AUTOMATIC RELATIVE ORIENTATION OF TERRESTRIAL LASER SCANS USING PLANAR STRUCTURES AND ANGLE CONSTRAINTS

Claus Brenner and Christoph Dold

Institute of Cartography and Geoinformatics
Leibniz Universität Hannover
Appelstr. 9A, 30167 Hannover, Germany
claus.brenner@ikg.uni-hannover.de, christoph.dold@ikg.uni-hannover.de

KEY WORDS: Relative orientation, registration, range images, terrestrial laser scanning

ABSTRACT:

The relative orientation of independently acquired terrestrial laser scan point clouds is an important task. If good starting values are available, well-known iterative algorithms exist to determine the required transformation. In this paper, we describe a method to obtain such starting values fully automatically, which is applicable to scenes containing planar elements. Our method first extracts planar patches in each scan individually and then assigns patch triples across scans in order to compute the rotation and translation component of the relative orientation. We assess the performance of our approach using a set of 20 terrestrial scans acquired systematically at increasing distance. For each scan, we automatically extract the 50 largest planar patches. We show that, although there are 1.15 billion possible patch triple assignments, we are able to compute efficiently a ranked list of possible transformations where the correct transformation is usually within the first few positions. For our test data and three test runs, it has been among the first 53 positions, even for scans with little overlap. Thus, instead of 1.15 billion candidate solutions, the score function needs only to evaluate on the order of 100 candidate solutions, which is an improvement by a factor of 10^7.

1 INTRODUCTION

In terrestrial laser scanning, an important problem is to find the relative orientation of independently acquired datasets, also called range image registration. This is a very well-known problem dating back to the first investigations on range images. It can be divided into two subproblems, coarse registration, which assumes no previous knowledge about the relative orientation of the two scans, and fine registration, where the assumption is that an initial orientation is known and the goal is to refine this in order to find the most accurate transformation parameters. Fine registration can be achieved using iterative techniques, usually based on the *iterative closest point* (ICP) approach. There is extensive literature on this subject. Originally described by Chen and Medioni (1991) and Besl and McKay (1992), many variants were proposed in the sequel, differing in the selection, matching, weighting and rejection of correspondences, e.g. (Zhang, 1994; Kapoutsis et al., 1999; Greenspan and Godin, 2001; Jost and Hügli, 2002; Sharp et al., 2002). An overview is given by Rusinkiewicz and Levoy (2001) and Gruen and Akca (2005). The ICP algorithm is nowadays also widely available in commercial software.

Any relative orientation based on the data itself requires two steps, (i) finding corresponding features in both datasets, and (ii) determination of the relative orientation which aligns those features. Iterative schemes like the ICP solve the correspondence problem by assuming that, applying the known coarse transformation, any point in the first scene is already close to his counterpart in the second scene. This allows to define corresponding features solely based on vicinity, with no or only limited interpretation of the scenes.

As for the coarse registration, finding the relative orientation of two overlapping scans without previous knowledge of the transformation is a hard (and mainly combinatorial) problem. For practical purposes, it is often solved in software by letting the user define a number of corresponding point pairs manually, which allows to compute the 3D Euclidean transformation. Automation of this step is not only interesting in terms of improvement of laser scan software. It also is related to fundamental problems such as object recognition (where one of the scans is replaced by a known model) and the problem of the 'kidnapped robot' in robotics (where the robot has to find its initial pose by determination of the relative orientation of its scan data and a known map).

Establishing correspondences between datasets without any previous knowledge requires features 'stronger' than points. Features should be stable with respect to partial occlusion, and should carry enough information to recover position and orientation (Faugeras and Hebert, 1986). In this paper, we investigate a coarse registration technique using correspondences of planar patches. We chose this feature since planar faces are often present in the vicinity of man-made structures. Furthermore, planar patches are relatively easy to extract from laser scanner data. We extend our previous work on that topic (Brenner et al., 2007) by an improved method to find patch correspondences.

