AUTOMATIC MATCHING OF TERRESTRIAL SCAN DATA AS A BASIS FOR THE GENERATION OF DETAILED 3D CITY MODELS

Christoph Dold[1] **and Claus Brenner**

Institute of Cartography and Geoinformatics, University of Hannover, Germany
{Christoph.Dold, Claus.Brenner}@ikg.uni-hannover.de

Working Group III/6

KEY WORDS: LIDAR, Laser scanning, Fusion, Matching, Registration

ABSTRACT

The request for three-dimensional digital city models is increasing and also the need to have more precise and realistic models. In the past, 3D models have been relatively simple. The models were derived from aerial images or laser scanning data and the extracted buildings were represented by simple shapes. However, for some applications, like navigation with landmarks or virtual city tours, the level of details of such models is not high enough. The user demands more detailed and realistic models. Nowadays, the generation of detailed city models includes usually a large amount of manual work, since single buildings are often reconstructed using CAD software packages and the texture of facades is mapped manually to the building primitives. Using terrestrial laser scanners, accurate and dense 3D point clouds can be obtained. This data can be used to generate detailed 3D-models, which also include facade structures. Since the technology of laser scanning in the field of terrestrial data acquisition for surveying purposes is new, the processing of the data is only poorly conceived. This paper makes a contribution to the automatic registration of terrestrial laser scanning data recorded from different viewpoints. Up to now, vendors of laser scanners mainly use manual registration mechanisms combined with artificial targets such as retro-reflectors or balls to register single scans. Since these methods are not fully automated, the registration of different scans is time consuming. Furthermore, the targets must be placed sensibly within the scan volume, and often require extra detail scans of the targets in order to achieve accurate transformation parameters. In this paper it is shown how to register different scans using only the measured point clouds themselves without the use of special targets in the surveyed area.

1 INTRODUCTION

The fact that many research groups and institutions are engaged in the field of 3D city modelling shows how important it is to expand existing databases with the third dimension. There are a lot of applications for 3D city models, but an area-wide use has been prevented by the lack of economic and fast methods to obtain such models. Up to now city models have only been created for single cities or even parts of cities, despite the fact that many institutions and companies are interested in them. City models are often provided only in a low level of detail. For example in Germany, *Phoenics* is stocked with a selection of digital city models that cover more than 30,000 km^2 (Phoenics GmbH, 2004). The most important applications of city models in the future are the integration of 3D data in car navigation systems and in urban GIS systems for disaster management, the support of the urban planning process, virtual walkabouts in cities for tourists, simulation and analysis purposes and the video game industry. This is a wide field and the applications differ in the demands on the city models. On the one hand detailed models, for example for urban planning, are required. Facades and also small parts of a building have to be modelled. On the other hand, for example for navigation systems, a complete coverage is required. Detailed models are only needed for single points of interest, standard buildings can be represented as a simple model. To cope with such diverse requirements, several levels of detail (LoD) have been introduced. A LoD concept for 3D city models is discussed in *Kolbe and Gröger* (Kolbe and Gröger, 2003).

In the past, several methods for the extraction of buildings for the generation of city models have been proposed. They can be classified according to the data source they use. In the following some previous work is mentioned briefly. *Brenner* developed a semiautomatic approach and uses ground plans of buildings and aerial laser scanning data to derive building models (Brenner, 2000). The software CyberCity Modeler can derive building models from photogrammetric stereo images (CyberCity AG, 2004). An early approach to combine DSM's with stereo images to extract parametric buildings is presented by *Haala* (Haala, 1996).

Today, users of 3D city models also ask for detailed and visually interesting models. Therefore, the airborne based data is not sufficient. Detailed terrestrial data is required to improve the quality of city models. Terrestrial laser scanning is one possibility to collect the amount of information which is needed for detailed building modelling. In the next chapter it is described how terrestrial laser data is currently acquired and the potential to simplify the measurement procedure is exposed. In chapter 3, a method for the registration of scans is presented using only the scan data itself. The method is based on the identification of 3D

[1]Corresponding author

planes in different scans. First planar surfaces are extracted from overlapping scans, then the transformation parameters are calculated. Finally, chapter 4 contains a complete example and the results are discussed. In the end an outlook to future research topics is given.

