AIRBORNE LASER SCANNING DATA FILTERING USING FLAKES

A. Borkowski, G. Jóźków

Institute of Geodesy and Geoinformatics, Wroclaw University of Environmental and Life Sciences – (borkowski, jozkow)@kgf.ar.wroc.pl

KEY WORDS: airborne laser scanning, filtering, flakes, variational modelling, filtering errors

ABSTRACT:

Generation of digital terrain model on the basis of airborne laser scanning data requires extracting exclusively these points from the raw point cloud which are reflections of a laser beam from the ground. This task is performed mainly automatically using specialist software for classification or filtration of laser scanning points. Various filtration methods have been suggested but neither of the proposed algorithms guarantees 100% efficiency. In this paper, the algorithm based on surface energy minimization is discussed. Total energy of the surface is the sum of internal and external energy. Internal energy describes geometrical properties of the modelled surface and for the introduced flakes model it equals the weighted sum of surface membrane kernel and surface thin plate kernel. External energy describes difference between estimated active surface and survey data and depends on the difference between the measured height and approximated height. As a result of total surface energy minimization, the active surface is adjusted to the terrain surface. The variational problem that occurs in the surface energy minimization was solved using the direct method (Ritz method). Flakes method was tested on authentic laser scanning data, which additionally contained reference data, i.e. correctly classified terrain points and object points. Comparison of reference data and sets of points obtained in filtration facilitated determination of the percentage values of filtering errors. The results confirmed that flakes method is effective. Filtering correctness value is similar to results which were obtained using other methods.

1. INTRODUCTION

Airborne laser scanning data is used principally for generation of digital models. The extraction of points subsets that belong to appropriate surfaces is probably the most difficult part of laser scanning data processing. Automatic elimination of points which are not a part of the modelled surface is called filtration. Development of new technologies and new measuring tools such as laser scanners provides users with more data. Density of points in point clouds increases as well. Due to a large number of scanning points manual classification of data is impossible. This problem is solved automatically by using special ALS data classification or filtration software. The applied algorithms are not completely efficient and therefore manual check and correction of the automatic process are necessary. For large data sets (which often contain over 100 million of points) the increase of correctness in the automatic process at the level of even a few percent contributes to the decrease of manual work. In this way, the final product, i.e. DTM is cheaper and can be obtained faster.

In the airborne laser scanning data processing nearly always the subset of terrain points is extracted. This subset is a basis for digital terrain modelling. Many researchers are interested in ALS data filtering because they view it as a method of terrain point extraction. They propose various solutions, some of which are used in commercial software. The automatic process is still imperfect and there is no guarantee that it is correct in 100%. Having analyzed the literature on the subject, one can group filtering methods on the basis of the used approach:

- linear prediction (Kraus, 2000; Kraus and Pfeifer, 2001; Briese et al., 2002),
- adaptive TIN models (Axelsson, 2000),
- mathematical morphology (slope adaptive filtering) (Sithole, 2001),
- data clustering analysis (Roggero, 2001),
- surface energy minimization (active contour models or flakes) (Elmqvist, 2002; Borkowski, 2004, 2005),
- spectral domain using FFT (Marmol and Jachimski, 2004) or wavelets (Vu et al., 2002; Borkowski and Keller, 2006; Wei and Bartels, 2006).

An overview of some filtration methods, their accuracy and restrictions can be found in the study (Sithole and Vosselman, 2004).

The algorithm presented below uses data transformed to regular values. Flakes model is a two-dimensional generalization of snakes model (Kass et. Al.., 1988), which was used in digital image processing. Minimization of total surface energy gives the flakes model. The flakes models, which is distinct from active contours model, (Elmqvist, 2002) is described in a different way. For the flakes model, the internal energy of the surface is a weighted sum of the membrane kernel and thin plate kernel.

The method discussed in this paper was tested on original airborne laser scanning data. Results of the automatic filtration were compared with reference data which facilitated determination of the amount of the wrongly classified points. The filtering error was computed as a share of incorrectly classified points in the whole data set. Percentage values of filtration errors indicate that filtration was correct.

2. FLAKES

Flakes model proposed for the surface is generalization of active contours model (snakes) introduced by (Kass et. al., 1988). If there are enough small and smooth elements of the surface which have some energy, the minimization of total energy of the surface will adjust this surface actively to survey data, and geometrical properties of the surface will be still defined (Borkowski, 2004, 2005).

