	3.4	5.3
samp61	35060	33240	614	52	1154	1.8	4.3	1.9
samp71	15645	13701	174	99	1671	1.2	5.6	1.7

Total percentages filtering errors (11) apart from one test were lower than 8%. In exceptions of three tests, percentage errors type I (12) and type II (13) were also very low and do not reach value of 9%. Results of filtering obtained using moving polynomial method are similar to the results obtained using other methods [16], (Fig. 7, Fig. 8).

There were not noticed typical errors – algorithm runs very well - all types of objects were quite good filtered. Polynomial surface was matched very well to all types of terrain. Executed tests showed that scanning gaps did not have influence to filtering process and largest quantity of incorrectly classified points. Large disproportion between percentage errors of type I and type II in some cases was caused by disproportion between objects' and terrain quantities of points in raw points cloud.

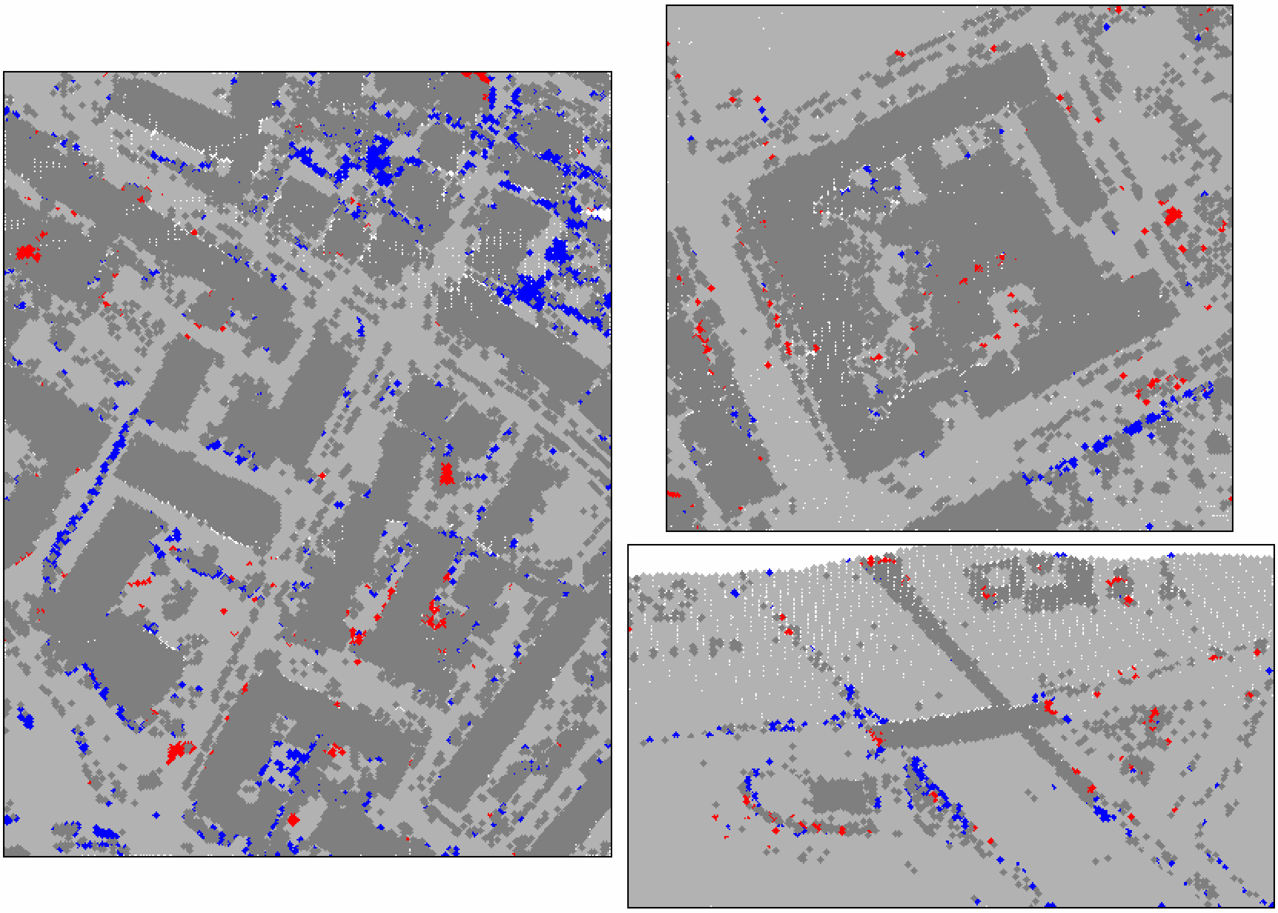

Fig. 5 *Results of comparison with referenced data (samp12, samp31, samp71)*

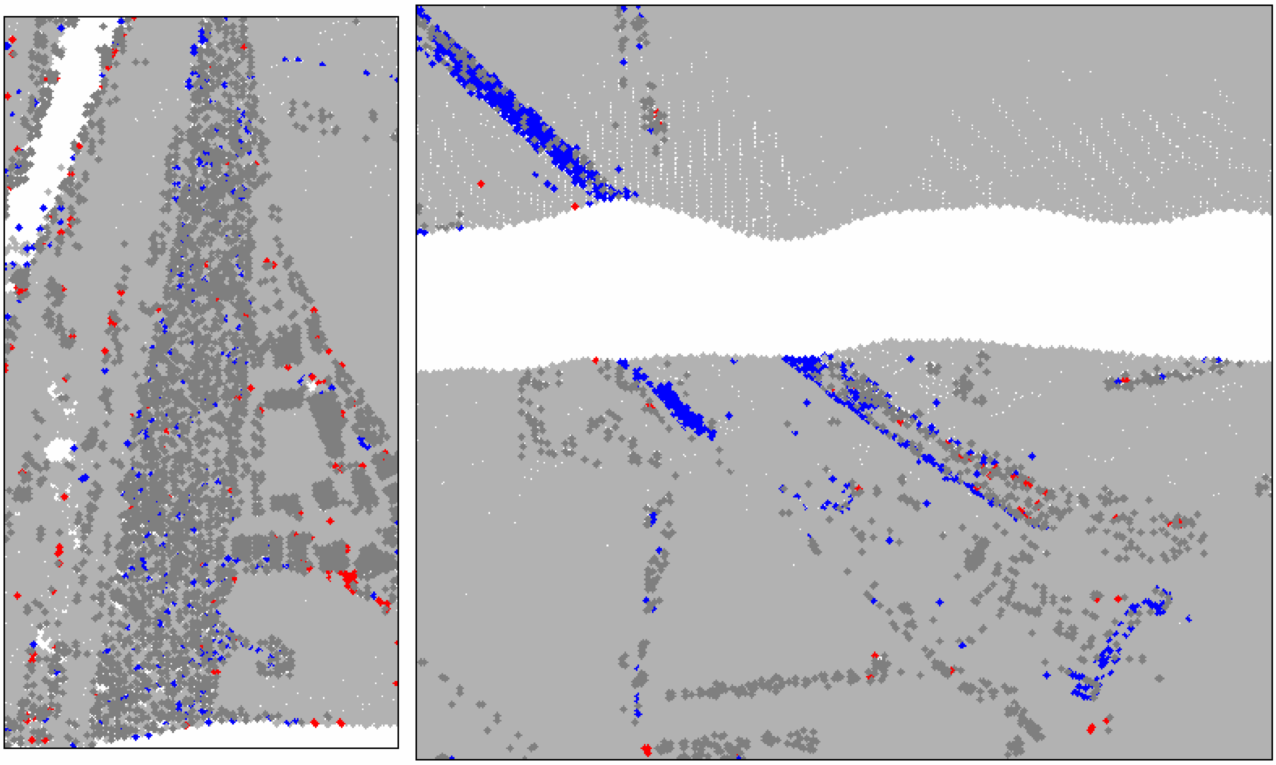

Fig. 6 *Results of comparison with referenced data (samp51, samp61)*

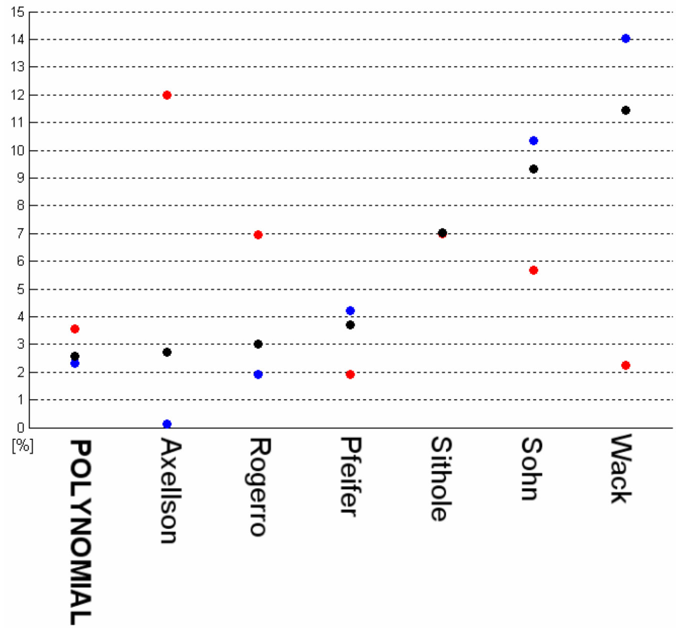

Fig. 7 *Obtained for samp51 percentage filtering errors using miscellaneous filters [16] and moving polynomial method (black – total error, blue – type I error, red – type II error)*