2 TERRESTRIAL DATA ACQUISITION

Terrestrial laser scanners for surveying purposes are becoming more and more popular. They are used for different tasks, mainly for monitoring measurements, documentation of monuments and historic buildings and for data acquisition of buildings as a basis for city modelling. Several manufacturers have developed new scanners and brought them on the market. The aboriginal costs are passable and the operation costs are low, so the distribution of the scanner equipment is advantaged. All vendors use a similar principle of the laser scanner's functionality: A laser-light pulse is emitted, parts of it are reflected at an object and received and registered by the instrument. The distance is measured by the time of flight of the laser pulse. The laser beam is usually deflected by rotating or oscillating mirrors in the vertical direction. The horizontal deflection results from rotating the whole scanner in discrete steps. Modern scanners cover an area of 360 degrees horizontally and are often termed as panorama scanners for this reason. The vertical measurement range usually lies between -45° to +45° depending on the manufacturer and the scanner model. Terrestrial laser scanners are classified by their maximum measurement range. Usually they are divided in three groups: close range, mid range and long range scanners. There are considerable differences in the accuracy of the single measurements and the spot sizes of the laser scanner instruments. The Institute for Spatial Information and Surveying Technology at the University of Applied Science in Mainz (i3mainz) accomplished tests with different scanner models, which investigate the quality of points recorded by laser scanners (Böhler et al., 2003). Among other things, the angular resolution, range accuracy, resolution information resulting from the angular resolution and spot size, edge effects and the influence of surface reflectivity are tested.

Our research group operates a Riegl LMS 360i scanner combined with a digital camera used for the acquisition of image data. The maximum measurement range of the scanner is 200 meters and a measurement rate up to 12000 points per second is possible. The accuracy of a single shot is 12 mm. The instrument is mainly used to obtain 3D point clouds and images from buildings and facades in order to derive detailed building models from the scanned data.

In order to scan buildings completely, several scan positions are necessary. Figure 1 shows a possible configuration of the measurement. For each position a new local coordinate system is defined by the position and orientation of the scanner. There are two possibilities to orient the systems in a global coordinate frame. First the coordinates of each scan position are known in the local system. In this case the instrument must be oriented to the North or

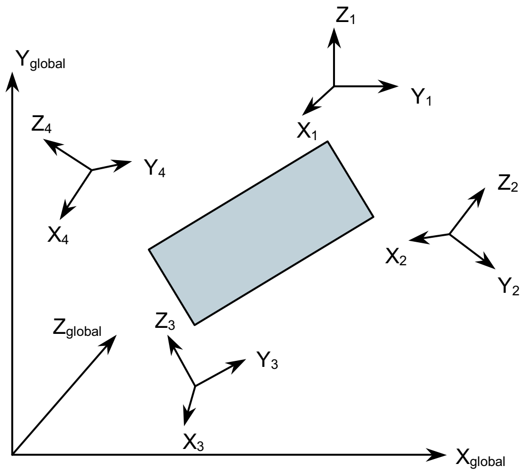

Figure 1: Scanner- and global coordinate systems

to a remote object. Then the measured points are calculated in the global system directly. In practice this method is only used in exceptional cases. The second method is to transform the measurements afterwards into the derived coordinate system. For this, the transformation parameters between the single scan positions must be determined. In most cases this is done by using identical points within two different scan positions. This process is also termed registration between scans. Thereby one scan position defines the reference system. The identical points are represented by special targets, for example cylindrical targets with retro film on them. Figure 2 shows an example of such a target.

