ADJUSTMENT AND FILTERING OF RAW LASER ALTIMETRY DATA

George VOSSELMAN¹ and Hans-Gerd MAAS²

¹Department of Geodesy, Faculty of Civil Engineering and Geosciences
Delft University of Technology, The Netherlands
  g.vosselman@geo.tudelft.nl

²Institute of Photogrammetry and Remote Sensing
Dresden University of Technology, Germany
  hmaas@rcs1.urz.tu-dresden.de

KEY WORDS: Laser altimetry, DEM, adjustment, filtering

ABSTRACT

Although laser altimetry has been used for the production of digital elevation models in different countries for several years, the completely automatic derivation of a DEM from the raw laser altimetry measurements is not yet mature. The two major problems are the detection of and correction for systematic errors in the laserscanner data and the separation of ground points from points resulting from reflections on buildings, vegetation or other objects above ground. This paper discusses strategies for dealing with these two steps in the production of a DEM from raw laser altimetry data. Results are shown of experiments with different data sets.

1 INTRODUCTION

In the last few years airborne laser altimetry has become the prime method for the acquisition of digital elevation models in several countries. The reduction of costs for DEM production and the increase of reliability, precision and completeness will have played a major role in preferring laser altimetry as the acquisition method above analytical or digital photogrammetry. Whereas laser altimetry also has some other advantages over photogrammetry (e.g. the penetration in forested areas), the evaluation of DEM’s derived from laser altimetry data shows that the production process of these DEM’s is not yet mature [Huisings and Gomes Pereira, 1998]. Most errors in these DEM’s can be attributed to two problems: (1) the detection and elimination of systematic errors in the acquired GPS, INS, laser ranging, and calibration data and (2) the elimination of points caused by reflections of laser pulses on vegetation and buildings. Both subjects will be addressed in this paper.

Kilian et al. [1996] noted the systematic errors in the laser data and introduced the concept of strip adjustment in analogy to the independent model adjustment in photogrammetry. The independent model adjustment transforms points from a model coordinate system to the terrain coordinate system by a 3D similarity transform. In a strip adjustment, the transformation from a strip coordinate system to the terrain coordinate system also needs to compensate for systematic deformations of a strip. Different geometric models have been proposed for these transformation. Kilian et al [1996] used twelve offset and drift parameters for position and orientation of a strip. Crombghs et al. [2000] described the use of three parameter transformation which compensates for an offset in height and tilts in flight direction and perpendicular to the flight direction. In both cases tie points were created by measuring corresponding points in overlapping strips. In the next section we discuss the need to explicitly model the error sources in the transformation from a strip coordinate system to the terrain coordinate system. Furthermore, we show that tie points need to be selected with care, since matching algorithms may very well produce biased estimates. Preliminary results of a strip adjustment will be presented.

The elimination of points caused by reflections of laser pulses on vegetation and buildings is usually referred to as filtering. Several filters have been proposed based on auto-regressive processes [Lindenberger, 1993], mathematical morphology [Kilian et al., 1996], least squares interpolation [Pfeifer et al., 1998] and slope thresholding in a TIN [Axelsson, 2000, Vosselman, 2000]. In section three, the performance of slope based filtering is discussed. Most filter algorithms are known to fail in case of larger buildings. A heuristic to eliminate these buildings is introduced. Results are shown for parts of the OEEPE test data sets of Vaihingen and Stuttgart.
All algorithms described in this paper operate on the original irregularly distributed laser points. Most algorithms will be faster and much easier to implement when using laser data in a grid. Producing such a grid, however, requires interpolation. Many interpolated points will be mixtures of ground points and points on vegetation or buildings. This complicates the filtering of the data and may also lead to biases in the estimation of tie points for strip adjustment. We therefore prefer to work with the original irregularly distributed laser data points. TIN’s are used to establish the neighbourhood relationships between the points.

Both process steps, strip adjustment and filtering, are required to create a seamless digital elevation model. They may, however, also interfere with each other. In section four suggestions are made on how to combine strip adjustment and filtering in the production processes of digital elevation models.

2 STRIP ADJUSTMENT

Airborne laser scanner data of larger areas is usually acquired in a strip wise manner. In typical applications with the goal of digital terrain model generation the strip width is in the order of 500 to 800 meter. Modern systems with high pulse rates and large range capabilities allow for the acquisition of strips with widths in the order of 1000 meter or more. Larger blocks are flown strip wise, with an overlap of 50 to 100 meter between neighbouring strips. Due to systematic and random errors of both the laser scanner system and the GPS/INS unit, errors will occur in the ground point data. These errors become obvious from ground control points and as discrepancies in the overlap region between neighbouring strips.

