LINE-BASED REGISTRATION OF TERRESTRIAL AND AIRBORNE LIDAR DATA

Wolfgang von Hansen, Hermann Gross, Ulrich Thoennessen

FGAN-FOM Research Institute of Optronics and Pattern Recognition
Gutleuthausstr. 1, 76275 Ettlingen, Germany
`wvhansen@fom.fgan.de`

Commission III

ABSTRACT:

When LIDAR data is considered as data source for 3D models of urban areas, there is the choice between airborne and terrestrial systems. They have different characteristics like resolution or scene extent. One of the most significant differences is that city models from airborne scanners typically lack any facade structures. It seems natural to combine both types of LIDAR data for better results. The key to this type of data fusion is a successful registration, i. e. the geometric transform of one data set into the coordinate frame of the other. This paper proposes a method that is based on feature lines that are extracted in a similar way from both data sets. The emphasis is on the coarse registration – the estimation of initial transformation parameters without any prior knowledge.

1 INTRODUCTION

There exist two different platforms for the acquisition of LIDAR data, terrestrial platforms and airborne platforms like planes or helicopters. The acquired data differs in resolution, scene extent, the viewpoint and the georeferencing. Airborne systems typically have a resolution of still only 1 point per square meter – even though multiple passes may yield up to 5–20 points per square meter – while terrestrial systems may feature thousands of points per square meter. The reverse is valid for the extent of the scene: Airborne scanners are able to capture large areas the size of a city in relatively short time while terrestrial scanners are much slower and can only be used for limited areas. Another disadvantage of terrestrial data usually is the lack of georeferencing information while airborne scanners are equipped with a navigation system consisting of high quality GPS and INS units.

If LIDAR data from both types of sensors exists for a region, it can be advantageous to fuse them into a common data set. The primary objective would be to enhance the point density of the airborne scanner and – for built-up areas – to add facade structures to the data set. The resulting 3D model would have a varying LOD (level of detail) as most of the scene is only captured by the airborne scanner and therefore is only available in a low resolution. The interesting parts have been imaged by both sensors and here the higher point density and the different vantage point of the terrestrial scanner greatly enhances the result. As a side effect, this would help georeferencing the terrestrial data.

The advantages of such a combination have already been emphasized by (Boehm and Haala, 2005). Their approach for an initial registration had been direct geocoding based on a GPS system and an electronic compass. Another example is (Ruiz et al., 2004) where LIDAR DTMs of a mountain region are augmented by several terrestrial scans. The correspondences have been established based on surveyed points.

In this paper, we investigate a method for an automatic coarse registration of airborne and terrestrial LIDAR data for the special case of urban areas. It is based on feature lines that describe discontinuities in both data sets. These typically occur at the boundaries of planar surfaces. The matching strategy does not require an initial position, but instead tests all possible solutions to find the one that best fits both data sets. Rotation and translation are found separately in two consecutive steps.

2 METHODOLOGY

2.1 General Approach

There are three main characteristics in which airborne and terrestrial LIDAR data of built-up areas differ:

1. Different point of view onto the scene. Airborne systems are typically looking downward with a rather small field of view so that shadow effects are minimized. From buildings, only the roof tops are imaged. There may be some 3D points on walls but they are far too few to allow a direct reconstruction of building walls.

   Terrestrial LIDAR systems are able to make 360°scans from between the buildings. The ground and building walls make up most of the scene content. Inclined roofs can be imaged while flat top roofs are outside the field of view from a scanner at street level.

2. There is a huge difference in the point density of the two systems. Airborne systems typically acquire only one point per square meter – even though multiple passes may yield up to 5–20 points per square meter. This is sufficient for building reconstruction based on planar surfaces in most cases, but the delineation of roof planes already suffers when there are a lot of smaller structures, like e. g. dormers, on the roofs.

   Terrestrial scanners, on the other hand, capture literally millions of 3D points in one scan. Being based on a polar coordinate system, the point density is higher close to the scanner and might go up to one million points per square meter for the Z+F scanner used for the data in this paper. At its maximum range of about 50 m, there are still several thousand points per square meter on planes which point to the scanner. Very small scene structures in the size of a few centimeters can be reconstructed if needed.

