Processing of laser scanning data for wooded areas

KARL KRAUS and WOLFGANG RIEGER, Wien

ABSTRACT

Airborne laser scanners have been increasingly used in recent years for the collection of terrain elevation data especially in wooded areas. The technology demands new processing technologies which have reached a stage of good practical application so far for the interpolatin of ground DTMs. Beyond pure ground DTMs, however, laser scanner data have great potential for other applications such as creation of digital building models or vegetation analysis. The paper presents the state of the art of interpolation techniques along with a new technique to separate wooded from non-wooded areas even in the special case of off-terrain objects (buildings, rocks) which would normally be eliminated from the ground model. Finally methods are outlined to extract information on forest stands, such as mean stand height, stand density, and tree species, directly from laser scanner data.

1. INTRODUCTION

The collection of high precision digital terrain models (DTMs) in wooded areas was an unresolved problem till recent years. Laser scanning is the first method that allows to accomplish that task. In the meantime ample experience has been collected, for example at the regional surveying offices in Germany (H. Hahn, M. Knabenschuh, W. Stössel, 1999). The obtainable accuracy of DTM points that were derived from laser scanner data in wooded areas can be given as (K. Kraus, 2000):

$$\sigma_H[\mathrm{cm}] = \pm (18 + 120 \cdot \tan\alpha) \tag{1}$$

Equation (1) is valid for a ground penetration rate of the laser signal of at least 25 % and a good mixture of vegetation and ground points for the whole region. The relatively high dependence of the elevation accuracy on the terrain slope $\tan\alpha$ is due to several reasons:

- In steep areas the topographic surface tends to be less well-defined than in flat areas.
- The ground extent of the scanner flex – which is about 20 cm at a flying height of 1000 m – can be ignored in flat areas but has significant influence in steep terrain.
- The inherent plane errors of the on-line GPS position and INS orientation strongly influence the elevations in steep terrain in contrast to flat terrain.

Photogrammetric data collection methods show a similar accuracy law as laser scanning DTMs (K. Kraus, 2000). The photogrammetric accuracy depends on the flying height. The accuracy in open terrain of a DTM that was plotted from wide angle images at a flying height of 1200 m, resulting in an image scale of 1 : 8000, is comparable to that of a laser DTM in forests.

The elimination of inherent systematic errors allows a significant improvement of the accuracy of the laser DTM particularly in flat terrain. The constant value of ± 18 cm in equation (1) can be reduced down to ± 10 cm (K. Kraus, N. Pfeifer, 1998).

The geomorphological quality of the DTMs that were derived from laser scanner data is not yet satisfactory especially along break lines of the terrain. In the publication (K. Kraus, N. Pfeifer 1998) several suggestions were made to improve the geomorphological quality. Currently the following method is the most effective one: The laser DTM is overlaid over photogrammetric imagery (with softcopy plotters or by image injection on analytical stereoplotters) and edited manually. The method implies the availability of both laser scanner data and aerial images of the region of interest and at least some degree of ground visibility in the aerial photographs. The operator is usually able to detect potential breaklines in the laser DTM and may collect them and eventually improve the DTM with help of the stereoscopic imagery.

2. CLASSIFICATION AND FILTERING OF THE LASER SCANNER DATA

For the processing of laser scanner data a special algorithm has been developed and implemented in the DTM program package SCOP (N. Pfeifer, K. Kraus, A. Köstli, 1999). This algorithm estimates the skewness of the error distribution of the laser scanner data in forests and assigns small weights to those points that show large positive errors during the interpolation with filtering. The process results in a classification of the laser points in terrain and off-terrain (mainly vegetation) points. Following, interpolation is restricted to one or the other set of points.
According to that method the following projects are currently processed at the Institute for Photogrammetry and Remote Sensing at the TU Vienna (I.P.F.).

- Riparian forests of the Danube below Vienna on behalf of the *Österreichische Wasserstraßen-direktion*. The region of interest has an extent of approx. 40 km². The laser scanner flight was undertaken by the company "TopoSys". It is expected to reach an overall accuracy between 10 and 20 cm even within the densely forested areas.
- Vienna Woods in behalf of the *Magistratsabteilung 41* of the municipal authority of Vienna. The total area of about 90 km² was previously surveyed with the laser scanner of the company "TopScan" (K. Kraus, E. Hynst, P. Belada, T. Reiter, 1997). A small part of the area has been captured by a laser scanner of the company "TopoSys" in late winter 1999. The I.P.F. is going to derive a DTM from that data and compare the DTMs from the two different laser data sets.
- A map sheet of the official map of the *Bayerisches Landesvermessungsamt*. The laser scanner data of that area were already processed with morphologic filters by the company "TopScan" (C. Weber, 1999). For comparison reasons these data are currently modelled with the filtering algorithms that have been developed at the I.P.F.

At the Photogrammetric Week results of the processing of these three projects are presented. The main focuses of the current contribution are,

- a method that enhances the I.P.F.´s interpolation with filtering with a skew error distribution function (section 3) in order to automatically separate small solid off-terrain objects, and
- the application of laser scanning to estimate forest stand parameters (section 4).

3. ENHANCEMENT OF THE INTERPOLATION WITH FILTERING

Although the implemented algorithm allows to automatically distinguish between areas that need to be processed with a skew error distribution, i. e. forested areas, and regions with a symmetric error distribution, i. e. non-forested areas, there are some critical constellations that may cause problems. N. Pfeifer, K. Kraus and A. Köstli (1999) processed a region in the *Sächsische Schweiz* near Dresden which is characterized by many rock needles positioned within the forested areas. The size of these needles often nearly does not exceed the size of single trees, thus the rock points are eliminated along with the vegetation points.
The area was processed at the TU Dresden and at the I.P.F. with SCOP. A contour plot of a surface created from all laser points was used to interactively delimit the rocky from the forested areas. This is possible since the shape of the contours differs between rock needles and trees. Wooded areas were processed with the skew error distribution function, rocks with the symmetric one. The figures in the publication (N. Pfeifer, K. Kraus, A. Köstli, 1999) show satisfactory results.