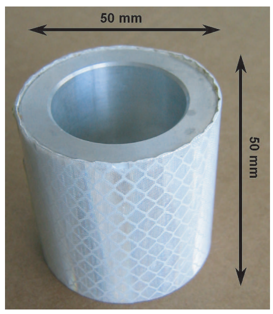

Figure 2: Reflector target

At least three targets must be placed in the field and scanned from the different positions. Afterwards the targets must be identified in both scans. This procedure is supported by the software, but there is still manual work necessary. Then the transformation parameters are calculated and the scans are registered. The whole registration process is time consuming and if the scanner for example is mounted on a car to scan a street of houses faster, the demand to distribute and collect targets is unacceptable. Using a navigation system for determining the scan position and orientation is expensive and the mounting of the scanner on the car is complex, because the deviation between scanner origin and navigation system have to be determined.

3 REGISTRATION PROCESS

3.1 Representation of laser scans as raster images

Terrestrial laser scanners are measuring in a regular pattern. The laser beam is deflected stepwise in vertical and horizontal direction and the step width can be defined. Since the scanner is usually mounted on a tripod and is not moving during the measurement, the local neighbourhood of recorded points is preserved. This fact enables storing the scan data in regular raster images, each directly measured or derived value being stored in a separate layer. The density of the measured points depends on the distance of the object and the defined step width. Neighbouring points within the scan pattern are usually also neighboured in reality. For storing such 3D images a standard image file format supporting float values, for example the "Tagged Image File Format (TIFF)" (Adobe, 1992) can be used. The advantage of storing the data in a regular raster is the possibility to use raster based image processing algorithms for the scan data. Algorithms used for 2D data can be adapted to the third dimension. Figure 3 illustrates the layer principle of the data storage and figure 4 shows an example of the different layers of a laserscan. The three layers contain the measured x, y and z values, with the local coordinate system being oriented towards the facade of the building. The x-axis runs along the facade, the y-axis is perpendicular to the facade and z represents the height.

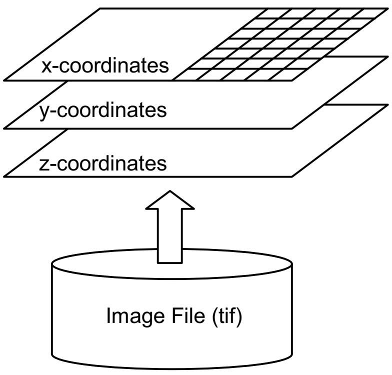

Figure 3: Data storage scheme

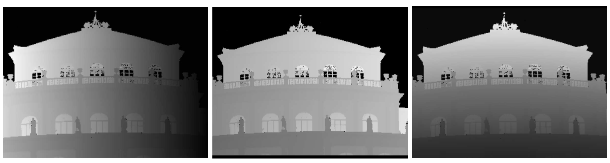

Figure 4: Laserscan of the opera in Hannover represented as x, y and z image

3.2 Extraction of 3D planes from the scan data

For the extraction of planes, a region growing algorithm has been adapted to the terrestrial scan data. The principle in general is to define a seed region first. From this initial region all neighbouring points are investigated and it is checked, if they fit to the seed. Rules for the fitting process are used to decide whether the investigated point is added to the seed or not. A region grows until no more neighbouring points are added to the region. Then a new seed region is selected from the remaining points and the process starts from the beginning. All steps are repeated, until all points are assigned to a region or other termination criteria become true. The regular storage of the scan data in an image containing layers with the coordinates enables the access to the measured values in a correct order and simplifies the 3D region growing. Neighbouring points are read out easily and can be used to estimate a plane which fits to the measured points. Therefore the measured x, y and z coordinates of a scan are used.

3.2.1 Planar Fit This section describes how to fit measured 3D points to a planar surface using the so-called technique of orthogonal regression: a plane is estimated by minimizing the square sum of the orthogonal distance to the given points (Duda and Hart, 1973). This method is also described in *Drixler* (Drixler, 1993) and several methods for the determination of adjusted planes are compared and discussed in *Kampmann* (Kampmann, 2004). The equation of a plane is:

$$a \cdot x + b \cdot y + c \cdot z + d = 0 \tag{1}$$

For the planar fit the parameters a, b, c and d must be estimated. Using the technique of orthogonal regression the vector containing the parameters a, b and c is determined by calculating the eigenvector belonging to the smallest eigenvalue λ_{min} of the matrix:

$$\mathbf{M} = \begin{pmatrix} \sum x_i^C \cdot x_i^C & \sum x_i^C \cdot y_i^C & \sum x_i^C \cdot z_i^C \\ \sum x_i^C \cdot y_i^C & \sum y_i^C \cdot y_i^C & \sum y_i^C \cdot z_i^C \\ \sum x_i^C \cdot z_i^C & \sum y_i^C \cdot z_i^C & \sum z_i^C \cdot z_i^C \end{pmatrix} \tag{2}$$