   While airborne scans of built-up areas still suffer from a limited point density, the opposite is true for terrestrial scans.
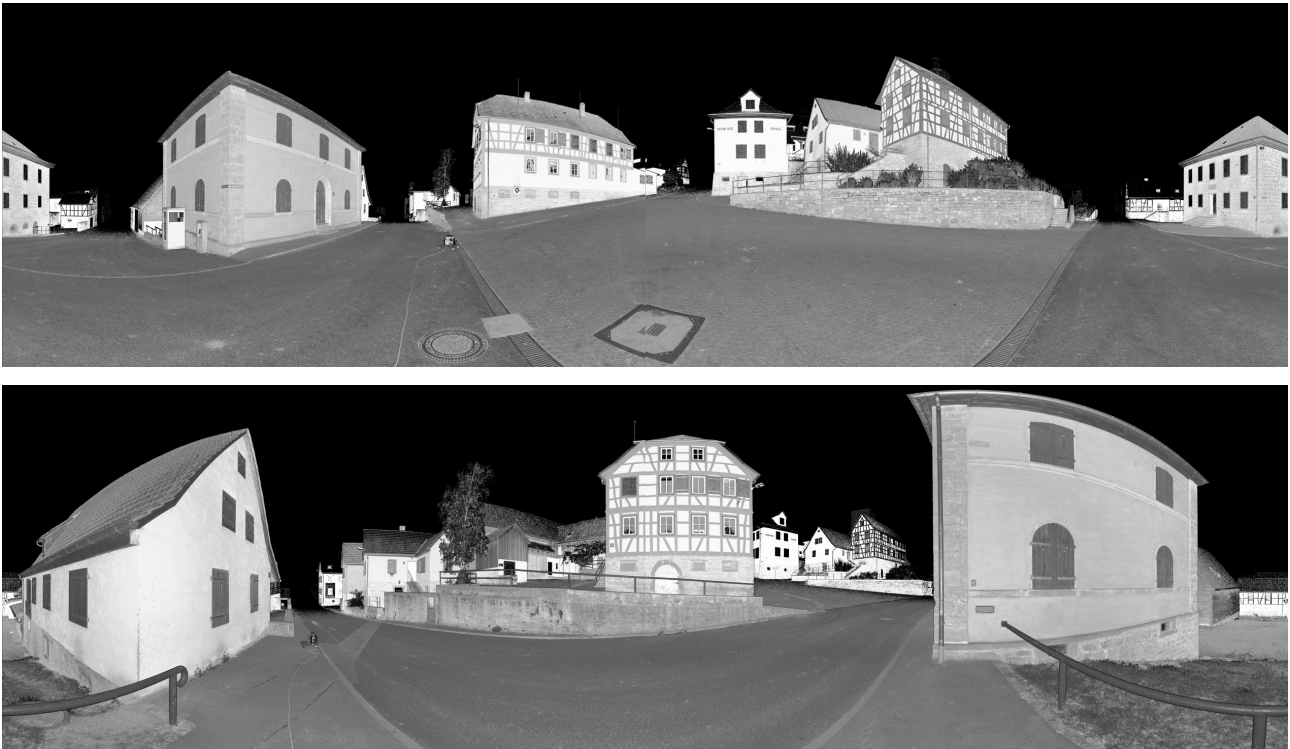
Figure 1. Panoramic intensity images from the two terrestrial scan positions.

Here, the number of points is so high, that proper data management becomes an issue and appropriate thinning techniques are required.

3. The scene extent of a single scan differs largely between the two systems. Airborne scanners acquire long strips that may have a width of a few hundred meters depending on flight altitude and required point density. Since a precise navigation system is mandatory for the operation of such a line scanner, registration of multiple strips into a larger data set is possible quite easily. This way, one large data set covering an extended area can be created.