This paper is organized as follows. In section 2, we present the mathematical background, in section 3 the basic problem and our approach are stated, and section 4 introduces our test data. Then, section 5 and 6 introduce and evaluate our solution for the determination of the rotation and the translation, respectively. Finally, section 7 draws conclusions and gives an outlook.

2 MATHEMATICAL FORMULATION OF THE PROBLEM

This section is based on the notation used in (Brenner et al., 2007), briefly repeated here to keep the paper self-contained. Two scenes (point clouds) S_1 and S_2 are given, each consisting of a set of points in 3D space. Any two corresponding points $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{R}^3$ with $\mathbf{x}_1 \in S_1$, $\mathbf{x}_2 \in S_2$, are related by an Euclidean (rigid) transformation

$$\mathbf{x}_1 = \mathbf{R}\mathbf{x}_2 + \mathbf{t}, \tag{1}$$

where $\mathbf{R}$ is a 3×3 rotation matrix, and $\mathbf{t} \in \mathbb{R}^3$ is the translation vector. Usually, due to errors, the transformed point of $\mathbf{x}_2$, denoted as $\mathbf{x}'_2$ (i.e., $\mathbf{x}'_2 = \mathbf{R}\mathbf{x}_2 + \mathbf{t}$), and its counterpart $\mathbf{x}_1$ from S_1, do not exactly coincide. Then, the transformation parameters for $\mathbf{R}$ and $\mathbf{t}$ can e.g. be found by (least-squares) minimization of $\sum \|\mathbf{x}_1 - \mathbf{x}'_2\|^2$. Given three or more point correspondences, closed form solutions exist to compute $\mathbf{R}$ and $\mathbf{t}$ (Sansò, 1973; Horn, 1987).

If no previous information is available, point correspondences cannot be established easily, since single points do not carry enough information. One way to solve this problem is to define descriptors (Johnson and Hebert, 1999). In contrast, we use a feature based approach which relies on planar patches. We assume the patches are given by their plane equations

$$\begin{aligned} \langle \mathbf{n}_i, \mathbf{x} \rangle - d_i &= 0 && (2) \\ \langle \mathbf{m}_i, \mathbf{x} \rangle - e_i &= 0 && (3) \\ \langle \mathbf{p}_i, \mathbf{x} \rangle - f_i &= 0 && (4) \end{aligned}$$

where $\mathbf{n}_i$, $\mathbf{m}_i$, $\mathbf{p}_i$ are normal vectors of unit length, d_i, e_i, f_i are the plane distances from the origin, and for each of the equations, $i = 1$ (plane in scene S_1) and $i = 2$ (plane in scene S_2) form a pair.

Three such plane pairs suffice to determine all six degrees of freedom of $\mathbf{R}$ and $\mathbf{t}$ in two steps. First, $\mathbf{R}$ can be found in closed-form by eigenvector analysis (actually part of the solutions in (Sansò, 1973; Horn, 1987)). Then, assume that scene S_2 has already been rotated, so that only the translation component $\mathbf{t}$ in Eq. 1 has to be determined. From Eq. 2,

$$\begin{aligned} \langle \mathbf{n}_1, \mathbf{x} \rangle - d_1 &= 0 \\ \langle \mathbf{n}'_2, \mathbf{x} - \mathbf{t} \rangle - d_2 &= 0. \end{aligned}$$

Since $\mathbf{n}'_2$ is already rotated, $\mathbf{n}_1 = \mathbf{n}'_2 = \mathbf{n}$, and $\mathbf{x}$ can be eliminated to obtain $\langle \mathbf{n}, \mathbf{t} \rangle = d_1 - d_2$. Doing the same for Eqs. 3 and 4 and stacking the equations yields

$$\begin{bmatrix} \mathbf{n}^\mathbf{T} \\ \mathbf{m}^\mathbf{T} \\ \mathbf{p}^\mathbf{T} \end{bmatrix} \mathbf{t} = \begin{bmatrix} d_1 - d_2 \\ e_1 - e_2 \\ f_1 - f_2 \end{bmatrix} \quad (5)$$

from which $\mathbf{t}$ can be determined.

