

# SEGMENTED FILTERING OF LASER SCANNER DSMS

K. Jacobsen, P. Lohmann

University of Hannover, Institute of Photogrammetry and Geoinformation, D-30167 Hannover, Nienburger Str. 1  
(jacobsen, lohmann) @ipi.uni-hannover.de

**KEY WORDS:** Filtering, segmentation of point clouds

## ABSTRACT:

The filtering of a digital situation model (DSM) containing points located on vegetation and buildings to a digital height model (DHM) is depending upon the structure of the objects not belonging to the bare ground. The limitation of an automatic global filtering of a DSM as a compromise between the different object classes is shown. Based on an automatic segmentation for each different class the required filter thresholds could be determined automatically for reaching an optimal filter result.

## 1. INTRODUCTION

Techniques for generating Digital Surface Models (DSM) like image matching and laser scanning have the disadvantage of not representing the wanted bare terrain, but the visible surface including vegetation and buildings. A number of filtering techniques have been developed to remove these "artefacts" in order to obtain the true Digital Elevation Model (DEM) (Lohmann et al., 2000). However the quality of nearly all procedures depends on an appropriate setting or determination of thresholds and control values (Jacobsen et al., 2002, Kraus, 1997, Voelz, 2001). As a result different parameters and threshold values, depending on the type of filtering, have to be used for different structured regions (Land use) within a DSM. This requires a segmentation of a DSM into areas of similar properties to optimize the performance of the filter.

## 2. SEGMENTATION OF A DSM

The segmentation of a DSM originating either from matching or laser scanner data should use classes which are almost general in their definition and therefore transferable from one scene to another. In most cases common Land use classes can be easily interpreted by a skilled human interpreter using optical imagery. Modern laser scanning systems are capable of recording height information together with an additional reflectance image of the terrain.

It has been shown (Baatz et al., 2002) that software tools like *Definiens eCognition* are able to produce satisfactory segmentations on high resolution remote sensing imagery using a hierarchical network of image objects together with a set of definable rules based on a weighting not only of pixel values but also features like shape and topology. This software was used for the current segmentation.

### 2.1 Segmentation of a laser scanning data set

The investigated data set was provided by the company TOPSCAN who carried out the laser survey on behalf of the Emscher Cooperative in an area "Ickern/Waltrop" close to Castrop-Rauxel in North Rhine Westphalia, Germany. It was generated using the OPTECH ALTM 1225 laser scanner

operated in first/last pulse mode, with the possibility to record an intensity image. Figure 1 (left) shows the grey coded last pulse heights together with the segmentation result and the intensity image to the right. The measurement point density is approximately 0.6 points per m<sup>2</sup>. This data set has been investigated for filtering using a segmentation with 8 general classes based on the height image alone and in combination with the intensity image.

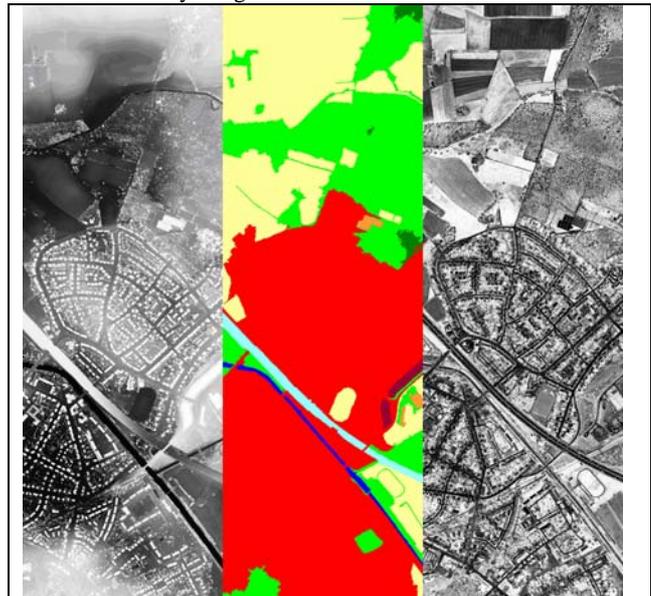


Figure 1: Segmentation result (left grey coded height, middle segmentation result, right intensity image of laser scanner data)

The hierarchical net within *eCognition* was formed by 3 levels and the selectable segmentation parameters were set iteratively to reach the best visual result (Lohmann, 2002).