Total energy of the surface E_{tot} is the sum of internal E_{int} and external E_{ext} energy:

$$E_{tot} = E_{\text{int}} + E_{ext} \ . \tag{1}$$

Internal energy (2) describes geometrical properties of the modelled surface. In the flakes method, internal energy is defined as the weighted sum of membrane kernel $z_x^2 + z_y^2$ and thin plate kernel $z_{xx}^2 + 2z_{xy}^2 + z_{yy}^2$:

$$E_{\text{int}} = \frac{\alpha}{2}(z_x^2 + z_y^2) + \frac{\beta}{2}(z_{xx}^2 + 2z_{xy}^2 + z_{yy}^2) \ , \tag{2}$$

where $z_x = \frac{\partial z}{\partial x}$, $z_{xx} = \frac{\partial^2 z}{\partial x^2}$, etc.

Weighting parameters α and β are chosen freely depending on implementation and geometrical properties (smoothness) of the modelled surface.
The variational problem of minimization of total energy of the surface:

$$E_{tot} = E_{\text{int}} + E_{ext} \rightarrow \min \ , \tag{3}$$

was solved using direct method (Ritz method). The searched surface is approximated with a certain functional model:

$$z(x, y) \approx f(x, y) \ , \tag{4}$$

which is a linear combination of the given base functions:

$$f(x, y) = \sum_{i=1}^{n} c_i \varphi_i(x, y) \ .$$

The base function was chosen as:

$$\varphi_i(x, y) = \varphi_i(x)\varphi_i(y) \ , \tag{5}$$

and (compare Figure 1):

$$\varphi_i(s) = \begin{cases} \frac{1}{\Delta}(s - (i-1)\Delta) & , \quad (i-1)\Delta \le s \le i\Delta \\ 1 - \frac{1}{\Delta}(s - i\Delta) & , \quad i\Delta \le s \le (i+1)\Delta \\ 0 & , \quad other \end{cases} \ . \tag{6}$$

The unknown weight coefficients (c_i) were replaced with surface altitudes on the grid. The linear conditional equation, which is the equivalent to variational problem in the point $P(i, j)$ was derived (* - Hadamard product):

$$B * Z + \frac{\partial E_{ext}}{\partial z_{i,j}} = 0 \ , \tag{7}$$

where
$$Z = \begin{bmatrix} z_{i-2,j+2} & z_{i-1,j+2} & z_{i,j+2} & z_{i+1,j+2} & z_{i+2,j+2} \\ z_{i-2,j+1} & z_{i-1,j+1} & z_{i,j+1} & z_{i+1,j+1} & z_{i+2,j+1} \\ z_{i-2,j} & z_{i-1,j} & z_{i,j} & z_{i+1,j} & z_{i+2,j} \\ z_{i-2,j-1} & z_{i-1,j-1} & z_{i,j-1} & z_{i+1,j-1} & z_{i+2,j-1} \\ z_{i-2,j-2} & z_{i-1,j-2} & z_{i,j-2} & z_{i+1,j-2} & z_{i+2,j-2} \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & e & c & e & 0 \\ e & d & b & d & e \\ c & b & a & b & c \\ e & d & b & d & e \\ 0 & e & c & e & 0 \end{bmatrix}, \ a = \frac{8}{3}\alpha, \ c = \frac{2}{3}\beta,$$

$$b = -\frac{1}{3}(\alpha + 17\beta), \ d = -\frac{1}{3}(\alpha - 2\beta), \ e = \frac{1}{6}\beta \cdot$$

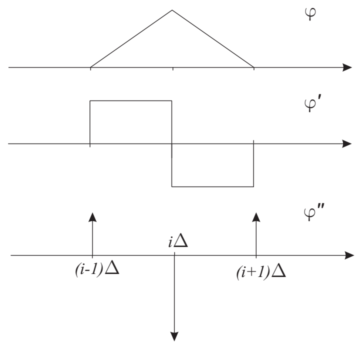

Figure 1. Base function and its derivatives

In the presented flakes model the external energy E_{ext} is described as discrepancy between survey data and estimated flakes surface. External energy (Figure 2) depends on the free parameter *s* and deviation r between measured height (z_d) and approximated height (z_t) in step *t* of iteration process:

$$\frac{\partial E_{ext}}{\partial z_t} = \begin{cases} r & for \quad r < 0 \\ re^{-r^2/s^2} & for \quad r \ge 0 \end{cases} \ . \tag{8}$$

System of equations (7) is given for each point and it is solved in the iteration process:

$$z_t = (A + I)^{-1}(z_{t-1} + E_{ext,t-1}) \ . \tag{9}$$

where z_t , z_{t-1} - heights of points estimated in step *t* and *t-1*,
A - band matrix depended on a , b , c , d parameters of the matrix B (Borkowski, 2004),
I - identity matrix,
$E_{ext,t-1}$ - external energy calculated in *t-1* step.