Note that the determination of the full transformation is done in two steps, first the rotation, then the translation. While at least three plane pairs are required to obtain the translation, only two plane pairs are sufficient to determine the rotation. This will be exploited below to reduce search space. In fact, a plane normal vector (of unit length) has two degrees of freedom, so that two plane pairs fix four degrees of freedom, one more than what is required to determine $\mathbf{R}$. As a result, given two corresponding normal vector pairs $\mathbf{n}_1$, $\mathbf{m}_1$ from S_1 and $\mathbf{n}_2$, $\mathbf{m}_2$ from S_2, due to measurement errors, the angle $\angle(\mathbf{n}_1, \mathbf{m}_1)$ and $\angle(\mathbf{n}_2, \mathbf{m}_2)$ are usually slightly different. Then, one can choose to determine $\mathbf{R}$ such that either $\mathbf{n}_1$ and $\mathbf{n}_2$ or $\mathbf{m}_1$ and $\mathbf{m}_2$ align perfectly. Using the eigenvector solution mentioned above, a preferable rotation $\mathbf{R}$ is found, which distributes the angle error equally to both corresponding vectors.

Noting that the determination of the rotation is a time-critical operation, the following alternative can be used, which achieves the same result without the need for an eigenvector analysis (based on (Horn, 1987)). Using $\mathbf{n}_1$ and $\mathbf{m}_1$, a Cartesian coordinate frame $\{\mathbf{u}_1, \mathbf{v}_1, \mathbf{w}_1\}$ is constructed by

$$\tilde{\mathbf{u}}_1 = \mathbf{n}_1 + \mathbf{m}_1, \ \mathbf{u}_1 = \tilde{\mathbf{u}}_1 / \|\tilde{\mathbf{u}}_1\| \quad (6)$$

$$\begin{aligned} \tilde{\mathbf{v}}_1 &= \mathbf{m}_1 - \langle \mathbf{m}_1, \mathbf{u}_1 \rangle \mathbf{u}_1, \ \mathbf{v}_1 = \tilde{\mathbf{v}}_1 / \|\tilde{\mathbf{v}}_1\| && (7) \\ \mathbf{w}_1 &= \mathbf{u}_1 \times \mathbf{v}_1, \end{aligned}$$

where Eq. 7 uses standard Gram-Schmidt orthonormalization. Due to Eqs. 6 and 7, $\mathbf{u}_1$ and $\mathbf{v}_1$ span the same plane as $\mathbf{n}_1$ and $\mathbf{m}_1$. Then, $\mathbf{M}_1 = [\mathbf{u}_1 \mathbf{v}_1 \mathbf{w}_1]$, writing $\mathbf{u}_1$, $\mathbf{v}_1$, $\mathbf{w}_1$ as column vectors, is an orthogonal matrix by construction. Doing the same for $\mathbf{M}_2$, one can see that

$$\mathbf{R} = \mathbf{M}_1 \mathbf{M}_2^\mathbf{T} \quad (8)$$

is orthogonal and in fact is the desired rotation matrix (since $\mathbf{M}_2^\mathbf{T} \mathbf{n}_2$ gives the components of $\mathbf{n}_2$ along the axes $\{\mathbf{u}_2, \mathbf{v}_2, \mathbf{w}_2\}$ and $\mathbf{M}_1$ maps this back to the first coordinate frame). Adding $\mathbf{n}_1$ and $\mathbf{m}_1$ in Eq. 6 ensures that the angle error is equally distributed to both corresponding vectors.

3 FUNDAMENTAL PROBLEMS AND APPROACH OF THIS PAPER

The foremost problem of coarse registration is the combinatorial complexity. If p plane patches are extracted in S_1 and S_2 independently and then all possible transformations are evaluated based on plane triples ($k = 3$), as described above, there are

$$\binom{p}{3} \cdot \binom{p}{3} \cdot 3!/2 \quad (9)$$

possible combinations. The first two terms are due to picking three planes (the triple) out of p, while the last factor reflects the possible permutations when assigning the triple from S_1 to S_2, reduced by a factor of two, since only triples of the same chirality need to be considered (i.e., a right-handed normal vector triple from S_1 can only match a triple in S_2 which is also right-handed). For $p = 50$ planes, which we use regularly, this yields 1.15 billion possible combinations which need to be tested.