The different levels have been generated in accordance to human interpretation concepts. As an example within level 1 the segment borders should coincide with the building outlines to allow the classification of single buildings. Therefore the building is the essential element to define the segmentation parameter of level 1. Buildings are visually best discriminated by a brighter colour (height) as compared to their surroundings. So the parameter colour dominates over the parameter shape in

this level. By similar reflections the other segmentation parameters are found.

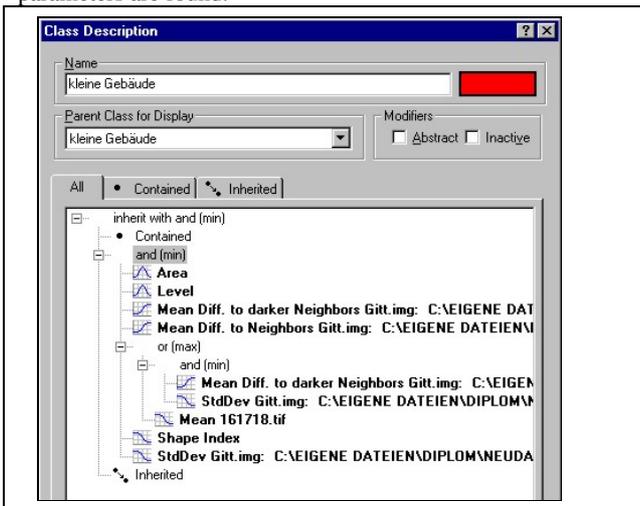


Figure 2: Class description within eCognition

Basic class or feature descriptions as shown in figure 2 and table 1 are used to define objects at each level and to aggregate them on the next higher level up to the desired output class. As for the class of settlement (level 3) it starts with buildings at level 1, aggregated to housing areas, blocks of flats including roads and backyards and row houses in level 2 (see figure 3).

<b>2. SMALL BUILDINGS</b>		→ Base to define class settlement (buildings, housings with backyards, small blocs of houses).
		<b>Feature Description</b>
Area		Segments between 80 Pixel (m <sup>2</sup> ) and 2000 Pixel → defines max. size of a small building.
Mean Difference to darker Neighbours (Height)		Terrain surrounding a building is lower → darker
Mean Difference to Neighbours (Height)		The majority of neighbouring segments should be lower.
Or (max)	and (min)	Mean Diff. to darker Neighbours (Height) Std Dev (Height) If the mean value of the intensity channel is high (typical for vegetated areas), the feature Std. Dev prohibits confusion with forest.
Mean (mean value intensity)		
Shape Index		The smooth geometry of buildings is used (length of border divided by area of building → < 2.6).
Std. Dev. (Standard Deviation of Height)		Texture (First order statistics) within a segment. Relatively small compared to class forest)

Table 1: Object description table (level 1 – “small buildings”)

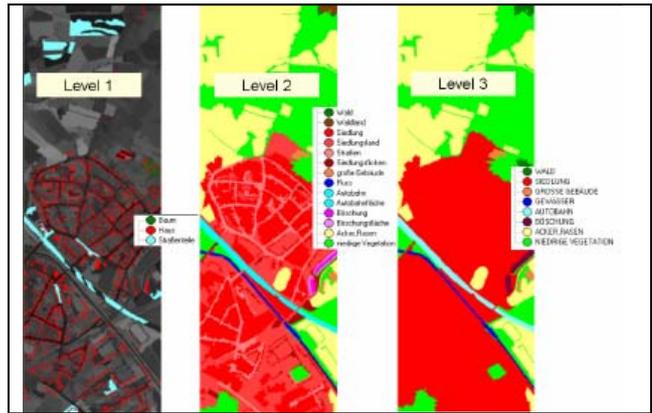


Figure 3: Different levels of aggregation in the segmentation process

The accuracy and reliability of the segmentation / classification was checked against a manual interpretation of the area and is summarized in figure 4.