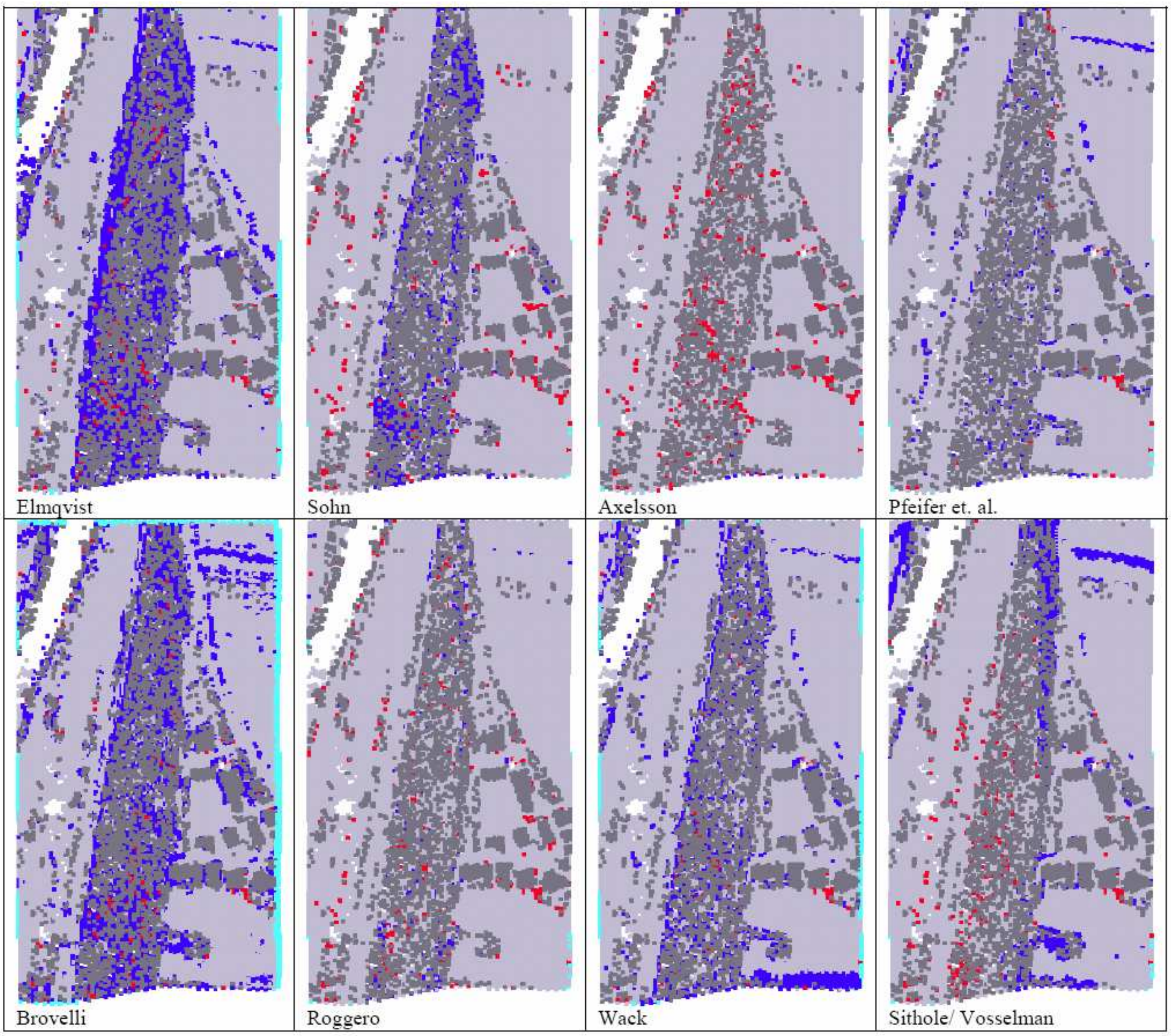

Fig. 8 *Results of comparison with referenced data (samp51) using miscellaneous filters [16]*

The lowest total percentage error was obtained for mild-sloped and flat terrain with large building, high single trees and some small objects (samp21 and samp31). The largest total error was obtained for strong-sloped terrain with miscellaneous objects on the slope (samp11). For this sample the percentage error of type I error was also the largest. For two samples (samp52 and samp53) the value of percentage error of type II is very high but amount of points that were error type II is very lower than amount of points that were error of type I. Large percentage values like these two were caused because of very high ratio bare earth points to objects point in raw points cloud. In manual correcting and checking of automatic filtering process, points of type II error are easier noticed than points of type I error.