Noting the positive effect of chirality in Eq. 9 (reduction by a factor of two), one may wonder if picking more planes may have a positive effect. If $k = 4$ planes are picked, the chirality can be computed for any sub-combination of 3 planes picked out of those four. That is, for $k = 4$ planes, four 'chirality numbers' ± 1 are obtained. Any pick of $k = 4$ planes in S_1 is thus one case in the set $\{(+1,+1,+1,+1), (+1,+1,+1,-1), \ldots, (-1,-1,-1,-1)\}$ (all of which may occur). Instead of all $4! = 24$ permutations of a plane quadruple picked from S_2, only those with the same four chirality numbers need to be considered. Depending on the actual sign combination, either 3 (8 cases), 4 (6 cases) or 12 (2 cases) permutations need to be considered, which yields an expectation of 1.5 cases on average (which is also obtained from $6!/2^4$). Thus, comparing the cases $k = 3$ and $k = 4$, one sees that $k = 4$ has an advantage only if the number of planes is relatively small ($p < 9$), in which case the computational cost is anyhow so low that one would not consider using the more complex approach. In summary, increasing k does not reduce the number of cases (for practical p), even if chirality is considered.

The second important problem is the rating of a solution. Ideally, a score function would be desirable which attains its maximum when the correct solution is found. If exhaustive search would be possible, the best solution would then be obtained by simply picking the transformation with the highest score. A candidate for this score function is the overlap of S_1 with the transformed S_2, for example based on counting the points in S_1 with close neighbors in S_2. While this works well when the scene contents of S_1 and S_2 are similar (e.g., scan positions are close together), it usually fails when they are very different (e.g., scan positions

far apart, occlusions, tilted scan). In the latter case, the score of the true transformation is low, and it may well be that a larger score can be achieved by using a wrong transformation.

Using additional criteria (such as point normals) to make the score function more selective is possible, however comes at an additional computational cost. While it is practicable to compute the score for hundreds of cases, it is usually not feasible to do so for 1.15 billion cases. Thus, the main idea is to build up a hierarchy of tests which cuts down search space and has the property that (i) the most inexpensive tests are applied first, (ii) the more expensive tests are only applied after a large number of false solutions has been ruled out already, and (iii) the tests, though simple, do not erroneously rule out the correct solution.

The goal of this paper is not to elaborate on the score function, but on this test hierarchy. Thus, we do not show that our algorithm finds *and indicates* the correct transformation (which requires a search *and* a score function which has a maximum at the correct transformation). Instead, we show that we are able to reduce the set of solution candidates substantially, while still retaining the correct solution in this set.

4 THE TEST DATA SET AND INITIAL PROCESSING

We selected an area called 'Holzmarkt' in the historic district of Hannover, Germany, for the evaluation of our algorithms. Twenty scans were acquired, of which 12 were taken (approximately) upright, another 8 with a tilted scan head. Throughout the text, the scan positions and datasets are denoted by 'SP01', 'SP02', etc. for the upright and 'SP03a', 'SP05a', etc. for the tilted scans. Fig. 1 shows all 12 scan locations in a cadastral map. The scan positions were chosen systematically along a trajectory with a spacing of approximately 5 meters. All scans were acquired using a Riegl LMS-Z360I scanner, which has a single shot measurement accuracy of 12 mm, field of view of 360°×90° and a range of about 200 m. Reference orientations for the scans were obtained by placing artificial targets in the scene, which were manually identified in the scans. The procedure yields errors in the range of a few millimeters, thus the reference is considered to be sufficiently accurate for our tests on coarse registration. We used the reference orientations to compute an approximate value for the overlap of scan pairs, ranging from 83.1% for scan pair SP01-02 down to 2.3% for SP01-12a, see (Brenner et al., 2007).