The indices C indicates the reduction of the coordinates by the centre of gravity:

$$x_i^C = x_i - \frac{\sum x_i}{n} \tag{3}$$

$$y_i^C = y_i - \frac{\sum y_i}{n} \tag{4}$$

$$z_i^C = z_i - \frac{\sum z_i}{n} \tag{5}$$

After calculating the eigenvector and consequently the parameters a, b and c, the fourth parameter d of the plane equation is calculated by:

$$d = -\frac{1}{n}(\sum x_i \cdot a + \sum y_i \cdot b + \sum z_i \cdot c) \tag{6}$$

The minimum eigenvalue λ_{min} of the matrix (2) corresponds to the square sum of the perpendicular distance between the estimated plane and the measured points. Due to this fact, the solution of the problem by orthogonal regression represents the best estimation.

3.2.2 Determination of a seed region The first step of the region growing is to find suitable seed regions. To this end an image containing root mean square (RMS) values is calculated. A filter mask is shifted over the xyz-image and all values within the mask are used to estimate a local plane. Afterwards the distance from each point within the mask to the estimated plane and the corresponding RMS value is calculated. As an alternative the minimum eigenvalue can be used. However, the error value is stored in a separate layer or image with the same dimension as the scan image at the position of the initial value of the filter mask. Small error values indicate that all local points are fitting to the estimated plane. Such positions are suitable as a seed region for region growing.

This method only requires a filter mask size as parameter. If the terrestrial laser scan is very dense, it is not necessary to calculate an error value for each measured point. In this case even small surfaces on a scanned object are covered by hundreds of scan points. For example, a facade of a building is scanned with an averaged resolution of 5 cm point density at the object. Then also small surfaces are detected as a seed, even if not every measurement value is taken into account.

3.2.3 Region Growing The region growing is a segmentation process, the scan points are assigned to planar regions. Therefore again the advantage of the regular raster is used. In practice, the initial seed pixels are given by the generated RMS-image, the smallest value is used at first. Now a plane is estimated using pixels around the seed region. The corresponding values are read out from the different layers within the given area and the coordinates are used to define the initial plane. The region grows by adding neighbouring pixels to the seed that fit to the plane. Thereby the distance between plane and 3D point is checked against a threshold that defines the maximum distance. After adding a certain number of points to a region, the plane parameters are recalculated using the already assigned 3D points. The region grows until no more points are added to the plane. Then the next seed region is selected and a new region is created. These steps are repeated until all points are assigned to a region, all possible seed regions have been used or a predefined maximum number of regions have been created. Optionally the created regions can be filtered by several criteria. For example small regions, where only few points have been assigned to, can be removed or the planes can be classified according to their normal vector.

3.3 Determination of the transformation parameters

After extracting planes from the scan data the next step in the registration process is to calculate the transformation parameters of a rigid motion between two different scan positions. *Grimson* (Grimson, 1990) and also *Jiang and Bunke* (Jiang and Bunke, 1997) describe the determination of the transformation parameters separated in rotation and translation. The complete transformation from a scan position S_2 into a reference scan position S_1 is given in homogenous coordinates by a 4 x 4 matrix:

$$\mathbf{T_{S_2-S_1}} = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \tag{7}$$

The 3 x 3 sub-matrix R contains the rotation and the 3 x 1 vector T the translation parameters. By using also homogenous coordinates for the measured points in a laser scan, the computation of transformed points is a single matrix multiplication.