Despite the huge number of points in a data set from a terrestrial scanner, the covered area is very limited. Only objects in the line of sight from the scanner position are imaged. For buildings, only the facades pointing to the scanner can be seen and occlusion often leads to facades that are partially mapped.

Despite the differences, some common structures on which the registration can be carried out must be identified. In addition, these have to be put into a representation that allows a feasible matching procedure. Finally, the two different products must be made compatible to overcome the differences mentioned previously.

Regarding the common scene content, only the ground surface is imaged by both sensor types. The walls are missing from the data sets from the airborne scanner while – under the assumption of flat roof tops – there are no roofs in the data from the terrestrial scanner. In addition, there is more clutter in the terrestrial data as for instance some of the insides of the buildings are mapped through the windows.

The methodology followed here is to bring the data sets into a similar representation from which common structures can be extracted. This will be achieved here by the generation of feature lines from the point clouds. This allows a representation that is independent from the original resolution of the data. The amount of data required to represent the scene is greatly reduced.

2.2 Features to match

Even though the actually mapped areas of the scene are quite different – building walls in one and roofs in the other data set – they share common lines at which these areas meet. These lines are either the intersection of two planar areas that are actually measured, like a building wall and an inclined roof in the terrestrial data, or they are the bounding lines of a set of contiguous points, like the roof surface in the airborne data. The extraction of these lines from the point cloud is based on a classification using the statistics of a point's surroundings as shown in (Gross and Thoennessen, 2006).

The matching itself is inspired by (von Hansen, 2006). A single pair of extracted feature lines is sufficient to compute all transformation parameters. These are a rotation around the vertical axis and a 3D translation. The simplification for the rotation is possible because both data sets already have aligned vertical axes. An inlier count after the transformation will be used to find the correct solution. Beyond this, we will not assume any initial values for the registration. This means that we will try to locate the position of the terrestrial scanner in a larger built-up area. Repetitive patterns might lead to wrong results, but hints at location could be exploited to rule out some of the wrong solutions. Another way could be to first register several terrestrial data sets in order to produce a larger set of feature lines and then to register this to the airborne data set. In addition we will outline a fine registration technique based on the feature lines.

2.3 Generation of lines from point clouds

2.3.1 Introduction A laser scanner delivers 3D point measurements in a Euclidean coordinate system. For airborne systems mostly the height information is stored in a raster grid with

Table 1. Eigenvalues for four typical situations

S	Type	λ_1	λ_2	λ_3
	End of a line	$\frac{1}{12}$	0	0
	Line	$\frac{1}{3}$	0	0
	Half plane	$\frac{1}{4}$	$\frac{1}{4}\left(1 - \frac{64}{9\pi^2}\right) \approx 0.07$	0
	Two planes	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8} - \frac{8}{9\pi^2} \approx 0.03$

a predefined resolution. Image cells without a measurement are interpolated by considering their neighborhood.

In contrary to the airborne data, the projection of terrestrial laser data along any direction is not very reasonable. Especially the combination of airborne and terrestrial laser scanning data requires directly the analysis in the 3D data.

2.3.2 Moments A 3D spherical volume cell with radius R is assigned to each point of the cloud. All points in a spherical cell will be analyzed. 3D moments as described by (Maas and Vosselman, 1999) are discussed and improved by (Gross and Thoennessen, 2006)

$$m_{ijk} = \frac{\sum_{l=1}^{N}(x_l - \bar{x})^i(y_l - \bar{y})^j(z_l - \bar{z})^k}{R^{i+j+k}N}, i + j + k = 2. \quad (1)$$

Neither the number of points nor the chosen physical unit for the coordinates, the radius and the weighting factor influences the values of the moments.

2.3.3 Filtering of points After calculation of the covariance matrix for each point in the data set considering a local environment defined by a sphere we have additional features for each point. These features are the center of gravity, the distance between center of gravity to the point, the eigenvectors, the eigenvalues and the number of points inside the sphere. They can be used for determination of object characteristics.