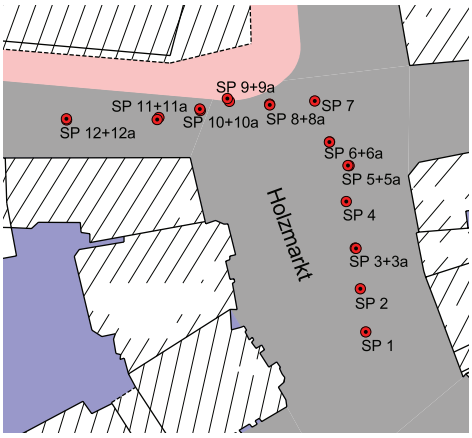

Figure 1: Placement of scan positions along a trajectory, shown in a cadastral map. Tilted scans are marked with an 'a' suffix.

For the extraction of planar patches, we used standard region growing, working on the regular raster of scan points. Region growing iterates the two steps of seed region selection and region expansion. Seed regions are prioritized according to their local planarity, which is computed using the residuals of a local best-fit plane. Once a seed region is selected, scan points along the region border are added if they lie in the plane (within a threshold of 6 cm), and the plane equation is updated. Fig. 2 shows an example segmentation.

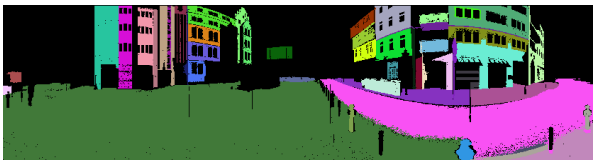

Figure 2: Planar segmentation of SP01, using random colors for the segments.

5 DETERMINATION OF THE ROTATION COMPONENT

5.1 The triple product and pairwise enclosed angles

For our test scene, we exhaustively computed all 1.15 billion plane triple combinations and the resulting transformations (this took several hours on a standard PC for each scan pair). Transformations were considered to be correct if the deviation from the reference is less than 5° in rotation and 1 m in translation. From table 1, one can see that at most, 0.212‰ of the triple combinations lead to a correct transformation, and this number even decreases rapidly with increasing distance between the scan standpoints.

	Triple assignments leading to correct transformation		Triples with compatible angles	Triples with compatible angles leading to correct transformation	
	#	‰	#	#	‰
SP 01-02	244635	0,212	1022507	42945	42,00
SP 01-03	208970	0,181	1020667	38947	38,16
SP 01-03a	153111	0,133	684729	20283	29,62
SP 01-04	147045	0,128	1091474	19043	17,45
SP 01-05	55116	0,048	698353	9681	13,86
SP 01-05a	41353	0,036	557906	4955	8,88
SP 01-06	48721	0,042	949832	8361	8,80
SP 01-06a	47843	0,042	1041477	8562	8,22
SP 01-07	14776	0,013	880668	3034	3,45
SP 01-08	15576	0,014	791156	2609	3,30
SP 01-08a	11372	0,010	840829	1048	1,25
SP 01-09	6306	0,005	605209	1125	1,86
SP 01-09a	11545	0,010	513071	778	1,52
SP 01-10	13372	0,012	754447	1357	1,80
SP 01-10a	4584	0,004	438870	596	1,36
SP 01-11	4232	0,004	758084	593	0,78
SP 01-11a	11160	0,010	653320	1572	2,41
SP 01-12	0	0,000	552271	0	0,00
SP 01-12a	0	0,000	402779	0	0,00

Table 1: Triple assignments leading to the correct transformation, angle compatible triple assignments, and angle compatible triple assignments leading to the correct transformation (for all scan pairs).

In order to raise this percentage, we used in (Brenner et al., 2007) the triple product to only consider plane triples above a threshold. A large triple product is desirable, since it leads to a good matrix condition number on the left hand side of Eq. 5. However, it is also problematic, since the appropriate value depends on the scene contents. If the scene does not contain planes leading to triple products above the threshold, no candidates are found. In this case, the threshold has to be lowered, which however quickly increases the number of false combinations as well.