Depending on their source, these errors can be distinguished into strip wise errors and local errors. Strip wise errors are caused by imperfections of the GPS/INS system, such as GPS cycle slips, by misalignments between GPS/INS system and laser scanner measurement head or by INS drift effects. They affect a whole strip or multiple strips. Local errors include the noise of the laser distance measurement unit, affecting single points, as well as the GPS noise, affecting several neighbouring scan lines. Strip wise errors have a systematic effect and become obvious from ground control point differences or discrepancies in the overlap region between neighbouring or crossing strips. Local errors of the laser distance measurement unit may become obvious from the analysis of neighbouring points on an object surface with known geometric properties, but will often remain undetected. Local errors caused by noise of the GPS coordinate determination may become obvious as local discrepancies in strip overlap regions, if tie points can be measured reliably with high density.

Basically, both strip wise and local effects have an influence on all three ground point coordinates. With GPS noise often being the primary error source, the effect in height direction will usually be larger than the effect in planimetry. In traditional DEM applications of laser altimetry with average point distances in the order of 3-4 meter over terrain of limited steepness, only the effect in height direction will be relevant. In modern applications such as 3-D building reconstruction based on laser scanner data with a density of several points per square meter [e.g. Maas and Vosselman, 1999], both effects in height and planimetry become equally relevant.

As the strip wise errors do often show a clearly systematic behaviour, they can be modelled and corrected for in a laser scanner strip adjustment procedure. In [Crombaghs et al., 2000], systematic errors in the height component have been shown, and a model for strip adjustment limited to the height coordinates, has been discussed. In the following, we will show a model for strip adjustment including height and planimetry components.

The strip adjustment procedure includes the following components:

- A mathematical model for the deformation of strips.
- A stochastic model (least squares adjustment, Gauss-Markov model).
- Ground control points.
- A method for measurement of tie points, defined as discrepancies detected in overlap regions between laser scanner data strips.

Due to the effort of providing reference values of ground control points, it is intended to reduce their necessity to a minimum. In an ideal case, four ground control points in the corners of a large region acquired by laser scanning should be sufficient.

In the following, the mathematical model used for strip adjustment in planimetry and height is formulated, and least squares matching applied to laser scanner data points in a TIN structure is discussed as a tool for the provision of tie point measurements.
2.1 Mathematical model

While a simple linear mathematical model for laser scanner strip adjustment restricted to the height component can be limited to three parameters [Crombاغs et al., 2000], the inclusion of planimetric effects requires significantly more parameters. [Kilian et al., 1996] used a twelve parameter model, including three offsets and three rotations of a strip, plus linear drifts of these parameters over time. Due to strong correlations between strip rotations and offset drifts, the number of parameters may be reduced to nine (eq.1). The errors are modelled in a strip coordinate system \((x, y, z)\) which has its origin in the centre of the strip and an \(x\)-axis which is approximately aligned to the flight direction.

\[
\begin{pmatrix}
X \\
Y \\
Z
\end{pmatrix} = R_{\text{strip_to_ref}} \cdot \left( R_e + xR_{et} \right) \begin{pmatrix}
x \\
y \\
z
\end{pmatrix} + \begin{pmatrix}
e_x \\
e_y \\
e_z
\end{pmatrix} + \begin{pmatrix}
X_{\text{strip centre}} \\
Y_{\text{strip centre}} \\
Z_{\text{strip centre}}
\end{pmatrix}
\]

(1)

In this formulation \(R_{\text{strip_to_ref}}\), \(X_{\text{strip centre}}, Y_{\text{strip centre}},\) and \(Z_{\text{strip centre}}\) define the transformation between the reference coordinate system \((X, Y, Z)\) and the ideal (e.g. error free) strip coordinate system \((x, y, z)\). The deviations from the ideal strip coordinate system are modelled by three offsets \((e_x, e_y, e_z)\), three rotations \((\alpha, \varphi, \kappa)\) and three time-dependent rotations \((\dot{\varphi}, \dot{\kappa}, \dot{\omega})\). It has to be noted that this model does only correct for systematic errors, which cause linear deformations of strips. Local effects, especially those caused by the limited precision of GPS, are not covered by this model and will be more difficult to deal with.