Tab. 1 shows the eigenvalues of the covariance matrix of four special point configurations. Other situations are presented in (Gross and Thoennessen, 2006). The ratios in the table have been calculated analytically. For instance, an ideal line has two eigenvalues which are zero and one that is greater than zero.

Fig. 2 (top) shows all points from an example data set with eigenvalues satisfying the criteria for planes. The color indicates the object height. In Fig. 2 (bottom), only the edge points corresponding to Tab. 1 row 4 are drawn for the same data set.

2.3.4 Line generation All points marked as edge point may belong to a line. These points are assembled to lines by a grouping process. We consider the greatest eigenvalue λ_1 and its eigenvector $\mathbf{e}_1$. Consecutive points with a similar eigenvector, lying inside a small cylinder are grouped together and approximated by a line.

We start with one of the points as trigger point $\mathbf{p}$ and look for more points in the direction of $\mathbf{e}_1$ that in addition have a similar first eigenvector. This set contains all points with nearly the same or opposite direction for the first eigenvector tested comparing the inner product of two vectors against a given threshold. We construct a line through the trigger point along its first eigenvector:

$$\mathbf{g} = \mathbf{p} + \mu\mathbf{e}_1 \quad (2)$$

The set of edge points inside the cylinder given by $\mathbf{g}$ with a given radius. The intersection of the cylinder with the point cloud includes all edge points with nearly the same first eigenvector as
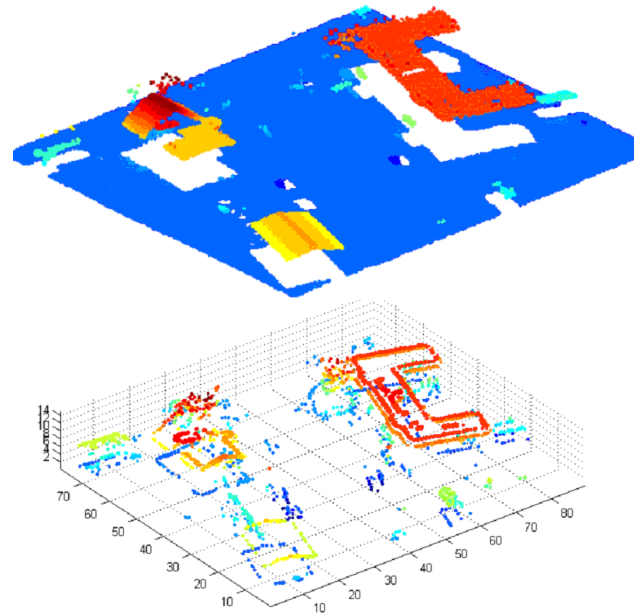

Figure 2. Top: Points identified as plane points. Bottom: Points with one high and two small eigenvalues.

the trigger point and not far away from the straight line given by the trigger point and its first eigenvector.

Collinear edges of different buildings in a row may belong to the same group. Therefore we examine the contiguity of the points in the neighborhood of $\mathbf{p}$. Once both start and end points are found, they are output as line segment. The same process is repeated for all points not assigned to a line until each point belongs to a line or can not generate an acceptable line.

2.4 Registration

2.4.1 Registration basics At this stage there exist two sets of line segments, one from the terrestrial scanner and one from the airborne scanner. These line segments are feature lines that describe scene structures. These should be only from objects like buildings, but vegetation may lead to some false responses. Their irregular nature makes it possible to rule them out as valid correspondences during registration, but they still require some processing time. For the terrestrial data set, the features can be from individual building facades, but also other structures that show systematic distance jumps. For instance, the shadow of a building on the ground induces an edge in space which is extracted as feature line but does not correspond to any existing scene structure. The major achievement of the representation as lines is the reduction of the raw point clouds to a small set of features that still carry geometric information and that can be matched between the data sets in order to perform the registration.

Registration is the geometric transform of one data set into the coordinate frame of the other so that corresponding parts overlap. This is a mandatory step that precedes any form of data fusion. Registration can be divided in two sub steps, coarse registration that is used to determine the initially unknown transformation parameters with little or no prior information, and fine registration that refines the parameters obtained by the coarse registration – or other external sources like e.g. a navigation system – to an optimal solution.