As it can be seen, the accuracy does not significantly improve by adding the intensity image, but the reliability of the class being produced raises. However in comparison to a manual interpretation it can be observed, that especially small area classes like creeks or ditches are missing. Therefore the selection of parameter / threshold setting according to land use segments in the following was improved by manual editing.

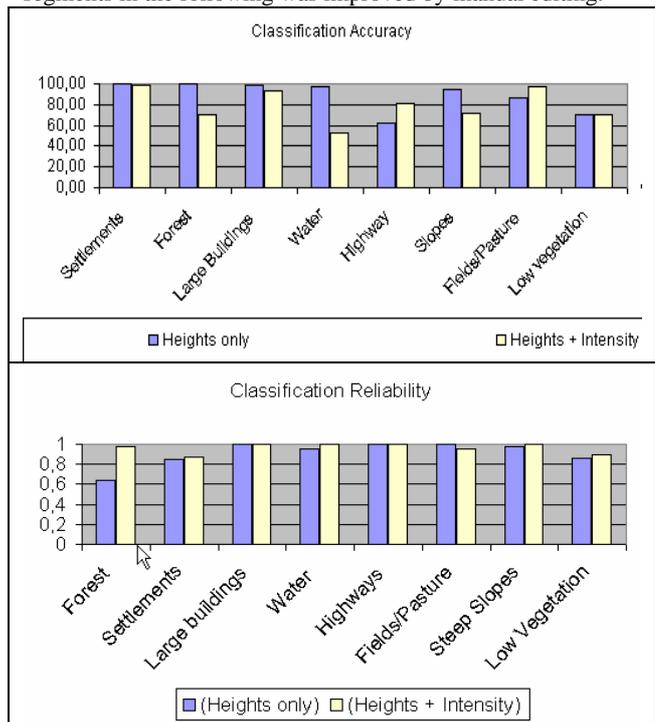


Figure 4: Classification/Segmentation reliability and accuracy

### 3. FILTERING OF THE DSM TO A DEM

In the University of Hannover the program RASCOR has been developed for an automatic filtering of a DSM to a DEM. RASCOR is able to determine the required thresholds for the filtering by an analysis of the data set. A problem is the different thresholds for different terrain types. Without separation of the terrain type only average parameters can be determined by the data analysis leading to some remaining

artefacts of the filtered DSM. The described DSM was at first filtered without respect of the object classes. The highway, two water courses and a large dump have been specified with break lines to avoid a not intended loss of points at the shoulder of slopes.