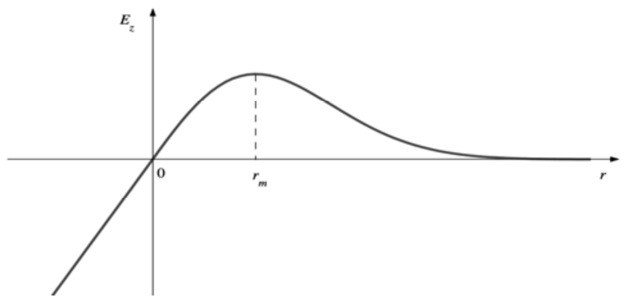

Figure 2. Model of external energy

It is essential to define starting surface z_0 and external energy $E_{ext,0}$ in the iteration process. This process ends if the surfaces estimated in the last two steps are nearly the same. In this case, the active surface has the minimal energy and is adjusted to the terrain surface.

In the last stage of filtration survey data is compared with the active surface which was estimated in the last step of iteration. If residues between them are small, points are classified as terrain points, otherwise as object points.

3. EMPIRICAL TESTS

3.1 Testing data

The presented algorithm was tested on real ALS data. Points were captured with an Optech ALTM scanner and both reflections (first and last) were recorded. Data also included reference data recorded as flags: 0 - terrain point, 1 – non-terrain (object) point. Such a solution facilitated extracting two subsets for each test set: a subset of correctly classified terrain points and a subset of correctly classified object points. These subsets permitted evaluation of the filtering correctness.

Flakes algorithm was tested carefully on 7 out of 15 test sets (Vosselman 2003). The chosen sets are characterized by the following features:

- number of points: 8608÷52119 points in sample,
- point spacing: 1÷1.5 m (for 4 sets) and 2÷3.5 m (for 3 sets),
- density of points: 0.67 points per square meter (for 4 sets) and 0.18 points per square meter (for 3 sets).

The detailed description of all 15 sets, rationales for their choice, type of terrain, non-terrain objects, the way how reference data was collected are presented in the work (Sithole and Vosselman, 2004). Exemplary test sets are presented in the (Figure 3, 4, 5 – heights coded in grey shades).