Our solution for the coarse registration is influenced by the approach from (von Hansen, 2006). There, it had been used for

Figure 3. Feature lines generated from airborne data.

a plane-based coarse registration of two terrestrial LIDAR data sets. Since we use lines as features here, the underlying geometry is almost the same. The main difference is that the two data sets from the two sensors differ significantly in their size: The terrestrial data set can be located anywhere in the airborne data set. Therefore, there is a high potential for false assignments but once the feasibility is shown in principle, enhancements can be applied to overcome such drawbacks.

In contrast to the previous work, solutions to rotation and translation will be found in two separate steps. For the rotation, the direction histograms will be exploited. Translation is then found using the existing generate-and-test strategy based on single pairs of corresponding lines. After the coarse registration we will use a line based fine registration step that improves the translation result.

2.4.2   Rotation   The rotation is estimated separately and prior to that of the translation. The basic idea is that the orientation of the lines in built-up areas is not random. Basically, there are buildings with rectangular structures having two main directions. As the individual buildings are often aligned, they lead to significant maxima in the histograms that can be used to identify relative rotation.

Estimation of rotation is based on the correlation of orientation histograms. For each data set, the direction in the horizontal plane of each line is computed and added to the histogram. There is an ambiguity in the direction of a line, so the histograms are 180°-periodic.

The correlation function between the two histograms is computed numerically. Maxima in the correlation function hint at correct solutions for the rotation angle. These maxima are detected by a threshold for the correlation coefficient. A parabola is fitted to the correlation function at each maximum. The location of the parabola's maximum is used as solution for rotation. The value of the maximum can be used to rank multiple solutions. However, this had not been done in this work. We estimated a translation solution for each angle found.

Since the correlation function is 180°-periodic, each maximum actually corresponds to two rotation angles that differ by 180°.

2.4.3   Translation   As there are at least two solutions for the rotation returned by the previous step, it is necessary to estimate a translation solution for each of these and pick the best one.

First, one data set is rotated so that corresponding lines should have the same orientation. Then, a solution for the translation is found via a complete search in a generate-and-test scheme: For each possible line correspondence, where orientations match, the
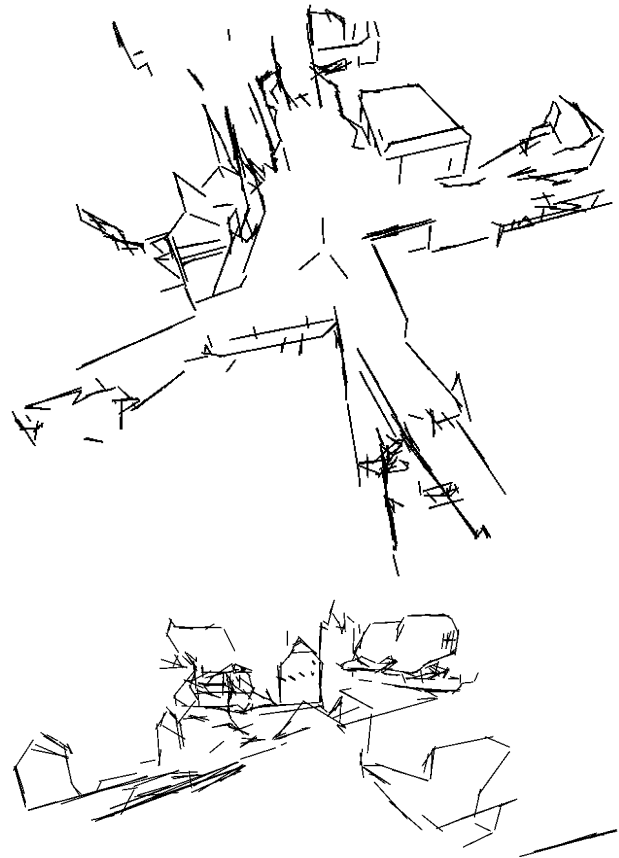

Figure 4. Feature lines generated from the two terrestrial scans.

translation is computed from the line's midpoints. For each translation, the number of inliers is counted: Pairs of lines that have matching orientations and where the midpoint of one line is close to the corresponding line. Larger shifts, up to the length of the line, along the lines therefore are possible and still count as inlier. This accounts for the weakly defined endpoints of each line segment.