3.3.1 Rotation The determination of the rotation matrix can be done by vector operations. Two pairs of corresponding planes are sufficient to determine the rotation. If there are several pairs of corresponding features the rotation may be calculated for all possible combinations and the median of the parameters can be computed. Alternatively the sum of the absolute angular deviation can be minimized. Any rotation can be expressed by a rotation axis and an angle of rotation about this axis. The direction of the axis is defined by:

$$r_{ij} = \frac{(n_i^{S_2} - n_i^{S_1}) \times (n_j^{S_2} - n_j^{S_1})}{\mid (n_i^{S_2} - n_i^{S_1}) \times (n_j^{S_2} - n_j^{S_1}) \mid} \tag{8}$$

$n_i^{S_1}$ and $n_i^{S_2}$ are the normalized vectors of corresponding planes in the scans S_1 and S_2. Two corresponding pairs of planes represented by the indices i and j are necessary for the determination of the rotation axis r_{ij}. The vector r_{ij} is orthogonal to $n_i^{S_2} - n_i^{S_1}$ and $n_j^{S_2} - n_j^{S_1}$.

If the rotation axis r is known, the rotation angle θ can be determined from the corresponding pair $(n_i^{S_1}, n_i^{S_2})$ by using the relationship:

$$n_i^{S_1} = \cos\theta n_i^{S_2} + (1-\cos\theta)(r \cdot n_i^{S_2})r + \sin\theta(r \times n_i^{S_2}) \tag{9}$$

Algebraic manipulation yields:

$$\sin\theta = \frac{(r \times n_i^{S_2}) \cdot n_i^{S_1}}{1 - (r \cdot n_i^{S_1})(r \cdot n_i^{S_2})} \tag{10}$$

$$\cos\theta = \frac{n_i^{S_2} \cdot n_i^{S_1} - (r \cdot n_i^{S_1})(r \cdot n_i^{S_2})}{1 - (r \cdot n_i^{S_1})(r \cdot n_i^{S_2})} \tag{11}$$

From this, the angle θ can be obtained. Using the angle θ and the rotation axix $r = (r_x, r_y, r_z)$, the rotation matrix R is defined by:

$$\mathbf{R} = \cos\theta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (1-\cos\theta) \begin{pmatrix} r_x^2 & r_x r_y & r_x r_z \\ r_y r_x & r_y^2 & r_y r_z \\ r_z r_x & r_x r_y & r_z^2 \end{pmatrix} + \sin\theta \begin{pmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{pmatrix} \tag{12}$$

3.3.2 Translation In case of determining the translation vector by using planes at least three corresponding pairs are required, because of restrictions to the position of planes in space. If the planes are corresponding, the normal vectors of the planes are approximately equal. The difference between the planes is the translation, the plane is shifted by the values Δx, Δy and Δz. The equations of corresponding planes can be written as:

$$a(x - \Delta x) + b(y - \Delta y) + c(z - \Delta z) + d^{S_2} = 0 \quad (13)$$

$$ax + by + cz + d^{S_1} = 0 \quad (14)$$

If the equations are equated one obtains:

$$a\Delta x + b\Delta y + c\Delta z = d^{S_2} - d^{S_1} \quad (15)$$

Generally, in matrix notation equation (15) can be written as:

$$\begin{pmatrix} d_i^{S-2} - d_i^{S1} \end{pmatrix} = \begin{pmatrix} a_i & b_i & c_i \end{pmatrix} \cdot \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \quad (16)$$

Each pair of corresponding planes yields one equation. A least squares solution is used to calculate the translation vector $\hat{x}$:

$$l + v = A \cdot \hat{x} \quad (17)$$

$$\hat{x} = (A^T A)^{-1} A^T l \quad (18)$$

4 EXAMPLE AND RESULTS

The method proposed in the previous chapter was applied to scan data gathered with a Riegl LMS 360i scanner. In a first test, a demo data set was recorded. A corner in a room was selected and scanned from two different scan positions. A corner provides three planar surfaces, which are perpendicular and thus provide enough information to determine the rotation as well as the translation parameters. The transformation parameters have also been determined using traditional methods. Retro-reflective targets were distributed, identified in the scans and the transformation matrix was calculated.