2.2 Measurement of tie points

Due to the aspired limitation of control points to a minimum, the major observations as input into the strip adjustment procedure are tie points in the overlap region between neighbouring or crossing strips. In the presence of systematic strip errors, these tie points will show discrepancies in their coordinates. Obviously, due to the scan pattern of laser scanner systems, datasets representing two strips of laser scanner data will not contain identical points. Therefore, points have to be interpolated. [Crombاغs et al., 2000] chose height tie points by interactively selecting flat areas in strip overlap regions and calculating the differences of the average heights of points within a certain radius in both strips. If strip adjustment shall be extended to planimetry, both height and planimetric tie point discrepancies have to be determined. A simple way to provide these measurements is the interpolation of laser scanner data of both strips to raster height images, allowing for the application of standard image matching tools [Kilian et al., 1996]. Suitable locations for matching with the aim of determining strip discrepancies must contain gradients in three non-coplanar directions. In many regions, such patches are only provided by buildings with according roof shapes. It can be shown that in such cases, matching applied to interpolated height images may lead to a strong bias in the determined shift parameters as a consequence of points in occlusion regions [Maas, 2000].

A solution for this problem is provided by the formulation of least squares matching (LSM) for irregularly distributed points in a TIN structure [Maas, 2000]. This formulation allows for extensions to exclude points in occlusion regions and to restrict matching to planar patches visible in both datasets, resulting in unbiased estimates for all three tie point coordinate discrepancies. Besides these three shift parameters, LSM does also provide estimates for the precision of the estimated parameters. The design matrix in LSM is derived from observations with stochastic properties; in combination with the noise properties of laser scanner data, this causes the values in the covariance matrix to be too large and thus the estimated shift parameter standard deviations to be too optimistic. This applies especially for high density laser scanner data and may have severe consequences for the analysis of the determinability of parameters. Methods to overcome this problem and to obtain realistic figures on precision and determinability of tie point shift parameters are also discussed in [Maas, 2000].

Applied to laser scanner data with moderate point density (1.8 meter average point spacing), standard deviations in the order of one decimetre for the two planimetric shift components \(\sqrt{\beta_{20}}\) point spacing) and one centimetre for the height discrepancies could be achieved [Maas, 2000]. The standard deviations of matching should be used as weight for observations in strip adjustment. In the same experiments, planimetric discrepancies of up to 40 cm and height discrepancies in the order of 10 cm could be detected. These shifts are significant.
While flat regions for determining height discrepancies are easy to find, the requirement for patches with gradients in three non-coplanar directions restricts the choice of suitable patches considerably. One consequence of this restriction should be the implementation of adaptive matching, reducing the number of parameters automatically in the case of non-determinability of some parameters. Another solution is provided by using reflectance data, which are provided by most laser scanner systems, for determining horizontal shifts as suggested by [Burman, 2000] and [Maas, 2000]. A matching strategy for automatically finding suitable regions for matching might be based on an interest operator implemented on irregularly distributed points in a TIN structure. As an alternative, available GIS data may indicate buildings, or buildings may be detected in the laser scanner data automatically, as shown in [Oude Elberink and Maas, 2000].

### 2.3 Adjustment results

The mathematical model described in 2.1 and the tie point measurement method described in 2.2 have been applied to two small data sets. Both data sets contained buildings that could be used for tie point measurements. No ground control points were available for these data sets. Therefore, a minimum number of control points was assumed, such that they did not deform the strips. The four strips of the Eelde test site are shown in figure 1. The two North-South strips cross the two East-West strips.

![Figure 1. Four strips of the Eelde test site.](image)

The preliminary adjustment results are summarised in table 1. Standard deviations were estimated from the residuals before and after the strip adjustment.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>FLI-MAP</th>
<th>ALTM1020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test site</td>
<td>JIJsselstreek (NL)</td>
<td>Eelde (NL)</td>
</tr>
<tr>
<td>Point density (pts/m²)</td>
<td>5-6</td>
<td>0.3</td>
</tr>
<tr>
<td>Flying height (m)</td>
<td>110</td>
<td>500</td>
</tr>
<tr>
<td>Number of strips</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Number of tie points</td>
<td>75</td>
<td>46</td>
</tr>
<tr>
<td>$\sigma_0$ before adjustment (cm)</td>
<td>15.2</td>
<td>35.6</td>
</tr>
<tr>
<td>$\sigma_0$ after adjustment (cm)</td>
<td>9.7</td>
<td>20.3</td>
</tr>
<tr>
<td>$\sigma_0$ improvement (%)</td>
<td>36</td>
<td>43</td>
</tr>
</tbody>
</table>

| $\sigma_X \sigma_Y \sigma_Z$ before adjustment (cm) |
| 16.4 | 20.0 | 5.0 |
| $\sigma_X \sigma_Y \sigma_Z$ after adjustment (cm) |
| 11.2 | 11.6 | 4.7 |
| $\sigma_X \sigma_Y \sigma_Z$ improvement (%) | 32 | 42 | 6 |