Following, a method is proposed that automatically allows to distinguish between forested and non-forested areas. Two laser data sets are needed, first reflected pulse and last reflected pulse. The algorithm was implemented on grid DTMs that were derived with interpolation from the two data sets with 1 m grid width. Better results will certainly be obtained with the original laser data.
If the terrain surface of the region is known, the laser data can be transformed into "object elevations" by simply subtracting the terrain elevation from the laser data. This is done for both the first and the last pulse data set. Thus it is possible to distinguish between different ground coverages, since the two pulses exhibit different elevation distributions, penetration rates, and roughness parameters both from one another and between different ground coverages (see section 4). Another very important data set is the difference between first and last pulses of the same area. This difference model shows how far the laser pulse may penetrate into the object on a point-to-point base rather than for a whole area. Several problems are still present:

- The terrain surface is not known accurately; rather, its derivation is the main aim of the process. Thus iteration is necessary: First, the DTM is derived according to the methods described above; from that model primary estimates of the object elevations are derived. Following, the interpolation is repeated by subsequent introduction of the iteratively refined knowledge about the ground coverage, which furthermore allows to improve the parameters.
- A larger area is necessary to estimate statistical parameters such as elevation distribution or roughness. Therefore segmentation is needed which is, however, not available at that stage; in contrary, it is part of the results of the process. The segmentation is based on the measures that need to be estimated from the segments, thus again iteration is mandatory. The difference between first and last pulse is of great help in that case, since rocks and buildings exhibit differences close to zero in contrast to vegetation.
- The size of the objects is of crucial importance. Trees and buildings are often of similar size whileas terrain features tend to be larger. Unfortunately, in areas like *Sächsische Schweiz* rock needles may be of similar size as trees.

Figure 1 shows the process of automatic separation of the wooded areas in contrast to the non-forested areas: Figures 1b and 1c show the difference data sets "summer first" minus "summer last" respectively "summer last" minus "ground model". All areas that exhibit differences larger than a threshold of 1 m in the first data set are assumed to be vegetation points, areas with values lower than 1 m in the second data set are assumed to be impenetrable for the laser scanner, i. e. buildings or rocks or similar. The threshold value is critical; it should further be tuned to the local slope in order to obtain more reliable results. Figure 1d shows the resulting off-terrain object mask which has additionally been processed by a 3 x 3 despeckle filter.
The first-last difference tends to exhibit "dams" along the edges of buildings or, more generally, along "walls" in the terrain (Figure 1c); this is due to the ground size of the laser spots (that may only partly hit roofs along the edges) and because most laser beams are sent in oblique direction rather than vertical and may arise from different flying stripes. Discretization into a grid further deteriorates the quality. Ideally both the first and last reflected signal for each laser pulse should be reflected in order to minimize the effective of the ground spot extent; this was not the case with the TopoSys laser scanner in the Rosalia example. Rather two flights were undertaken which additionally accounts for a portion of the dams.
Figure 1f shows a hill shading visualization of the ground model with the buildings extracted from the data set "summer first". All models are regular grid models with 1 m grid width.
The mask (Figure 1d) can certainly still be improved, however, the process is fully automatic and can be the base for a more sophisticated break line detection task.

The usage of aerial or satellite imagery may be of great help to better distinct between vegetation and non-vegetation. Another, most promising approach, is to collect information about the intensity of the reflected radiance with the laser scanner, as can be done with the ScaLARS (**Sca**nning **L**aser **A**ltitude and **R**eflectance **S**ensor) that was recently introduced as a prototype (A. Wehr, C. Hug, 1999). The intensity of the reflected radiance in near infrared (810 nm) differs widely between vegetation and non-vegetation as is well-known, thus distinction is easily possible.