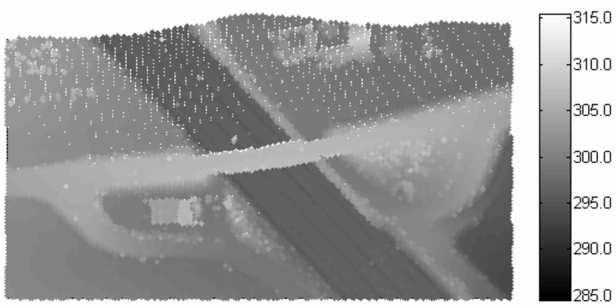

Figure 3. Test data – test 71

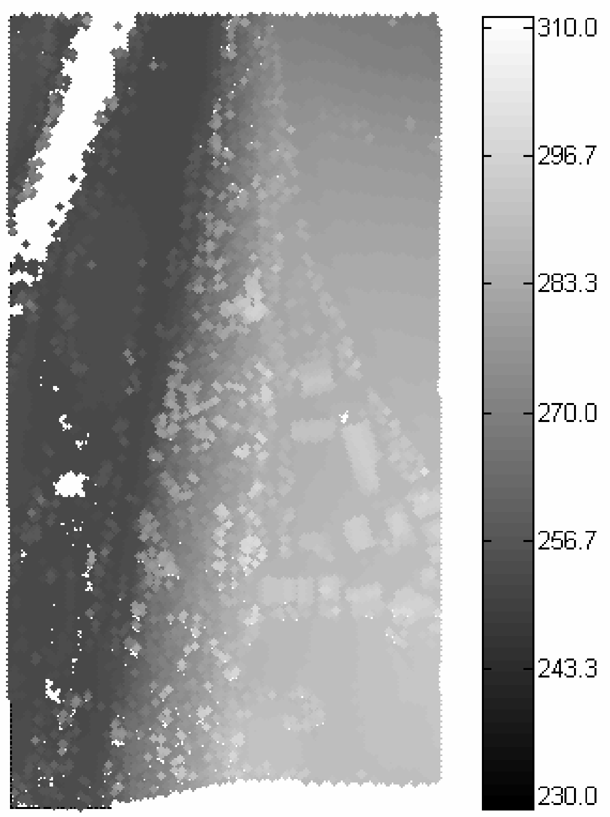

Figure 4. Test data – test 51

Figure 5. Test data – test 12

3.2 Correctness evaluation

Filtering correctness was evaluated by comparing reference data with set of points which were outcomes of automatic filtration. Reference data belong to two subsets separated from test sets:

- correctly classified terrain points – set P,
- correctly classified object points – set Q.

As a result of the automatic filtration, in which survey data was used, two sets were obtained:

- points classified as terrain points – set R,
- points classified as object points – set S.

Verification of correctness of flakes method depended on comparing pairs of sets described above (Figure 6). Operation on sets (intersection or complement) performed on both set pairs brought another four sets:

- set A – intersection of P and R sets (P ∩ R) – terrain points classified correctly by the algorithm (Figure 7, 8, 9 – medium-grey colour of points),
- set B – complement of P and R sets (P \ R) or complement of S and Q sets (S \ Q) – terrain points which algorithm classified incorrectly as object points (filtering error type I), (Figure 7, 8, 9, – bright grey colour of points),
- set C – complement of R and P sets (R \ P) or complement of Q and S sets (Q \ S) – object points which algorithm classified incorrectly as terrain points (filtering error type II), (Figure 7, 8, 9 – black colour of points),
- set D – intersection of Q and S sets (Q ∩ S) – object points which algorithm classified correctly (Figure 7, 8, 9 – dark grey colour of points).

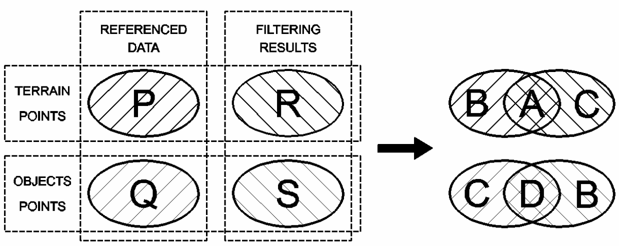

Figure 6. Comparison of reference data with filtration outcomes

Counting the number of a, b, c, d points in A, B, C, D sets respectively and calculating the share of incorrectly classified points in the whole set, it is possible to evaluate percentage errors of filtering. Three kinds of percentage errors were determined (Sithole and Vosselman, 2004).
Percentage error type I – participation of points that were errors of type I in the set of real terrain points:

$$\sigma_1 = \frac{b}{a+b}\,. \tag{9}$$

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

$$\sigma_2 = \frac{c}{c+d}\,. \tag{10}$$

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

$$\sigma = \frac{b+c}{a+b+c+d}\,. \tag{11}$$

Percentage effectiveness of filtering can be estimated analogically as the proportion of correctly classified points to the whole set of survey data. Percentage effectiveness of filtration equals the difference between 100% and percentage error of filtration.

3.3 Results

Numeric tests of the filtering algorithm described in this paper were done on seven test sets. Many different parameters of the active surface were tested. Results of comparison of automatic filtration with reference data (number of A, B, C, D sets) and percentage values of filtration errors are presented in the (Table 1).
Percentage filtration errors of type I (9) as well as of type II (10) and total (11) for each set were lower than 10%. Results of ALS filtration with the flakes method are similar to the results obtained using other methods (Sithole and Vosselman, 2004), (Figure 10).
Except for terrain break lines, no typical filtration errors were recorded – algorithm runs very well - all types of non-terrain objects were correctly removed from the source data set. Flakes surface may be adjusted well to all types of terrains. Tests showed that scanning gaps did not affect filtration process and did not increase the amount of incorrectly classified points. In some cases large disproportion between percentage errors of type I and type II was caused by disproportion between object and terrain points in the “raw points cloud”.
The lowest percentage error of filtration was recorded for the terrain with mild slopes, single trees and a high building (test 21). The largest percentage error was recorded for the area with many artificial structures and break lines (test 22). For this test set the percentage filtration error of type II was also the largest. The reason, why the percentage error of type II had such a big value, was a big ratio of terrain points to object points in raw laser scanning data.
Apart from one test sample (test 54 – more object points than terrain points in source data), there were always more points which were type I errors than points of type II error. It is caused by the fact that the algorithm based on flakes method, like most of algorithms, eliminates as many object points as possible. At the same time, many terrain points are also eliminated as scanning errors of type I.
The loss of redundant points (type I error) in the process of DTM generation with ALS data is less dangerous than incorrect data (type II error). However, during manual correction and control of automatic filtration process, points of type II error are detected easier than points of type I error.
Errors of type I for the flakes method were usually:

- points on the edge of elaboration area,
- points near the terrain break lines,
- randomly located single points.

Points classified as filtration errors of type II were usually:

- randomly located single points,
- points that were reflections from very low and very small objects,
- groups of points that create artefacts.

Test	a	b	c	d	σ_2 [%]	σ_1 [%]	σ [%]
12	24260	2431	693	24735	9.11	2.73	5.99
21	9939	146	112	2763	1.45	3.90	1.99
22	20939	1565	854	9348	6.95	8.37	7.40
31	15100	456	285	13021	2.93	2.14	2.57
51	13641	309	190	3705	2.22	4.88	2.80
54	3828	155	197	4428	3.89	4.26	4.09
71	13519	356	119	1651	2.57	6.72	3.04

Table 1. Results of correctness evaluation for the filtering method based on flakes

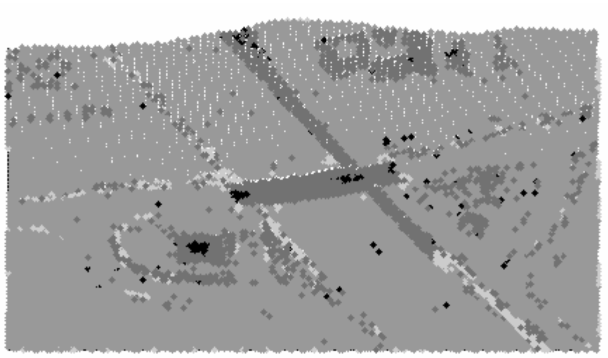

Figure 7. Results of comparison with reference data – test 71

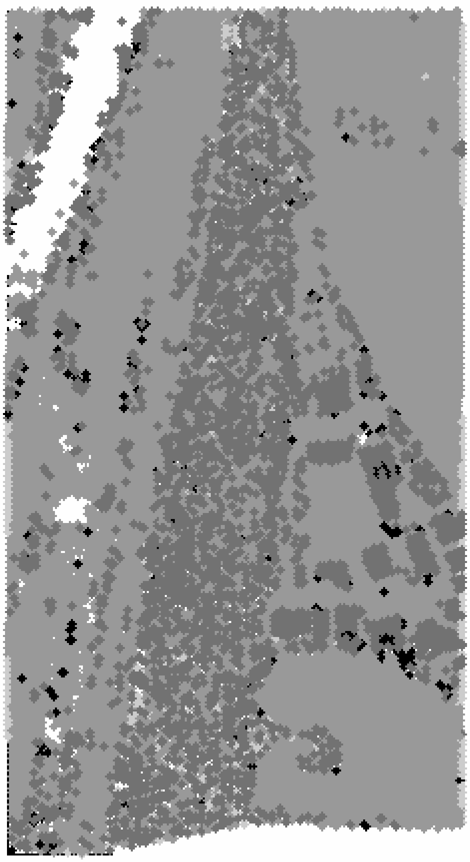

Figure 8. Results of comparison with reference data – test 51

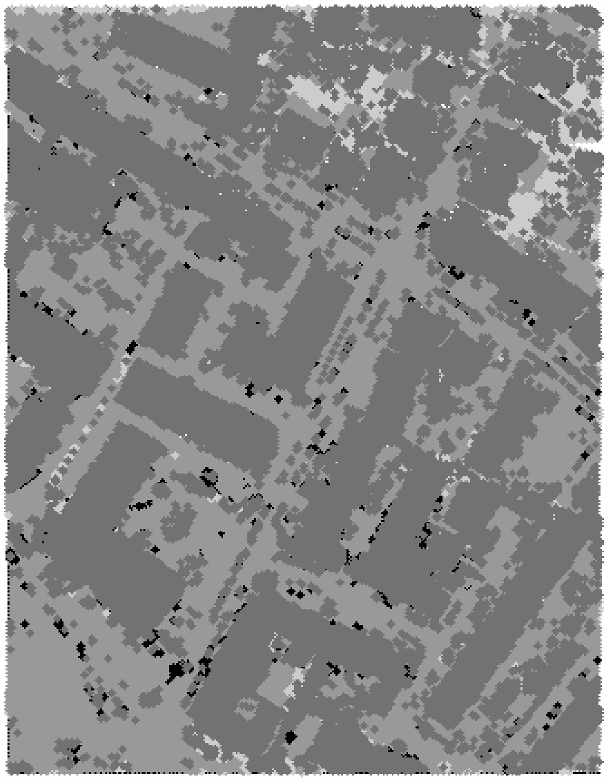

Figure 9. Results of comparison with reference data – test 12

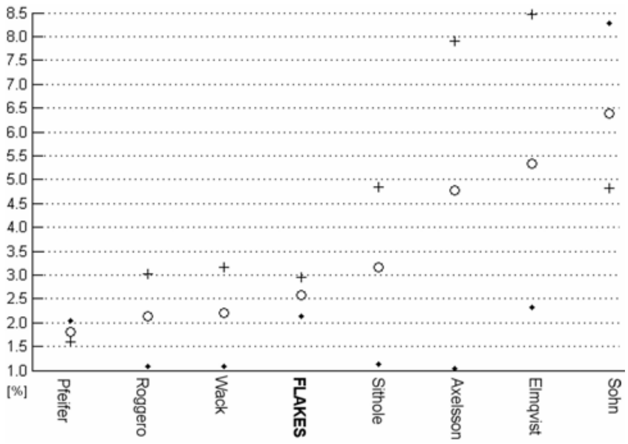

Figure 10. Percentage filtration errors for test 31 recorded by using miscellaneous filters (Sithole and Vosselman, 2004) and flakes method (circle – total error, cross – type I error, dot – type II error)

4. CONCLUSION

This paper presents the algorithm of automatic extraction of terrain points from unprocessed airborne laser scanning data. This algorithm is based on approximation of active flakes surface to survey data.

Numeric tests were done on real airborne laser scanning data. The outcomes of filtration were compared with reference data and values of filtration percentage errors were calculated. Computed figures confirm the usefulness of the flakes method for airborne laser scanning data filtration. For each test set filtration correctness exceeded 90%. Depending on the structure and type of land cover, the total percentage error of filtration

was in the range from 1.99% to 7.40%. Values of errors obtained using flakes method are similar to the results obtained using other filtering algorithms (Figure 10).
Wrongly classified points on the edge of the elaboration area (filtering error of type I) could be eliminated by adding points from the adjacent scan to the data. During manual check of automatic filtration, points that create artefacts (filtering error of type II) can be quickly detected and eliminated from the set of terrain points. This set is the basis for DTM generation using airborne laser scanning data. Tests confirmed that photogrammetric gaps are not responsible for classification errors on the edges of gaps.