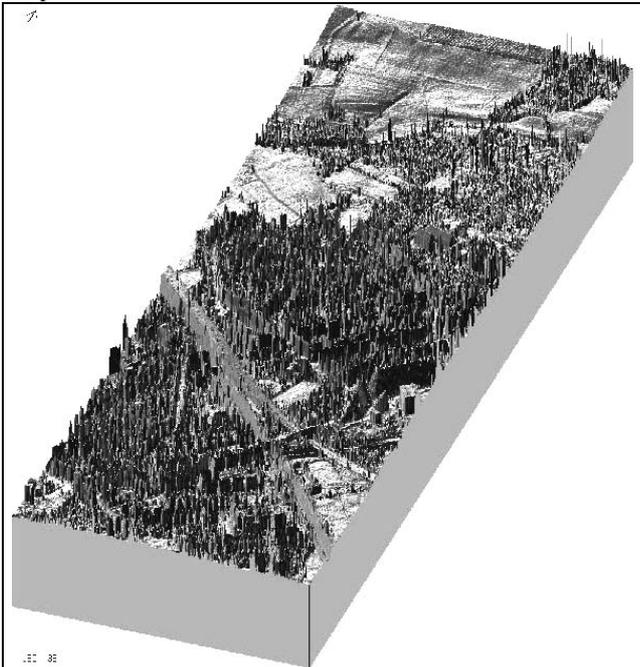


Figure 5: 3-D-view of the used DSM

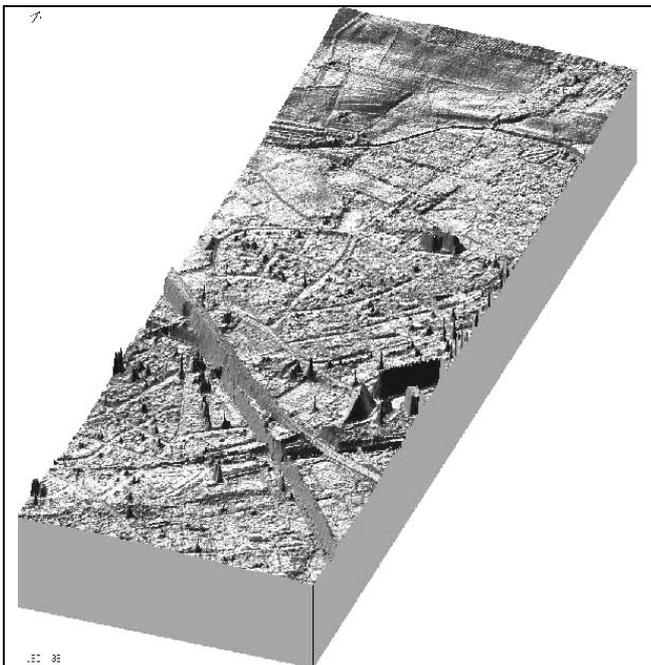


Figure 6: 3-D-view to the DEM generated by global filtering

By the filtering all points not belonging to the bare surface should be removed from the DSM to achieve a DEM, but no other points.

37% of the DSM-points have been removed by the automatic global filtering (see figure 5, 6 and 7). Few trees and buildings at the project border and two large industrial buildings are still included in the DEM. Without use of break lines, a higher number of points at the shoulders of dams and cuts are unintentional excluded.

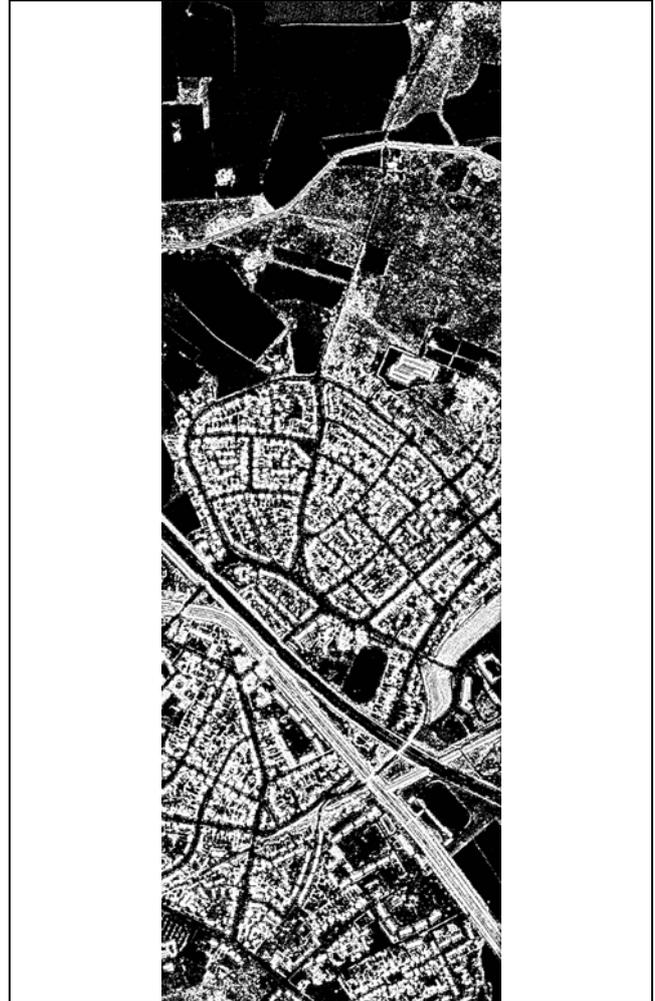


Figure 7: result of the filtering – white = removed or not existing points

### 3.1 Filtering by RASCOR

Program RASCOR is using a sequence of different methods for the filtering of a DSM. Only data sets with raster arrangement are accepted. If the DEM has a random arrangement, the Hannover program DTMCOR has to be used which is mainly based on the more time consuming least squares interpolation (prediction). The operational use showed, from a random arrangement a raster arrangement can be interpolated and this can be analysed by RASCOR with sufficient results even under the condition of not using the original data. But this was not required for the used data set which was directly available in raster form.