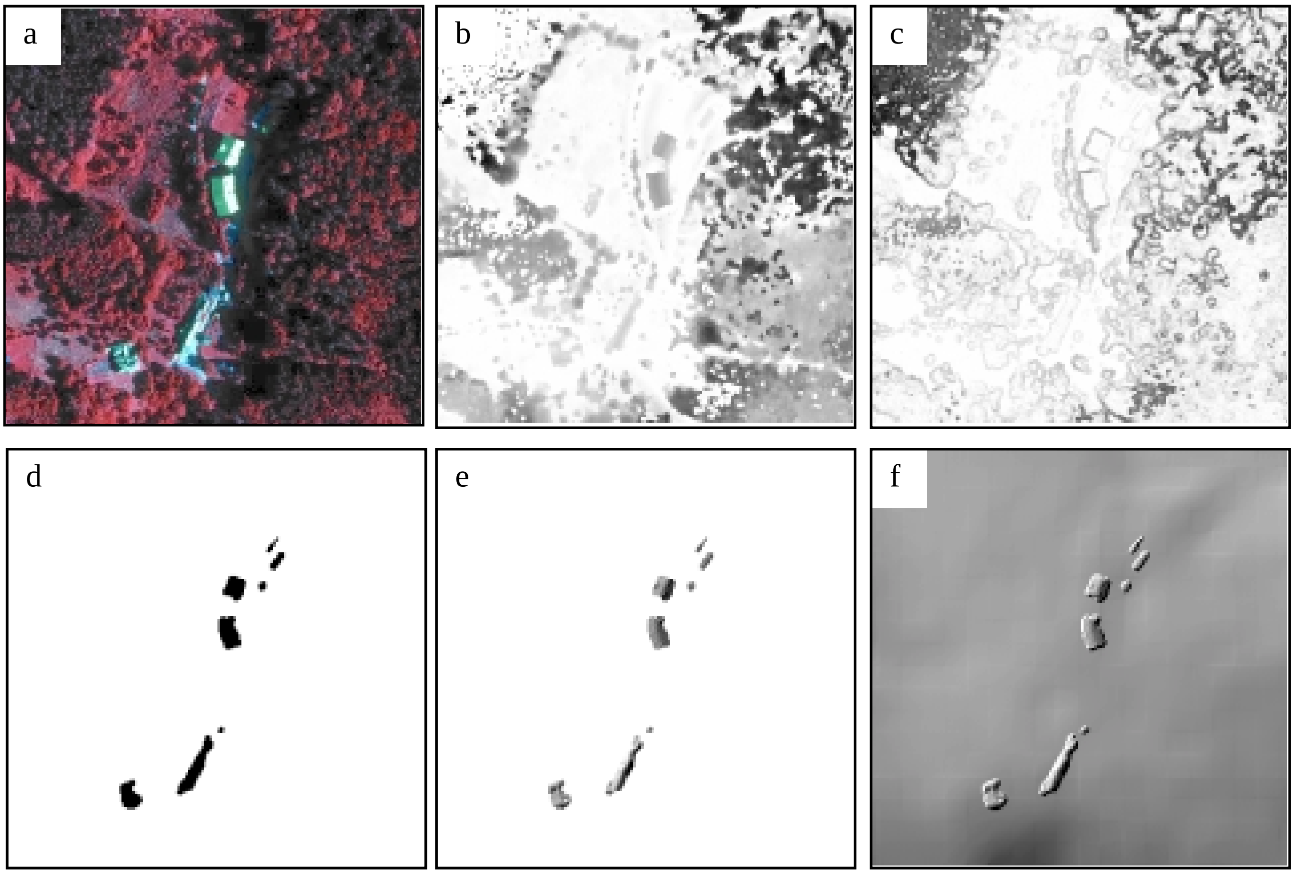

Figure 1: Separation of vegetation (trees) and solid off-terrain objects (buildings): (a) Ortho image of the area; (b) "Summer last" minus ground model, (c) "Summer first" minus "summer last"; (d) mask of buildings derived from (b) and (c) where b is higher than 1 m and c is lower than 1 m, a 3 x 3 despeckle filter was used afterwards; (e) extracted building points; (f) hill shading of DTM (ground model) with buildings implanted. Coding: Higher elevations are darker (b, c, e).

4. APPLICATION OF LASER SCANNING IN FORESTRY

The Agricultural University of Vienna operates a research forest south of Vienna. A part of that area was captured with a laser scanner by the company "TopoSys" at August 19, 1998 in two flights (recording the first respectively last reflected signals for each laser pulse, resulting in data sets "summer first" respectively "summer last") and at March 26, 1999 (leafless season) with the last signal recorded, resulting in a data set "winter last". The last flight was used to create a terrain surface model (DTM) by means of the interpolation and filtering with a skew error distribution function. Figure 2 shows contours of a part of the area (300 x 200 m²) with a contour interval of 2 m, interpolated from the "winter last" data set with the filtering with a skew error distribution function. During the processing a classification in terrain points and off-terrain points was obtained.

Following, the relative (against the terrain) elevations Δh were calculated for all laser dots of the flight "summer first" by simply subtracting the corresponding (interpolated) terrain elevation from each laser point. These relative elevations Δh were used to classify the laser dots:

- All laser points with $\Delta h < 2$ m were eliminated, i. e. all points of the flight "summer first" at the terrain surface as well as those in meadows, bushes, and reforestation areas.
- All laser points with $\Delta h > 40$ m were eliminated, because these points cannot be tree points in that region (amazingly relatively many points were eliminated by this criterion. Presumably most of these dots result from birds).

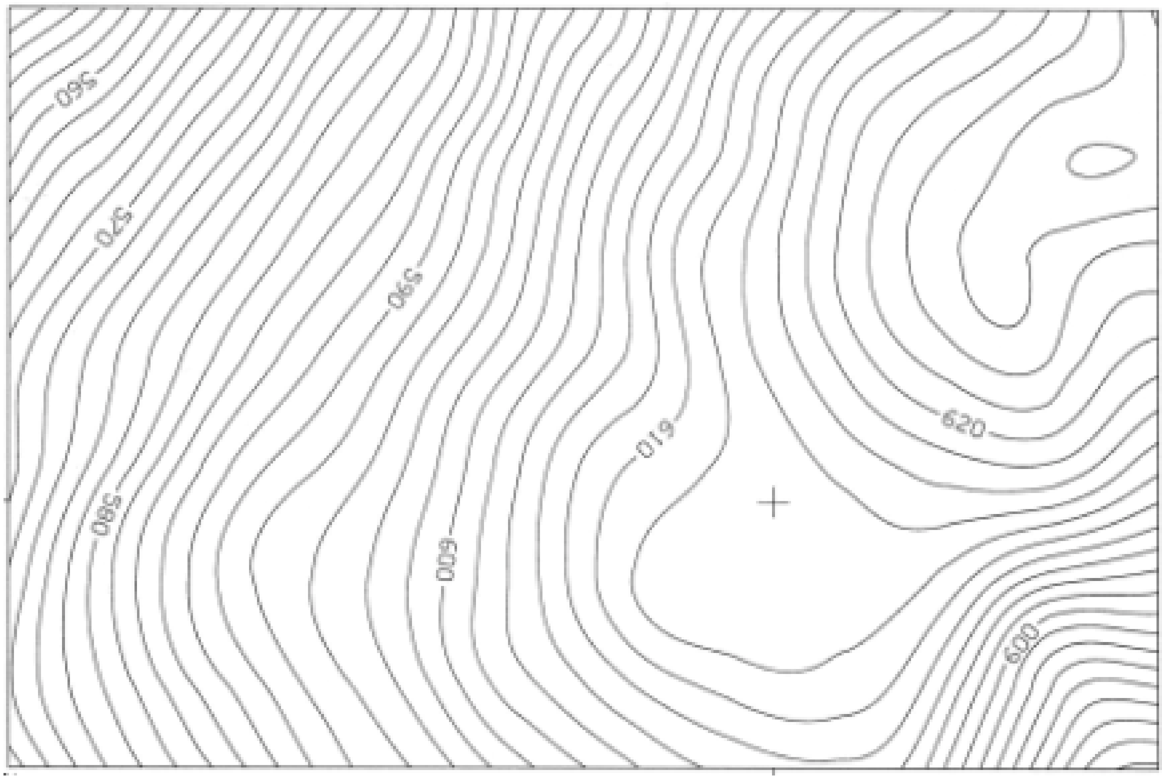

Figure 2: Terrain surface obtained from "winter last" with interpolation and filtering with a skew error distribution function.