The proper choice of weighting parameters α and β of internal energy (2) allows including structure lines. The disadvantage of flakes filtering method is the necessity of regularization of original data, which extends the computing time. Nevertheless, the outcomes of filtration for the flakes method confirm that it is useful for ALS data filtration.

ACKNOWLEGMENTS

Interpolation procedures have been implemented using 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.

REFERENCES

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

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

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

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

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

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

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

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

Kass M., Witkin A., Terzopoulos D., 1988. Snakes: active contour models. *International Journal of Computer Vision*. 1, pp. 321-331.

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

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, pp. 23–30.

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

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, pp. 227–232.

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, pp. 203–210.

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.* Vol. LIX, pp. 85-101.

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

Vosselman G., 2003. ISPRS test on extracting DEMs from point clouds "A comparison of existing automatic filters", Delft. http://www.itc.nl/isprswgIII-3/filtertest/Reference.zip

Vu T. T., Yamazaki F., Tokunaga M., 2002. Wavelet-based filtering the cloud points derived from airborne laser scanner. Proceeding of the 23rd Asian Conference on Remote Sensing. Kathmandu, 25-29 November. http://www.gisdevelopment.net/aars/acrs/2002/adp/adp008.asp

Wei H., Bartels M., 2006. Unsupervised segmentation using Gabor Wavelets and statistical features in LIDAR data analysis. 18th International Conference on Pattern Recognition. Hong Kong, 20 August.