RASCOR starts with an analysis of the height distribution itself. Based on the structure of the achieved histogram of height distribution an upper and lower limit of the accepted height can be identified automatically. This methods requires flat areas, it does not work in rolling and mountainous terrain. It is followed by an analysis of the height differences of neighboured points. The accepted height limit of neighboured points is depending upon the slope and the random errors. With this method only small objects and the boundary of larger elements can be eliminated, but it is still very efficient.

Even large buildings can be found by a sudden change of the elevation in a profile to a higher level and a later corresponding change down if no vegetation is located directly beside the

buildings. This method can be used for laser scanning, but it is not optimal for DEMs determined by automatic image matching where the buildings are looking more like hills.

Other larger objects not belonging to the bare ground are identified by a moving local profile analysis; at first shorter and after this longer profiles are used. The required length of the moving local profiles is identified by an analysis of a sequence of shorter up to longer profiles. In flat areas the individual height values are checked against the mean value of the local moving profile, in rolling areas a linear regression is used, in mountainous areas polynomials have to be used. It will be combined with data snooping taken care about a not even point distribution caused by previously eliminated points. All these methods are applied in X- and Y-direction. Elements which have not been removed by this sequence of tests are analysed by moving surfaces which may be plane, inclined or polynomial. The size of the moving surfaces is identified by the program itself by checking the data set with a sequence of cells with different size. The optimal cell size is used for further filtering. As final test a local prediction can be used, but it is usually only finding few points not belonging to the surface after the described sequence of tests.

In the case of the check for height differences of directly neighbored points, the upper point will be eliminated if the tolerance limit will be exceeded. The other methods are using a selectable weight factor for points located below the reference defined by the neighbored points. This will keep points located in a ditch or cutting in the data set. Usually points determined by laser scanning do not have blunders causing a location below the true position, but this may happen in the case of a DSM determined by automatic image matching, justifying a weight factor.

In forest areas at first only the trees are removed by the program, smaller vegetation is remaining, so a second iteration is necessary. A second iteration in other cases may remove also terrain points leading to a more generalised DEM. This may be useful for the generation of contour lines, but it is not optimal for the correct description of the terrain.

### 3.2 Filtering individual segments

If the filtering will be done separately for the individual segment types, caused by the different height structure, different tolerance limits will be defined by the program and also the number of points used for the test in a local profile or area may be different to the global filtering.

#### Settlement and large buildings

Optimal results have been achieved with the tolerance limits by standard analysis and with two iterations. The first iteration is removing the buildings but not the points located on vegetation and other smaller objects. They are removed with the second iteration (figure 8). Large industrial buildings can be identified with the method – jump of height in a profile up and later down. In total 60% of the points are taken out in the settlements.

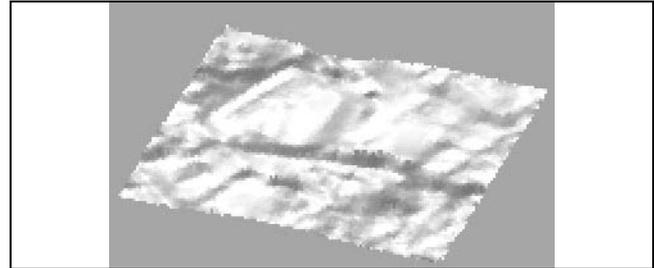
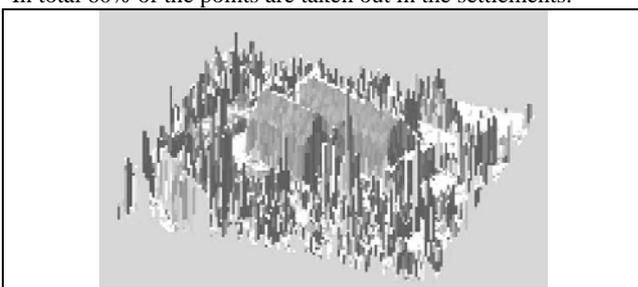