Table 1. Strip adjustment results

In both cases no larger systematic errors were present in the data as was already noted by visual inspection. By the adjustment, the standard deviations of the coordinate differences between the tie points reduced by about 40%. Interestingly, the planimetric precision improved much more than the height precision. For the production of accurate 3D city models strip adjustment may therefore be even more important than for the production of DEM's.
The height precision achieved with the FLI-MAP scanner is quite good and is in agreement with other accuracy analyses [Brügelmann, 2001]. The height standard deviation of the ALTM1020 scanner is also quite acceptable, considering that the estimated standard deviation contains the noise in the heights of two different strips and possibly small biases caused by the tie point transfer.

3 FILTERING VEGETATION AND BUILDINGS

For the production of digital elevation models, the many points that are measured on vegetation, buildings and other objects above the ground surface need to be removed from the data set. Several algorithms have been developed for this purpose [Kilian et al., 1996, Pfeifer et al. 1998, Axelson, 2000, Vosselman, 2000]. These filter algorithms make assumptions on the spatial distribution of points in the terrain. By verifying these assumptions points are classified as ground point or not. Huisings and Pereira [1998] show that filtering becomes difficult when objects to be removed (like buildings) are similar in shape to objects that are part of the terrain (e.g. dikes). Whereas the filters can correctly eliminate most points above the ground, there are many cases left in which the laser data by itself provide insufficient information to reliably classify the points without the usage of additional information. This holds for example for points on low vegetation and points on large buildings.

Parts of the OEEPE data set of Vaihingen and Stuttgart have been processed using the slope based filter described in [Vosselman, 2000]. This filter can be implemented using mathematical morphology. In order to deal with larger buildings without the use of large structure element for the mathematical morphology, a post processing step was performed in which larger higher objects were detected.

3.1 Slope based filtering using mathematical morphology

In contrast to the min/max operators, the slope based filter defines the maximum allowed height difference between two points as function \( \Delta h_{\text{max}}(d) \) of the distance \( d \) between these points. When this function are known, the DEM points are defined as a subset of all laser points \( A \) by

\[
DEM = \left\{ p_i \in A \mid \forall p_j \in A : h_{p_i} - h_{p_j} \leq \Delta h_{\text{max}}(d(p_i, p_j)) \right\}
\]  

(2)

It can be shown that this is equivalent to accepting all points below the eroded surface, if the structure element for this erosion is defined by

\[
k(\Delta x, \Delta y) = -\Delta h_{\text{max}} \left( \sqrt{\Delta x^2 + \Delta y^2} \right)
\]  

(3)

Hence, for each point, the eroded height at the same (X,Y) location is computed and compared to the original height. If the eroded height is lower, the point is rejected [Vosselman, 2000].

The size of the kernel has a large effect on the computational effort. To reduce this effort heuristics can be applied that produce virtually the same results in only a fraction of the time used for the strict implementation. Depending on the amount of point above the ground surface the filtering speed varies between 2 and 10 million points per hour (on a SGI O2 10k).

Slope based filtering only accepts a point as a ground point if its height is not much above neighbouring points. The discriminative power of the filter therefore depends on the height differences in the terrain and the point density. The higher the point density, the steeper the slopes between ground points and points on nearby buildings and vegetation will be. Therefore filtering results improve with an increasing point density [Vosselman, 2000]. Steep slopes in the terrain clearly deteriorate the filter results. In order to avoid that all points on these terrain slopes are eliminated, one has to accept large height differences between nearby points. Consequently, similar height differences caused by vegetation, cars or lower buildings will also be accepted.

3.2 Removal of buildings

Buildings are not completely eliminated by morphological filters if the kernel size is smaller than half the size of a building. Figure 2 shows the typical pattern of accepted (white) and rejected (black) points for such cases. Points on the roof and near the edge of a building are rejected due to large height differences between these points and ground points. Due to the limited size of the kernel, points in the interior of the roof are only compared to other points on the roof. In case of flat roofs, these points are not eliminated.
One way to eliminate the interior roof points is to make use of a GIS. This requires an accurate registration of the laser data to the GIS. Errors in the registration will lead to the removal of ground points if these points fall inside a building contour. The usage of a GIS for this purpose, of course, also relies on the completeness of the GIS.