4 CONCLUSION

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. Results of executed filtering tests were compared with referenced data therefore values of filtering percentage errors were calculated. Obtained values confirmed the usefulness the presented method in filtering of airborne laser scanning data. For each testing sample the correctness of filtering was on the level about 90%. Depending on the terrain structure and kind of terrain covering the total percentage error of filtering got values from 1.3% to 11.8%. Values of errors obtained using moving polynomial method are close to the best results obtained using other filtering algorithms (Fig. 7, Fig. 6 compare with Fig. 8).

There were not noticed typical errors of filtering – algorithm runs very well onto all types of data (miscellaneous types of terrains and terrain structures, size and kind of non-terrrain objects). Scanning gaps or edges of terrain do not causes in incorrect classification of points.

Description of the presented algorithm is very simple. All computing are executed onto original data (gridding of data is not necessary). Throughout the free choice of parameters of weight and damping functions moving polynomial surface approximate well the local terrain structures. Algorithm can be modified to take into account information a-priori as few fixed points i.e. on the edge of dyke. These points are surely bare earth points and in the process of estimation of polynomial parameters weight of these points will be never modified. The disadvantage of the presented algorithm is the amount of numeric computing – polynomial parameters are calculated in each point in multiple iterations.

5 ACKNOLEGMENTS

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 4T12E01729.

Literature

[1] Axelsson P. *Processing of laser scanner data - algorithms and applications.* ISPRS Journal of Photogrammetry and Remote Sensing, LIV(2). 1999. 138–147.

[2] Axelsson P. *DEM generation from laser scanner data using adaptive TIN models.* International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII-1/B4. 2000. 110-117.

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

[4] Borkowski A. *Filtracja danych lotniczego skaningu laserowego z wykorzystaniem metody aktywnych powierzchni.* PTIP, Rocznik Geomatyki, Vol. III, Zeszyt 4. 2005. 35-42.

[5] Borkowska A., Jóźków G. *Wykorzystanie wielomianowych powierzchni ruchomych w procesie filtracji danych pochodzących z lotniczego skaningu laserowego*. Archiwum Fotogrametrii, Kartografii i Teledetekcji, Vol. XVI, 2006.

[6] Borkowski A., Keller W. *An Attempt to ALS-data filtering in wavelet domain.* 8th Bilateral Geodetic Meeting Poland-Italy. Wrocław, 22-24 June 2006. *http://www.geo.ar.wroc.pl/8bgmpi/*

[7] Briese C., Pfeifer, N., Dorninger P. *Applications of the robust interpolation for DTM determination.* ISPRS, Commision III, Symposium Photogrammetric Computer Vision. Graz, 9-13 September 2002. 55 – 61.

[8] Elmqvist, M., Jungert, E., Persson, A. und Soderman, U. *Terrain modelling and analysis using laser scanner data*. International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV-3/W4. Annopolis, Maryland, 22-24 October 2001. 219–227.

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

[10] Filin S., Pfeifer N. *Segmentation of airborne laser scanning data using a slope adaptive neighborhood.* ISPRS Journal of Photogrammetry and Remote Sensing. 2006. 60, 71–80.

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

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

[13] Marmol U., Jachimski J. *A FFT based method of filtering airborne laser scanner data.* ISPRS Congress, Commision 3. Istambul, Turkey 12-23 July 2004. http://www.isprs.org/commission3/wg3

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

[15] Sithole, G. *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 2001, 203–210.

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

[17] Vosselman G., Maas, H.-G. *Adjustment and filtering of raw laser altimetry data.* OEEPEWorkshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models. Stockholm, 1-3 march 2001.

[18] Vosselman G. *ISPRS TEST ON EXTRACTING DEMs FROM POINT CLOUDS "A comparison of existing automatic filters"*. Delft, 2003. http://www.itc.nl/isprswgIII-3/filtertest/Reference.zip

Reviewer

Andrzej Borkowski, Ph.D., D.Sc., Wroclaw University of Environmental and Life Sciences, The Faculty of Environmental Engineering and Geodesy, Institute of Geodesy and Geoinformatics, Grunwaldzka 53, 50-357 Wroclaw, Poland, +48713205609, borkowski@kgf.ar.wroc.pl