Figure 8: Using 2 iterations to remove large buildings and surrounding vegetation

#### Highway

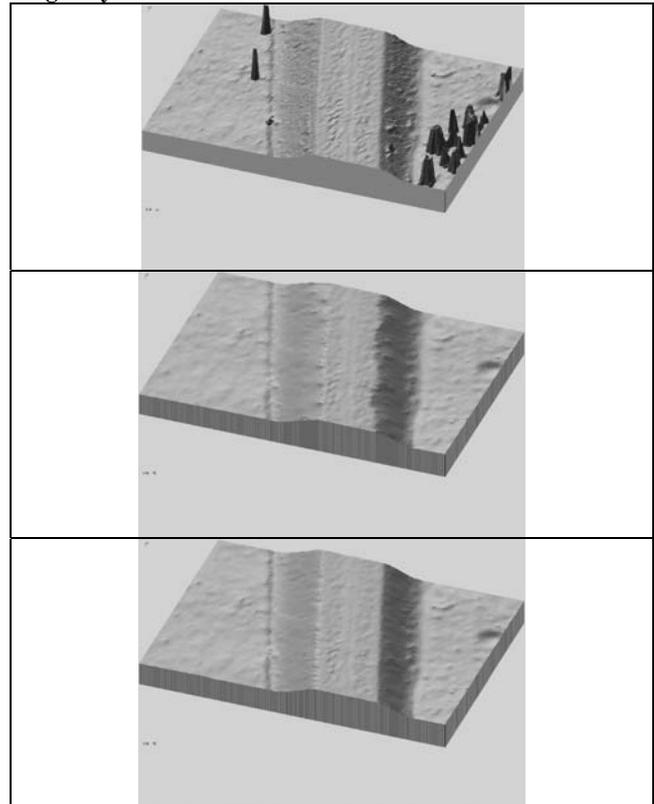


Figure 9: highway: up: original data, center: standard filtering, down: filtering with break lines

Highways are very often located on dams or in cuts, so it corresponds to railways and dams. A general filtering without taking care about the special classes tends to an elimination or reduction of the dams, especially the upper shoulders. If only these special segments are filtered, the filter parameters will take care about the special structure but nevertheless the upper dam shoulders are influenced. These elements have to be handled in a special manner. RASCOR allows the definition of such special areas where only a local filter will be used for the elimination of points located on cars and on bushes on the slope. Another possibility is the use of the segment limits and also the roads as break lines. The program will not take care about points located behind break lines.

As it can be seen in figure 9, a standard filtering of a highway dam is removing the trees and also bushes on the slope, but it takes also points out of the shoulder. A standard filtering removed 21.6% of the points. If break lines are taken into account, only 10.4% of the points were excluded and the dam looks quite more like it should be.

### Slopes

Slopes do show similar problems like highways and dams. They can be handled as special areas, but it is better to use break lines instead of this.