The points classified as tree points in "summer first" were joined with the terrain points of "winter last". The result is shown as a contour plot in Figure 3 with a contour interval of 5 m. The forested areas and even single trees can easily be recognized along with the different tree heights. In the perspective view the impression is further increased (Figure 5), especially in contrast to the plain ground model (Figure 4). The grid width of the "canopy model" is 5 m, thus the surface does not represent the real canopy surface or single trees within the forest stands, but rather a kind of a mean canopy surface with the effective roughness reflected quite well.

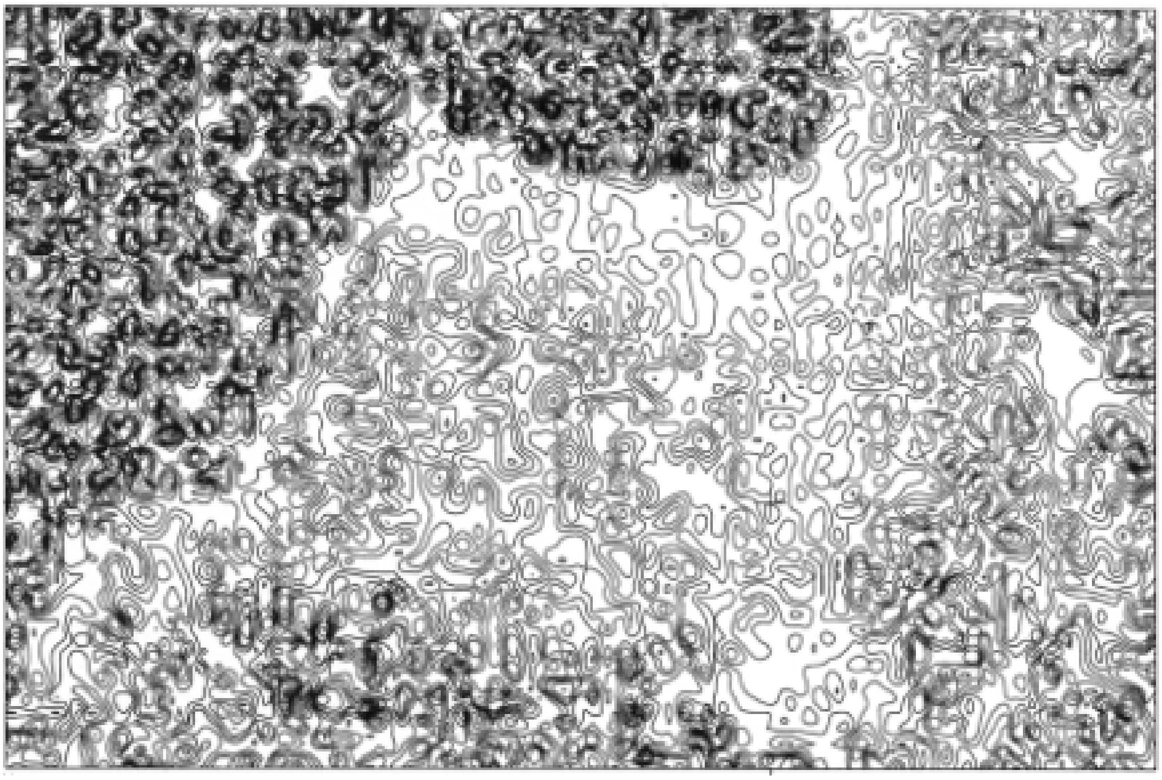

Figure 3: Terrain points of "winter last" joined with canopy points of "summer first".