In order to form a more selective and scene independent criterion, we investigated the use of the three angles enclosed by the three normal vectors instead of their triple product. To evaluate how accurate the angles between any two pairs of plane normal vectors

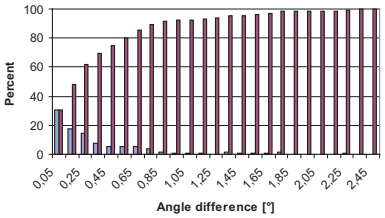

Figure 3: Histogram and cumulated histogram of the angle differences of manually selected plane pairs.

are, we manually identified a small set of corresponding planes between scans. For any possible plane pair in one scan S_1, we computed the angle between the plane normal vectors. Knowing the corresponding vectors in S_2, we computed the enclosed angle as well and derived the difference. In total, 328 pairs were considered. From Fig. 3, one can see that for more than 90% of the normal vector pairs from S_1, the corresponding pairs in S_2 form the same angle within a 1° tolerance. This leads to the conclusion that tight bounds can be imposed on the angles when searching for corresponding plane triples. Table 1 shows that out of the 1.15 billion triple combinations, only between 400,000 and 1 million compatible combinations remain. The rate of triple combinations which lead to correct transformations is as high as 42‰. Thus, for SP01-02, by using angle constraints, we can reduce the amount of search required by a factor of $42‰/0.212‰ \approx 200$. This is also the average factor over all scans.

5.2 Searching for the correct orientation

As noted in section 2, the rotation is fully determined by two normal vector pairs, using Eq. 8. Thus, only p over 2 pairs need to be picked, and (c.f. Eq. 9 for $k = 2$) a total of $p^2(p-1)^2/2$ plane pair combinations exist. For $p = 50$, this yields 1,225 pairs in each scan, and 3 million combinations. If only vector pairs including the same angle (tolerance 1°) are regarded, this reduces to 140,000 compatible combinations, or 4.8%, on average. From table 2 one can see that the number of compatible normal vector pairs is relatively stable. However, if the rotation matrix is computed for each of the compatible combinations and compared to the (known) reference orientation (allowing a 2° tolerance), one can see that the number of those pairs leading to a correct orientation decreases with increasing scan numbers, from 8,034 (5.5%) down to almost zero. Thus, even scans far apart yield a large number of compatible normal vector pairs, but the percentage leading to the correct transformation decreases. Note that there is no need to test the 3 million cases by exhaustive enumeration. Instead, all 1,225 angles between pairs in S_2 can be sorted into angle bins (we used 1° bins for this purpose). Then, for each plane combination in S_1, the subset of candidates in S_2 can be retrieved quickly.

In the next step, the goal is to pick a correct orientation from the approximately 140.000 candidates – or more precisely, to rank the candidates in such a way that the correct solution is among the first few proposals. Since the percentage of correct solutions can be around only 1% (for the cases we wish to be able to succeed), random picking would imply that we can expect only one correct solution among (the first) 100 picks.

In order to improve this rate, we computed the rotation matrix for all compatible combinations. Note that using Eq. 8, this does not require matrix inversion or eigenvalue analysis, so it is computationally inexpensive, even for 140.000 candidates. For each candidate rotation matrix, we recovered the three rotation angles ω, ϕ, κ. Fig. 4 shows a plot of all rotation candidates, in (ω, ϕ, κ) space, for the scan pair SP01-02. For the figure, the rotations were normalized using the known reference orientation, so that the correct rotation is at $(\omega, \phi, \kappa) = (0, 0, 0)$. At this point (center in Fig. 4), one can see a dense point cloud (according to table 2, 5.53% of the points should be located there). In order to test this, we sorted all candidate rotations (ω, ϕ, κ) into bins (using a bin size of 2°). After this, the bins are extracted highest count first. Similar (ω, ϕ, κ) values are merged during this step if they differ in all angles by less than 2° (this operation is similar to histogram smoothing considering neighboring cells).