Figure 5 shows the measurement setup. The scan positions are visualized by the scanners, the two scans are displayed in different colours and are already registered. The calculated transformation matrix can be treated as reference. The matrix is structured as described in equation (7) and reads as follows:

$$\mathbf{T} = \begin{pmatrix} 0.4590 & -0.8880 & -0.0297 & 3.5393 \\ 0.8879 & 0.4596 & -0.0210 & -1.9763 \\ 0.0323 & -0.0167 & 0.9993 & -0.5283 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (19)$$

For the registration process without the targets, first planar surfaces are extracted from the scan data. The distance threshold (see chapter 3.2.3) for the region growing has to be selected greater than the accuracy of the scanner. In this case the noise of the scanner does not affect the segmentation. In the example the threshold is selected to 2 cm. The region growing process results in three extracted planes for each scan, if the small regions are neglected. The next step is to assign corresponding planes. At first for the test data set this is done manually.

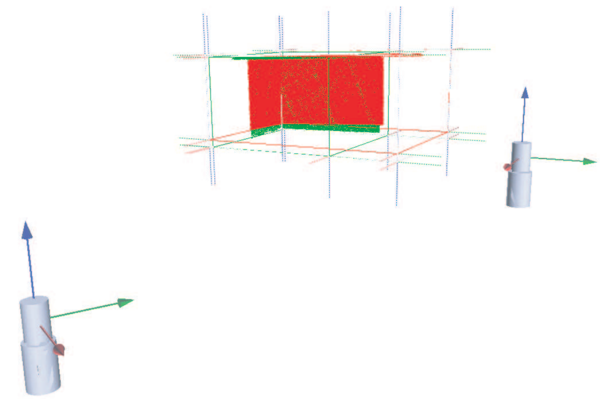

Figure 5: Registered scans and scan positions

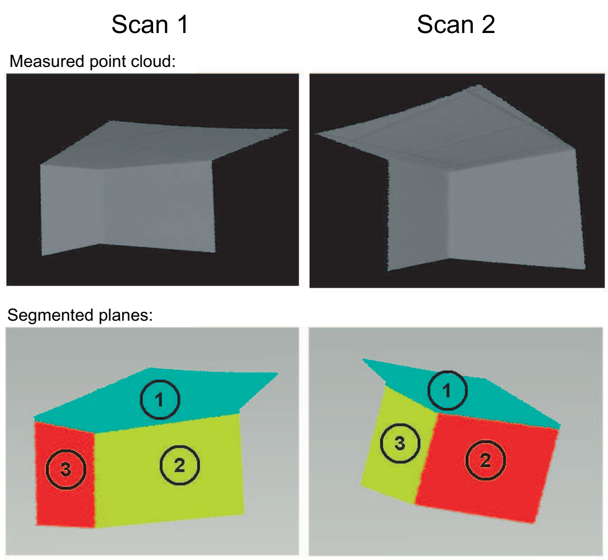

Figure 6: Test data

Figure 6 illustrates both, scans and the result of segmenting the points to planes. The numbers in the figure indicate the corresponding planes. The plane parameters were calculated using all segmented points by the method described in chapter 3.2.1 and are estimated to:

The plane parameters in table 1 are used to compute the rotation and the translation between the two scan positions. For the rotation matrix three different combinations with pairs of corresponding planes are possible, whereas all the planes are required to calculate the translation component. The method described in section 3.3 applied to all combinations (2 and 1, 3 and 1, 2 and 3) yields an averaged transformation matrix including rotation and translation of:

Scan / Plane	a	b	c	d
1 / 1	-0.0302	-0.0162	0.9994	-0.8710
1 / 2	0.9993	0.0169	0.0342	2.8249
1 / 3	0.0135	-0.9998	-0.0122	-3.9721
2 / 1	0.0082	0.0043	0.9999	-1.4600
2 / 2	0.4721	-0.8815	0.0071	6.3114
2 / 3	-0.8835	-0.4683	0.0098	-1.9604