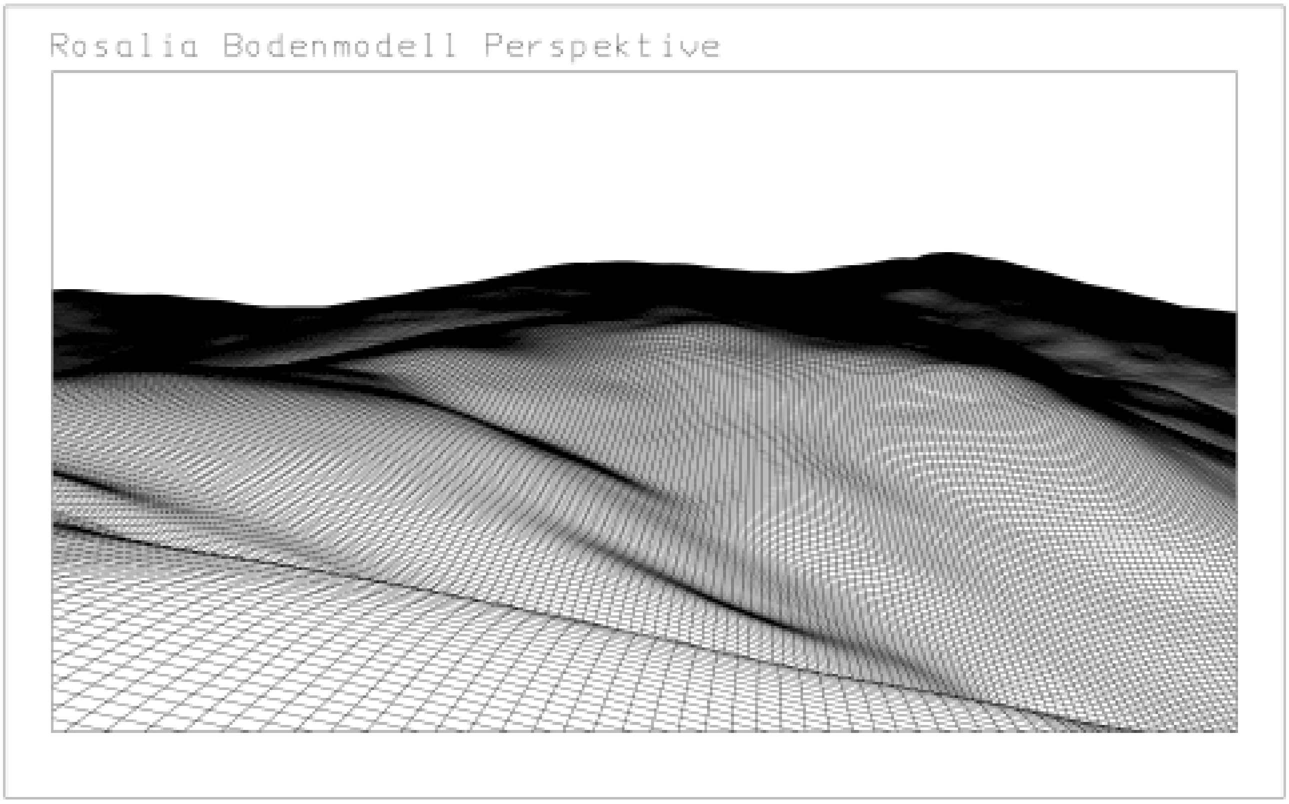

Figure 4: Perspective view of ground model.

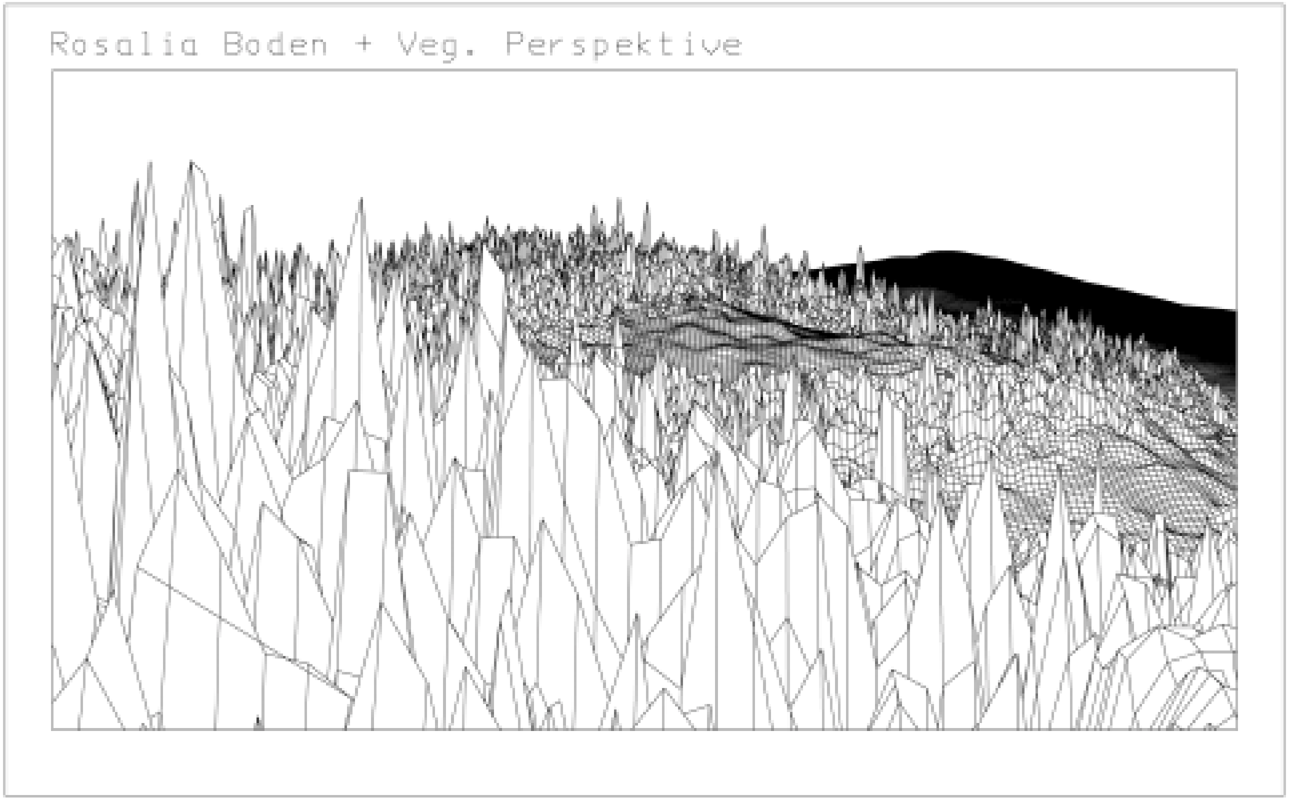

Figure 5: Perspective view of ground model with vegetation implied.

It is near at hand to interpolate a canopy surface simply from the mentioned tree points as collected with the first reflected pulse (the question of the definition of a canopy surface is not discussed here). Two methods are obvious:

- Either the interpolation with the skew error distribution function is used, yet with low weights for points with large negative errors (i. e. points below the canopy).
- Or all elevations are multiplied by –1, and the interpolation with the skew error distribution function is used the same way as with the derivation of the terrain surface.

Figure 6 shows a "canopy surface" that was calculated by the second method. If the terrain surface of Figure 1 is subtracted from that canopy surface, a map of the elevation distribution of the trees can be obtained, which is shown in Figure 7 with a contour interval of 1 m. The trees in the upper left corner exhibit elevations of approx. 30 m, those in the lower right corner only approx. 18 m.

Further parameters that are necessary for forestry applications can be derived from the mentioned laser scanner data. These are in a first step geometric parameters of forest stands, but also information about single trees in a later stage. Forest stand parameters that can be obtained from the (geometric) laser data are mean stand elevation, tree species (in classes), elevation layers, and leaf area indices. Certainly, the geometric conditions for some of these parameters are correlated with one another, thus additional information may be necessary beyond the pure laser elevations, particularly in stands with mixed tree species.