### Watercourses

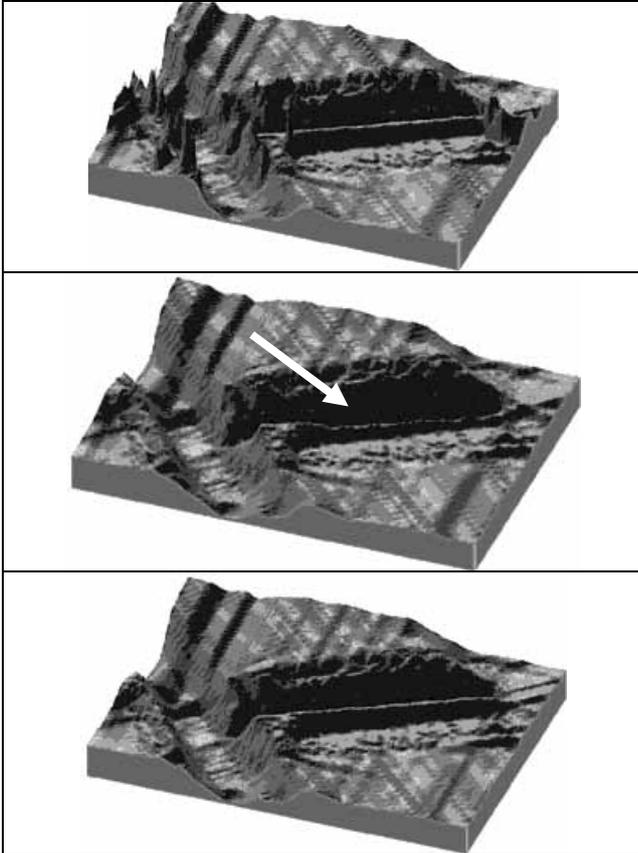


Figure 10: watercourse upper: original DSM center: global filtering lower: result of filtering separately for water bodies

Caused by total reflection of the water surface, usually there is no signal available for water bodies, resulting in missing height information about the water surface. But just beside the water, height information is available which is sufficient for the height definition of the water surface. Typical characteristics of watercourses are slopes and also parallel dams. Cuts are not causing problems because they are not changed by RASCOR, but problems exist with dams beside the watercourse if they have not been separated.

As it can be seen in figure 10, a global filtering of the whole data set is affecting the dam parallel to the water body while the individual filtering is not causing problems. The same effect like the separate handling of the water bodies will be achieved by the introduction of break lines.

### Canal

A canal has a similar structure like the watercourses. The canal existing in the data set has no parallel dams and only few vegetation elements, by this reason the filtering was without any problem. Even the global filtering was able to achieve correct results (figure 11).

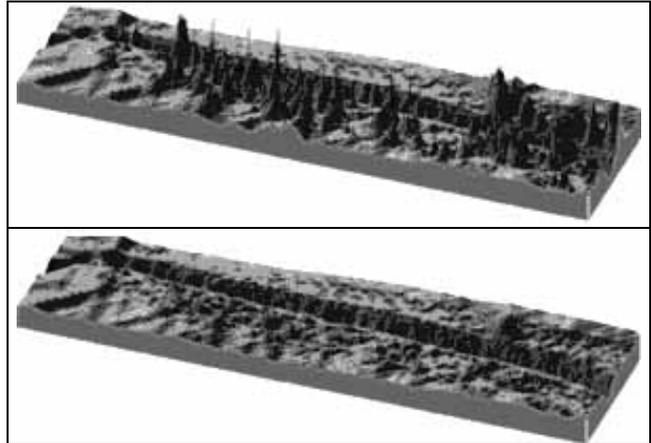


Figure 11: canal: upper: original DSM lower: after global filtering

### Bridges

Bridges do show a sharp change of the height without any slope. Of course, if special segments are used just for the bridges, they will not be changed because points are only located on top of the bridges without more or less variation in height. From the definition it is better to use break lines instead of a special segment. The break lines can include also the neighbored dams, so a handling may be easier.

### Grass

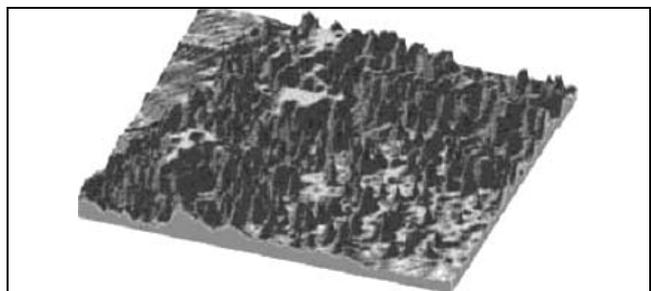
Meadows and willows are simple segments; they do include only few points not belonging to the bare surface like located on bushes and trees, so they can be handled with the global and the segmented filtering without any problems.