The translation that leads to the largest number of inliers is taken as solution. The correct rotation angle is the one that corresponds to the best translation.

2.4.4   Fine registration   The rotation is already defined quite precisely for two reasons. First, the orientation histograms had been computed from all lines and therefore contain some averaging against noise. The correlation function has got a similar property. Second, the angles had been estimated based on fitting a parabola to the data, leading to higher accuracy.

The translation, on the other hand, is estimated only from a single line correspondence. Any errors contained in this particular correspondence will propagate to the result. Therefore, the translation had been re-estimated from all correspondences that counted as inliers as follows:

Each line is represented by its midpoint $\mathbf{x}$ and direction $\mathbf{v}$. For a pair of corresponding lines, the translation is

$$\mathbf{x}'_i = \mathbf{x}_i + \mathbf{t} + \alpha_i \bar{\mathbf{v}}_i \qquad (3)$$

where $\mathbf{t}$ is the global translation and α_i an individual slack variable. The midpoint $\mathbf{x}_i$ of the first line is translated such that it coincides with the midpoint $\mathbf{x}'_i$ of the second line. There is an
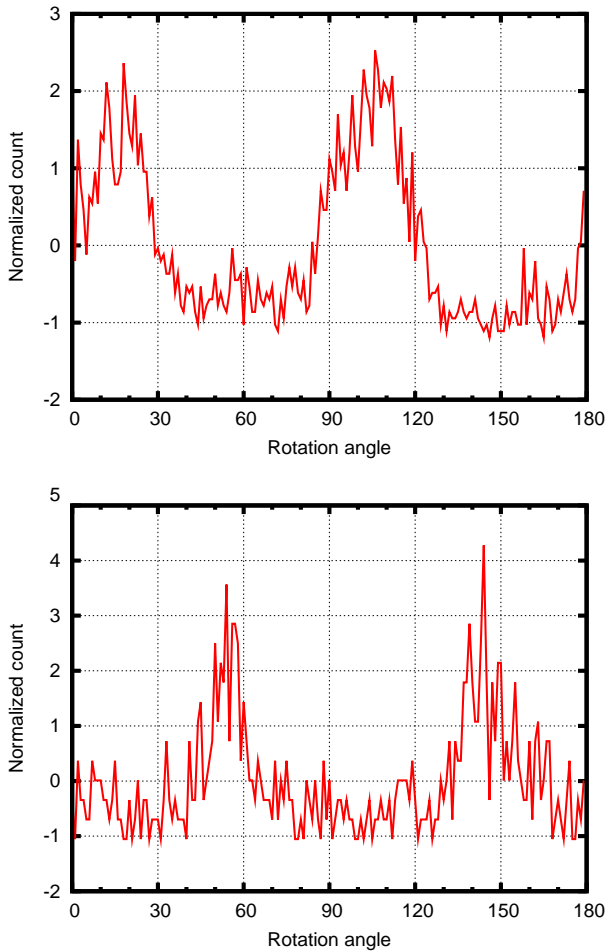
Figure 5. Normalized orientation histograms for the airborne (top) and terrestrial (bottom) feature line sets.

ambiguity in the choice for the orientation vector $\bar{\mathbf{v}}_i$. Here we took

$$\bar{\mathbf{v}} = (\mathbf{v} + \mathbf{v}')/2 \qquad (4)$$

as both orientations are known to have similar values for matching lines. The slack variables α_i compensate the shift along the lines – their actual values can be ignored. Eq. 3 can be reformulated as

$$\mathbf{t} + \alpha_i \bar{\mathbf{v}}_i = \mathbf{x}'_i - \mathbf{x}_i \qquad (5)$$

which can be solved as a linear equation system with $\mathbf{t}$ and α_i as unknowns.