Crucial for the derivation of most parameters is an analysis of the elevation distribution of the laser dots within the forest stand. In the case of the research forest of the Agricultural University Vienna all three laser data sets (winter last, summer first and last) were freed from the terrain elevation by subtraction of the ground model, resulting in "ground coverage models" with 1 m grid width. An additional model is obtained as difference of "summer first" minus "summer last". From these models histograms were derived for several stand types. Figure 8 shows the relative elevation plotted against the relative point frequency (class width 20 cm) and the corresponding distribution functions for several stand types (tree species and age classes).

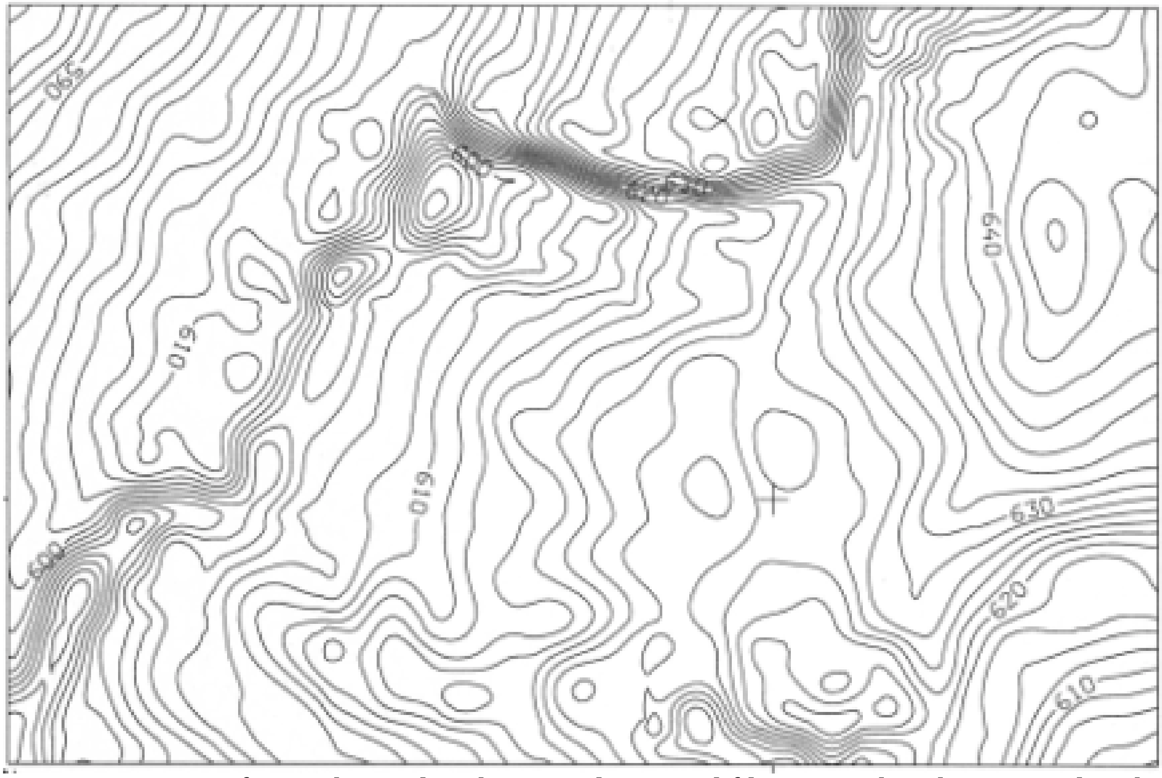

Figure 6: "Canopy surface", obtained with interpolation and filtering with a skew error distribution function on the inverse model (elevations multiplied with –1 before and back after interpolation).

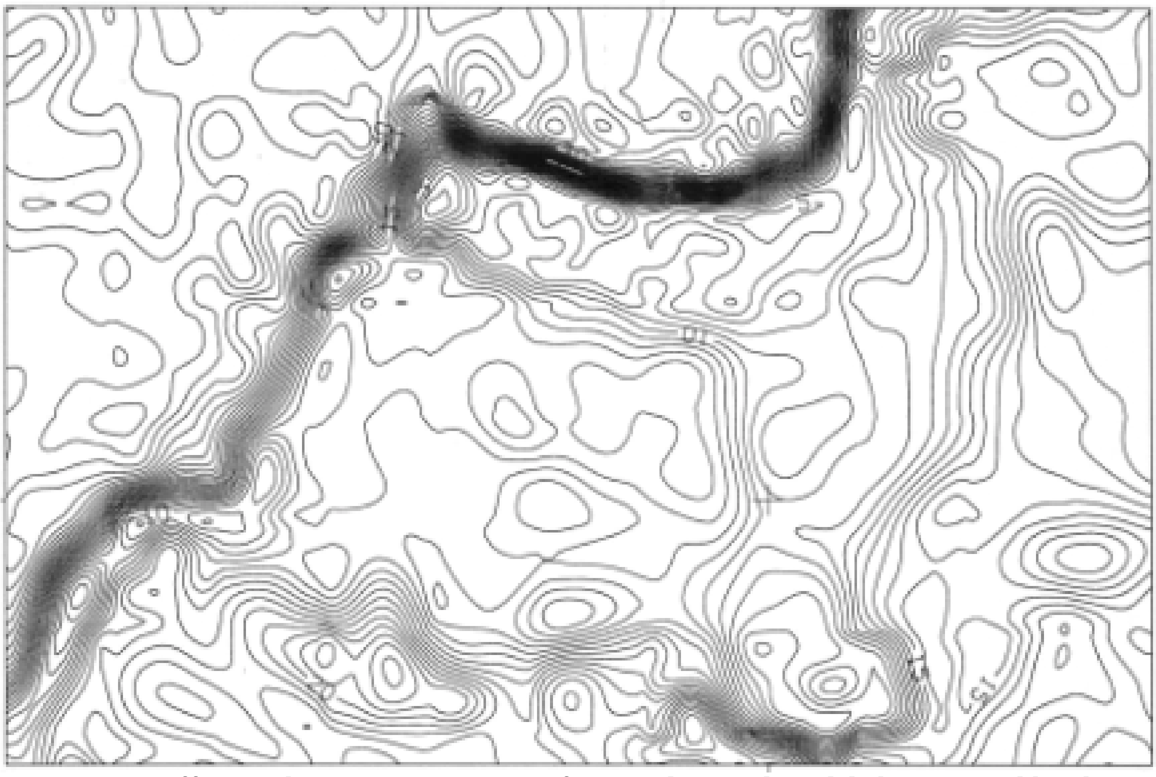

Figure 7: Difference between "canopy surface" and ground model shows stand heights.