### Vegetation

As vegetation areas a mixture of grass and bushes are defined. These segments could be handled in a similar manner like the segment grass. Corresponding to this, the segmented as well as the global filtering yielded to very similar results.

### Forest

A global filtering of the whole area was done only with one iteration because a second iteration affected especially dams and highways. For the forest area a second iteration is recommended. The filter parameters are not so sensitive. With a second iteration (figure 12) also the global filtering was leading to sufficient results in the forest area, but as mentioned, other areas have been influenced.



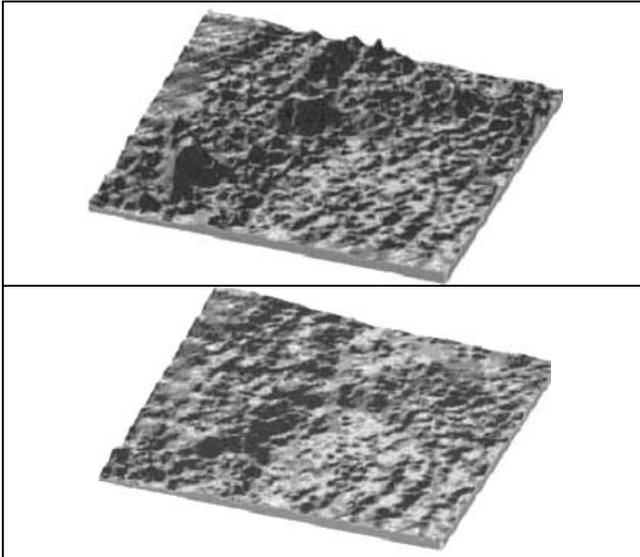


Figure 12: forest: upper: original DSM center: filtered with 1 iteration lower: filtered with 2 iterations

Passini, R., Betzner, D., Jacobsen, K., 2002, Filtering of Digital Elevation Models. ASPRS annual convention, Washington 2002

Schultze, D. 2003. Segmentierte Filterung von Laserscanner-Daten. Diplomarbeit, Universität Hannover

Voelz, H., Mathematische Grauwertmorphologie zur Filterung Digitaler Oberflächenmodelle, Diplomarbeit, Universität Hannover, <http://www.ipi.uni-hannover.de/html/lehre/diplomarbeiten/2001/2001.htm>

#### 4. CONCLUSION

The automatic filtering of a DSM to a DEM without respect of the different structure of the available classes is not leading to optimal results. Most terrain classes do show advantages if they are filtered individually. Especially settlements including large buildings, forest, water courses and highways as well as dams are resulting in quite better results like a global filtering with parameters analysed from the whole data set. But dams as well as highways and water courses should be handled with break lines in order to maintain their specific shape. A major difference between a segmented and a global filtering can be found in the settlement and the forest area. Because of this a grouping of the other segments is possible, which simplifies the whole process.

#### References:

Baatz, M. & M. Mimler, 2002 "Bildobjekt-Primitive als Bausteine - Extraktion von Objekten "of interest" bzw. anthropogenen Objekten basierend auf der expliziten Kanteninformation von Bildobjekt-Primitiven". In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren - Innovative Methoden. Wichmann Verlag, Heidelberg: 179-188.

Jacobsen, K. 2001, New Developments in Digital Elevation Modelling. Geoinformatics, June 2001, pp 18 – 21

Kraus, K., 1997, Eine neue Methode zur Interpolation und Filterung von Daten mit schiefer Fehlerverteilung. Österreichische Zeitschrift für Vermessung & Geoinformation 1/1997, pp 25-30

Lohmann P., Koch A., Schaeffer M., Approaches to the Filtering of Laser Scanner Data, IAPRS (33), B3/1, 534-541, Amsterdam, 2000

Lohmann P., Segmentation and Filtering of Laser Scanner Digital Surface Models, IntArchPhRS, Com.II, Xi'an, Vol.XXXIV Part 2, pp 311-315