Table 1: Plane parameters

$$\mathbf{T} = \begin{pmatrix} 0.4562 & -0.8895 & -0.0273 & 3.5397 \\ 0.8893 & 0.4568 & -0.0215 & -1.9579 \\ 0.0316 & -0.0145 & 0.9994 & -0.5140 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{20}$$

In comparison to the reference values determined with reflector targets the results show differences in the rotation as well as in the translation component. In order to evaluate these results, both transformation matrices are applied to a measured set of points and the differences in all coordinate axes are calculated. The outcome of this is an average shift of the points about:

$$\begin{aligned} \Delta X &= 0.013\, m \\ \Delta Y &= 0.023\, m \\ \Delta Z &= 0.009\, m \end{aligned} \tag{21}$$

This result shows, that the described method is suitable to determine the transformation parameters between two overlapping terrestrial lasers scans. The accuracy of the calculated parameters is sufficient to achieve initial values for a fine adjustment afterwards. Furthermore it is expected, that the accuracy of the transformation parameters will be improved by using more corresponding surfaces. Scans of building facades contain many planar surfaces which contribute to the accuracy of the determined transformation parameters.

5 SUMMARY AND OUTLOOK

In this paper an approach has been described to register terrestrial laser scans without using special targets as identical points to achieve the transformation parameters. A segmentation is used to derive meaningful planar regions in each scan. The parameters of the planar surfaces are determined by a robust estimation algorithm. Afterwards at least three corresponding pairs of planar patches are selected and the transformation parameters are computed separated in rotation and translation. A complete example is given and the results are compared with reference values achieved by traditional methods.

In the future, it is planned to automate the matching procedure of planar surfaces. A constrained tree search will be used to find corresponding regions in different scans, validating them using geometric properties. Finally, the fine adjustment of the different scans will be done using 3D correspondences.

REFERENCES

Adobe, 1992. TIFF - Revision 6.0. Adobe Systems Incorporated, 1585 Charleston Road, P.O.Box 7900, Mountain View, CA 94039-7900. http://www.adobe.com/Support/TechNotes.htm.

Böhler, W., Bordas Vicent, M. and Marbs, A., 2003. Investigating laser scanner accuracy. In: IAPRS, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 5/C15, Antalya, pp. 696–701.

Brenner, C., 2000. Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen. PhD thesis, Universität Stuttgart, Institut für Photogrammetrie, Deutsche Geodätische Kommission, C 530.

CyberCity AG, 2004. http://www.cybercity.tv/ (accessed on 27.04.2004).

Drixler, E., 1993. Analyse der Form und Lage von Objekten im Raum. Vol. Reihe C, Heft Nr. 409, Deutsche Geodätische Kommision, München.

Duda, R. O. and Hart, P. E., 1973. Pattern Classification and Scene Analysis. John Wiley and Sons, New York.

Grimson, W. E. L., 1990. Object Recognition by Computer. The MIT Press.

Haala, N., 1996. Gebäuderekonstruktion durch Kombination von Bild- und Höhendaten. PhD thesis, Universität Stuttgart, Institut für Photogrammetrie, Deutsche Geodätische Kommission, C 460.

Jiang, X. and Bunke, H., 1997. Gewinnung und Analyse von Tiefenbildern. Springer-Verlag Berlin Heidelberg.

Kampmann, Georg und Renner, B., 2004. Vergleich verschiedener Methoden zur Bestimmung ausgleichender Ebenen und Geraden. Allgemeine Vermessungsnachrichten 2/2004, pp. 56–67.

Kolbe, T. H. and Gröger, G., 2003. Towards unified 3D city models. In: Proceedings of the ISPRS Comm. IV Joint Workshop on Challenges in Geospatial Analysis, Integration and Visualization II in Stuttgart.

Phoenics GmbH, 2004. http://www.phoenics.de (accessed on 28.04.2004).

ACKNOWLEDGEMENT

The presented work has been done within in the scope of the junior research group “Automatic methods for the fusion, reduction and consistent combination of complex, heterogeneous geoinformation”. The project is funded by the VolkswagenStiftung.