Stand heights obtained from the difference model need to be calibrated dependent on tree species and age. Nevertheless the method is practicable and simple and gives a good impression of different tree heights, e.g. the change of tree heights from ridge to valley areas.

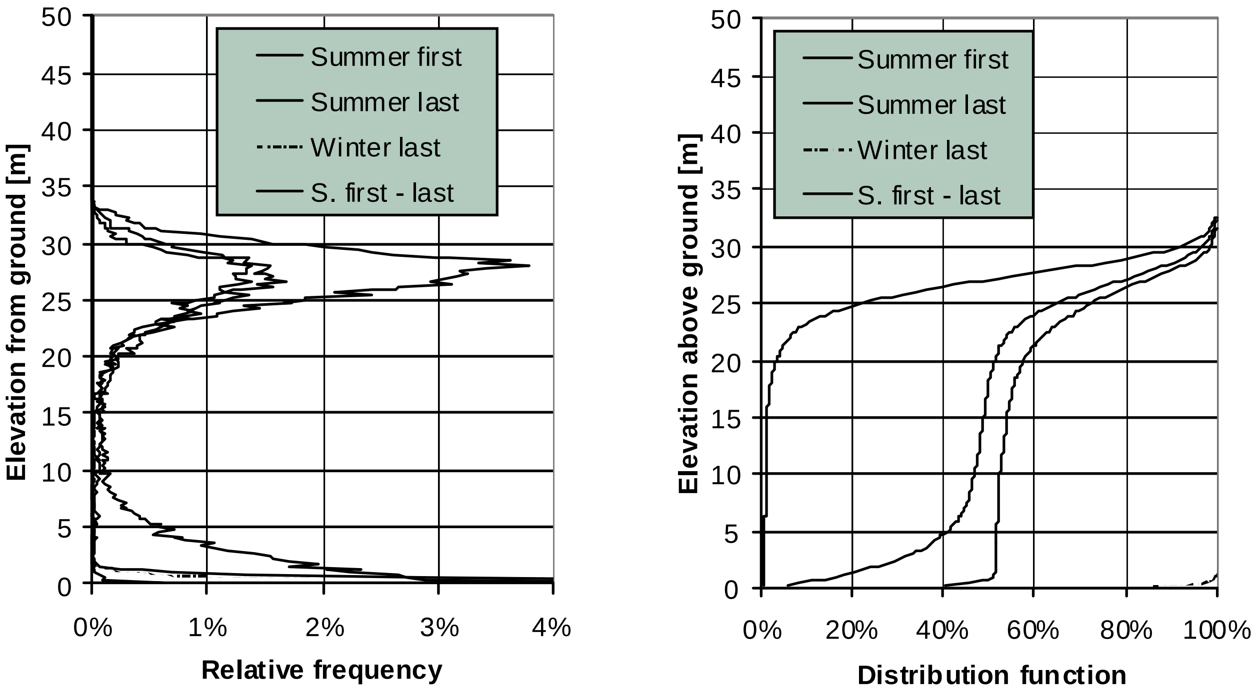

Figure 8a: Histogram and distribution function of elevation distribution of a mature beech stand.

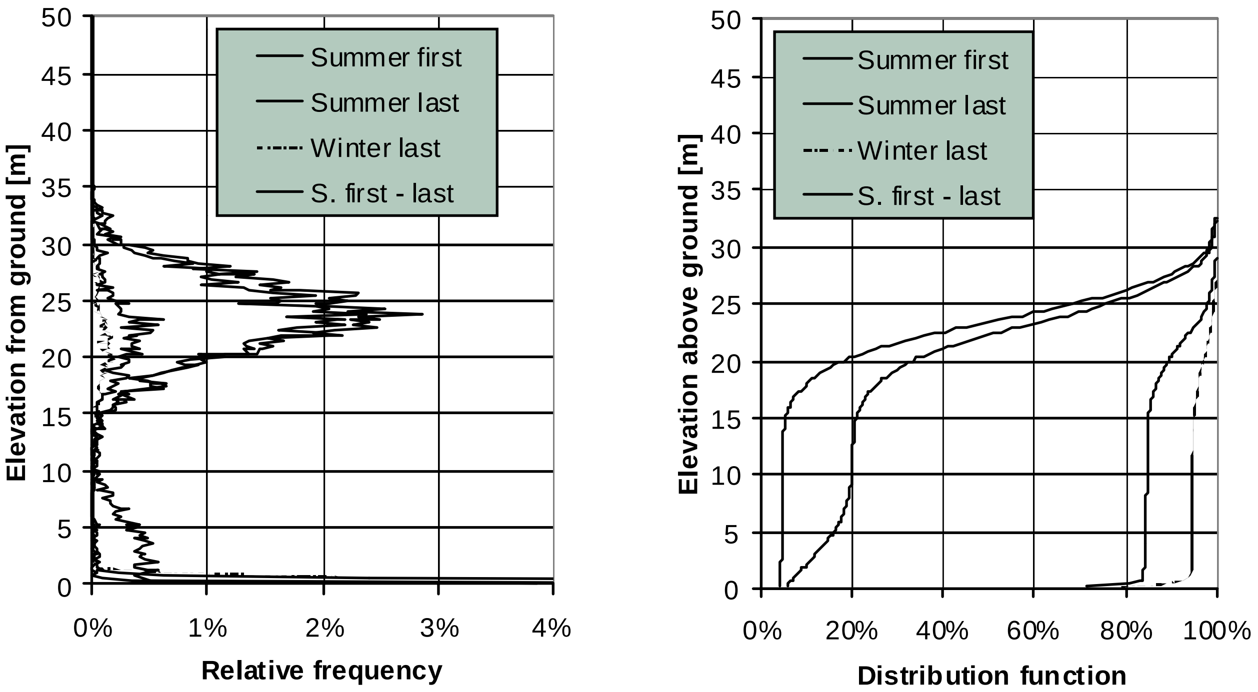

Figure 8b: Histogram and distribution function of elevation distribution of a mature spruce stand.