Each line correspondence adds three equations and one unknown. The system for N correspondences contains $3N$ equations and $3 + N$ unknowns. Therefore, at least two line correspondences are required for a solution to fine registration.

3 EXPERIMENTS AND RESULTS

As input one data set from an airborne scanner and two data sets from a terrestrial scanner (Fig. 1) have been used. The task was to determine a solution for the relative transformation between both terrestrial and the airborne data set based on extracted feature lines.

3.1 Extraction of feature lines

The extracted feature lines are shown in Figs. 3 and 4. In the airborne data, one can see the outline of the data set from the
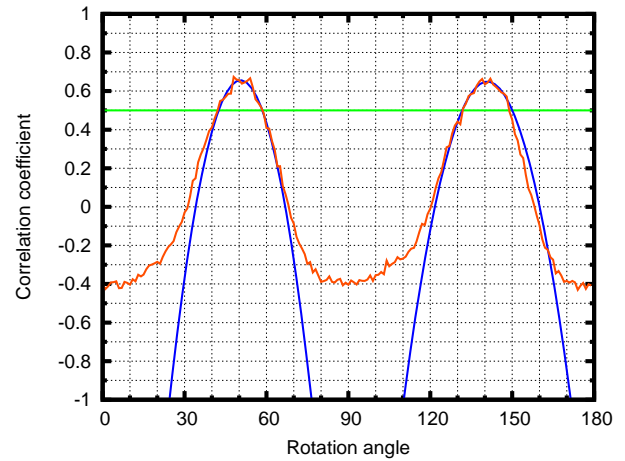

Figure 6. Correlation result for the two histograms in Fig. 5.

lines drawn as borders. In the terrestrial data set, there are some artifacts pointing radially away from the center. Note that not all feature lines that one would expect can be extracted from the data. For instance, in the airborne data, there are no vertical lines at the building corners, as there are no points acquired at these positions.

3.2 Rotation

The orientation histograms for the airborne and one of the terrestrial data sets are shown in Fig. 5. The count is normalized such that the mean is zero and the standard deviation one. This is a prerequisite for the correlation. Both histograms have two peaks 90° apart from the other. These correspond to the rectangular structures in the scene. The peaks are sharper for the terrestrial data as there is less noise contained in its orientations.

The correlation result is shown in Fig. 6. The correlation function is plotted in red. The horizontal green line is the threshold for maximum detection. The parabola fitted to the maxima are in blue. Because of the 90°-ambiguity from the rectangular scene structures, there are two correlation maxima. Along with the 180°-ambiguity of line orientations, a total of four angles are returned as possible solution for the rotation.

3.3 Translation

The translation estimation returned 11 inliers for the first and 10 for the second data set. In both cases, the inlier rate was lower for the wrong angles. The fine registration lead to shifts of 19 cm for the first and 50 cm for the second data set. This shows roughly the quality of the coarse registration result.

The results are shown in Fig. 7. One can see that many lines exist in the terrestrial data sets (red) that do not have a corresponding partner in the airborne data set (green).

3.4 Remarks

Even though the scheme works for the two data sets shown here, difficulties are expected when there are repetitive patterns in the scene that do not allow a unique solution. There is also the possibility of false solutions because there is no test for contradiction: Only the number of correct correspondences is counted, but not the number of feature lines that cross after transformation which is physically impossible.

Experiments showed that the solutions are not too stable. The main reason is that interior and exterior of the buildings can no

longer be distinguished from the line representation. Additional information should be used to rule out impossible solutions. One approach would be to extract raster ground and building maps from the point cloud, superimpose them and look for contradictions.

Once a fine registration between multiple terrestrial data sets has been carried out, these could be fused to larger data sets with more unique structures to be joined with airborne data.