For processing data with large buildings we use a simple heuristic based on the analysis of heights in the connected components of accepted and rejected points. Patterns as shown in figure 2 are detected in the TIN data structure. They consist of a connected component of accepted points, surrounded by a connected component of rejected points, which on its turn is again surrounded by a connected component of accepted points. These components will be named RI (roof interior), RE (roof edge), respectively G (ground).

If the components RE and RI belong to a building it is expected that the points in RE were rejected because of height differences with points of G, but not because of points in RI. For a building, the heights in RI are expected to be higher or equal to the heights in RE. Furthermore, there should be a significant height difference between the points in RI and the points of G that are near the building. Resulting from these considerations, the heuristic for the removal of buildings is defined as follows. All points of a component RI are rejected if

- all points of RI are higher than the lowest point in RE, and
- the median height of points in RI exceeds the median height of those points of G that are adjacent to RE by some assumed minimum building height (e.g. 2 m).

Using this heuristic most buildings can be removed from the dataset. The weak point of the heuristic lies in the assumption that the component of the interior roof points is separated from the component of the ground points. I.e., there should be no TIN edge that connects a ground point to a roof point that was not eliminated by the morphological filter. Sometimes, however, these components are connected. This is the case if there are some reflections on low vegetation and the wall of a building such that there is a path of points from a ground to a roof point along which there are no steep slopes between the successive points. The occlusion of an area beside a building may have a similar effect. In this case the distance between the roof points and the nearest ground points increases. Consequently, the slope between these points decreases and may fall below the slope threshold. Ground points and roof points are then connected by long TIN edges.

### 3.3 Processing the OEEPE datasets Vaihingen and Stuttgart

Two parts with buildings and/or vegetation were selected from the OEEPE datasets of Vaihingen and Stuttgart (figure 3). Clearly visible are the effects of a flight planning error: the strips are not, or not completely, overlapping. The
Voßelman, 2000. When the data from the strips is filtering. Therefore lead to unpredictable results. Hence, in areas with a lot of vegetation, tie points can only be measured after filtering.
Figure 4. Part of dataset Vaihingen. Top: original data. Middle: after slope based filtering. Bottom: after building removal.
Figure 5. Part of dataset Stuttgart. Top: original data. Middle: after slope based filtering. Bottom: after building removal.
Based on these considerations we suggest to split the filter process in two parts. The procedure for the production of a digital elevation model from raw laser altimetry data would then consist of the following steps:

- Filtering the vegetation while keeping the buildings. This can be achieved using morphological filter with a small kernel. In case of very dense vegetation and a low point density the separation between points on vegetation and points on buildings may, however, become difficult.
- Measurement of tie points and adjustment of the strips. Use can be made of tie points on slanted roofs as well as of tie points on the ground surface of areas with vegetation.
- Merging the data sets from the different strips to one seamless data set.
- Filtering of the buildings. Because of the higher point densities in the overlaps between the strips, some points on low vegetation, that were not removed in the first step, may now be detected and filtered.
- Finally, for many applications a data reduction may be required.

5 DISCUSSION

Although digital elevation models produced by laser altimetry are already successfully used for many applications, processing procedures are not yet standardised. Strip adjustment is often not a part of the procedure used in practice. Software for strip adjustment is not yet commercially available and the mathematical models for the adjustment are still under development. One might argue that strip adjustment will not be necessary in future. Once all laser altimetry surveys are more controlled and routinely calibrated, the acquired laser data may show no or only very small systematic errors. For the time being, however, one can not be sure that systematic errors can be neglected. Performing a strip adjustment is recommended as good practice.

Filtering procedures are used more often in practice, simply because filtering by hand is just to labour-intensive. Several algorithms have been proposed that use quite different concepts. For the further development of the filtering procedures it would be interesting to compare these filters on different types of terrain.

Both the selection of good tie points and the filtering become easier if the point density increases. With the increasing pulse rates of airborne laser scanners, the obtainable point densities can be much higher than the densities that are usually required for the resulting digital elevation models. To improve the processing results one should therefore consider to scan the terrain with higher point densities than strictly required for the DEM production. The higher costs for the acquisition of the raw laser data may very well be offset by the better possibilities to automatically process these data.

ACKNOWLEDGMENTS

The laser altimetry data used for the research in this paper was kindly provided by the Survey Department of the Ministry of Transport, Public Works and Water Management (Delft/NL) and by Geodelta B.V. (Delft/NL) and Fotonor (Sandefjord/NL).

REFERENCES