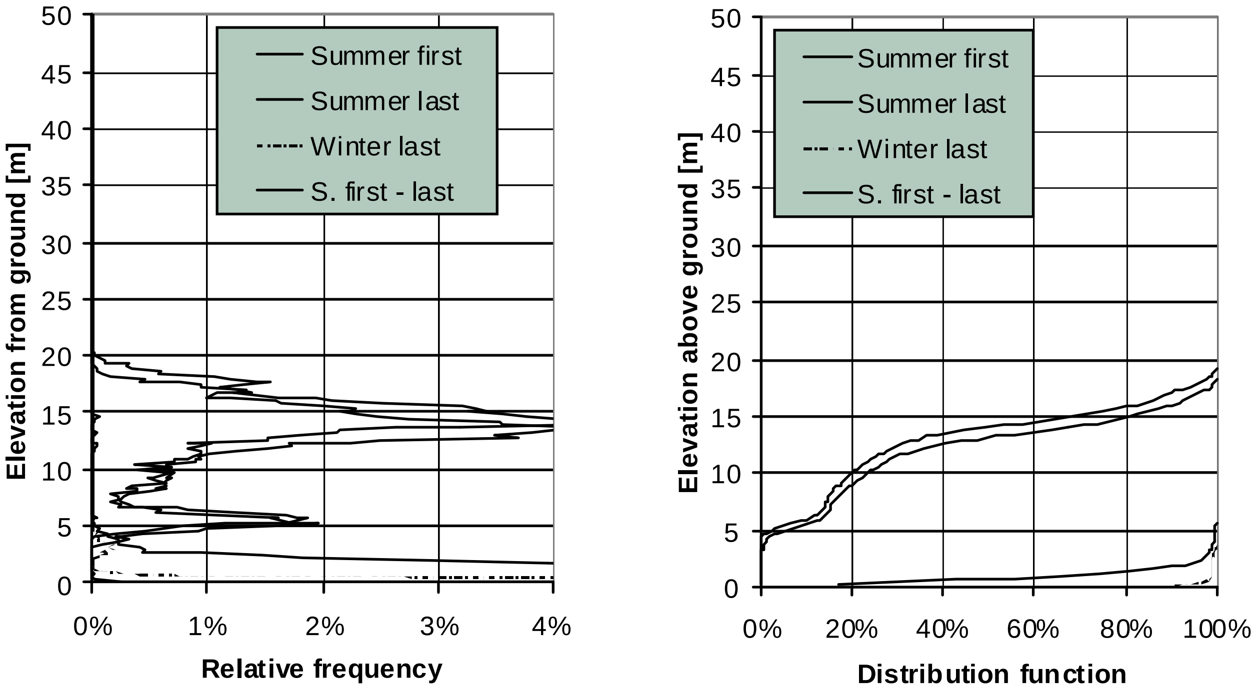

Figure 8c: Histogram and distribution function of elevation distribution of a young beech stand.

Stand density, tree species, age, and leaf area indices are correlated with the penetration rate as given by the integral distribution function (Figure 8). Separation of these parameters is, however, not easy. The values need to be calibrated with ground data. A combination with false color infrared imagery seems to be necessary at that stage in order to free the ground penetration rates from influences of tree species. Texture measures such as local variance of elevation, applied on the different derived data sets (especially the difference data sets) along with the elevation and difference data sets themselves, can be used as input for a multispectral classification. Yet, this field of research is still at the beginning, giving hope for a wide range of new applications possible in near future.

5. CONCLUSIONS

Airborne laser scanners have great potential not only to collect terrain elevation data in wooded areas, but they also yield a lot of information relevant to many disciplines beside photogrammetry. The distinction of buildings or other solid off-terrain features from vegetation is under development and still needs improvement, yet there are promising results with an iterative process that combines first and last pulse data sets and derived ground models. Laser scanner data bear a lot of information about vegetation which can be derived by statistical approaches. Histograms and elevation distribution functions can be used to describe different tree species and age classes as well as density. Correlation with ground truth data is necessary to calibrate these data.

6. ACKNOWLEDGEMENTS

The DTMs and the contour plots were calculated and produced by Thomas Reiter; he also contributed valuable ideas to the processing methods.

Those parts of the publication that are related with the interpolation with filtering have been funded through the project P11336-ÖMA of the Austrian Science Foundation (FWF), the application of laser scanner data in forestry have been funded by the FWF project P12812-INF.

7. REFERENCES

Hahn, H., Knabenschuh, M. and W. Stössel (1999): Erfahrungsbericht zum Laser-Scanning-Verfahren. GIS, Jahrgang 12, Heft 5, pp. 38-42.

Kraus, K., Hynst, E., Belada, P. und T. Reiter (1997): Topographische Daten in bewaldeten Gebieten - ein Pilotprojekt mit Laser-Scanner-Daten. Österreichische Zeitschrift für Vermessung und Geoinformation, Jahrgang 85, pp. 174-181.

Kraus, K. and N. Pfeifer (1998): Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry & Remote Sensing, Vol. 53, pp. 193-203.

Kraus, K. (2000): Photogrammetrie. Band 3: Topographische Informationssyteme. Dümmler Verlag, in Druck.

Pfeifer, N., K. Kraus and A. Köstli (1999): Restitution of airborne Laser-Scanner-Data in wooded areas. GIS, Vol. 12, Copy 5, pp. 18-21.

Wehr, A. and C. Hug (1999): Topographische Geländeaufnahmen mit ScaLARS. GIS, Jahrgang 12, Heft 5, pp. 6-11.

Wever, C. (1999): Laserscannermessungen - ein Verfahren setzt sich durch. GIS, Jahrgang 12, Heft 2, pp. 12-17.