4 SUMMARY AND OUTLOOK

In this paper we have shown a method for the coarse registration of terrestrial and airborne LIDAR data of built-up areas. It is based on line features created inspection of local point density statistics. The solution had been found using orientation histograms for the rotation and a generate and test scheme for the translation parameters by matching all possible combinations of two line segments. The test to find the optimal solution was based on the inlier count.

First results showed that this technique was able to find solutions for the two data sets tested here. Additional work is required for the testing of more data sets and to find solutions when the scene structure is less unique than in the examples shown here.

REFERENCES

Boehm, J. and Haala, N., 2005. Efficient integration of aerial and terrestrial laser data for virtual city modeling using lasermaps. In: G. Vosselman and C. Brenner (eds), Laser scanning 2005, IAPRS, Vol. XXXVI-3/W19.

Gross, H. and Thoennessen, U., 2006. Extraction of lines from laser point clouds. In: W. Förstner and R. Steffen (eds), Photogrammetric Computer Vision, IAPRS, Vol. XXXVI Part 3.

Maas, H. and Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. ISPRS Journal of Photogrametry & Remote Sensing 54, pp. 153–163.

Ruiz, A., Kornus, W., Talaya, J. and Colomer, J. L., 2004. Terrain modeling in an extremely steep mountain: A combination of airborne and terrestrial lidar. In: O. Altan (ed.), Proc. of the XXth ISPRS Congress, IAPRS, Vol. XXXV-B5.

von Hansen, W., 2006. Robust automatic marker-free registration of terrestrial scan data. In: W. Förstner and R. Steffen (eds), Photogrammetric Computer Vision, IAPRS, Vol. XXXVI Part 3.
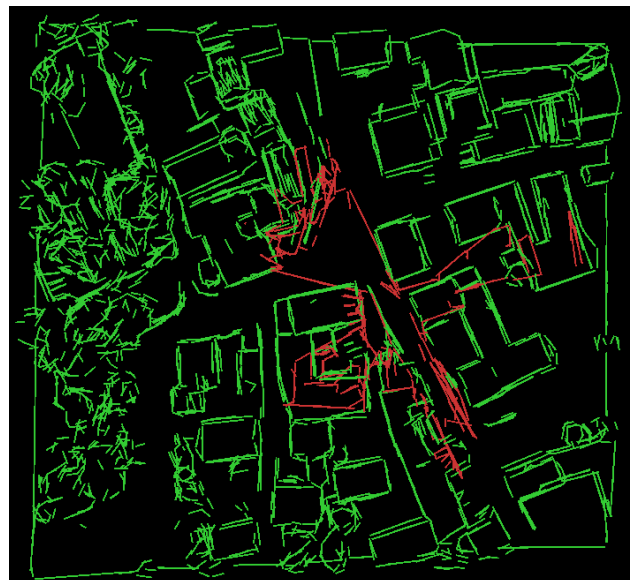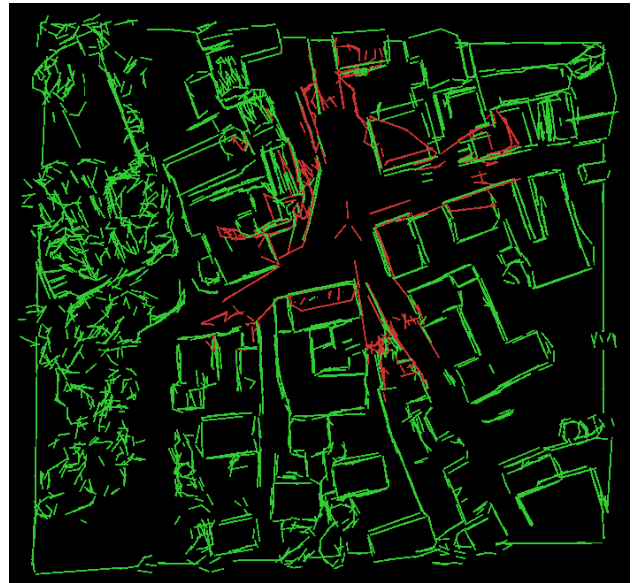

Figure 7. Registration results for the two terrestrial scans.