Pair	Compatible	%	Correct	%
SP01-02	145202	4,84	8034	5,53
SP01-03	147944	4,93	7497	5,07
SP01-03a	115260	3,84	5566	4,83
SP01-04	164200	5,47	5852	3,56
SP01-05	145098	4,83	3496	2,41
SP01-05a	121400	4,04	2885	2,38
SP01-06	166238	5,54	4218	2,54
SP01-06a	165922	5,53	4414	2,66
SP01-07	173934	5,80	2513	1,44
SP01-08	167550	5,58	2639	1,58
SP01-08a	168050	5,60	2728	1,62
SP01-09	141868	4,73	1651	1,16
SP01-09a	140498	4,68	926	0,66
SP01-10	157464	5,25	2115	1,34
SP01-10a	113540	3,78	1007	0,89
SP01-11	138768	4,62	929	0,67
SP01-11a	147310	4,91	1642	1,11
SP01-12	105978	3,53	2	0,00
SP01-12a	94758	3,16	148	0,16

Table 2: Angle compatible normal vector pairs, percentage relative to total number of combinations (3 million), number of correct rotations computed from the pairs, and percentage relative to the compatible cases.

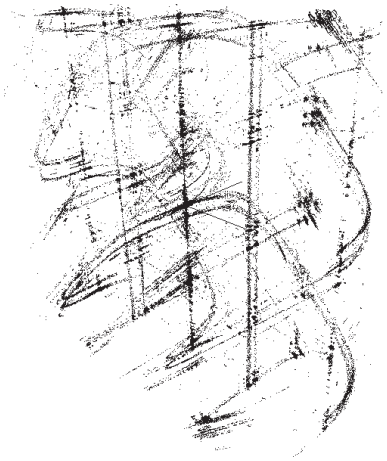

Figure 4: Plot of all rotation candidates for the scan pair SP01-02, in (ω, ϕ, κ) space. Each orientation is represented by a point. The correct orientation is at the center of the figure, where the ω and ϕ axes can be seen. The κ axis points upward.

As a result of this procedure, we obtain a list of orientations, sorted in descending order of bin hits. Fig. 5 shows the number of hits for the 20 bins with highest count, for the scan pair SP01-02 and SP01-09a. In the case SP01-02, the first bin (8,034 hits) has a much higher count as the second bin (1,752 hits). In fact, the first bin represents the correct orientation and the bin count is equal to the value in table 2. This situation is not always as clear. For example, in the case SP01-09a, the counts are generally lower and there is no clear peak at the first bin. In this case, the correct orientation corresponds to the 8th largest bin.

To give a better overview, Fig. 6 shows a plot of the 20 bins with

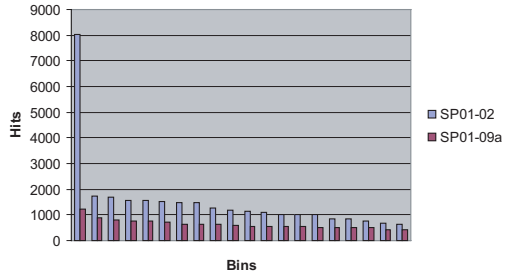

Figure 5: Histogram of the first 20 (orientation) bins with largest bin count for the scan pairs SP01-02 and SP01-09a.

highest count, for all scan combinations. As can be seen, for low scan numbers, there is a clear peak at bin 1, which is also the reference orientation. For SP01-04 and up, the peak gets wider, but still the correct orientation is at the first bin. The first exception to this is SP01-07 (which has 51% overlap), where the correct transformation is in the second bin (count 2,332). Closer examination reveals that the first bin (similar count of 2,362) represents a turn by κ=180° around the up- (Z-) axis with respect to the reference orientation. SP01-09a (29% overlap) is the first case where the correct orientation is not among the first two bins. SP01-11 is still worse, but note this pair has only 9.9% overlap. SP01-11a has 12.2% overlap and the correct solution is in bin 1. For SP01-12 and SP01-12a, the reference orientation was not part of the first 100 bins, however their overlap is only 4% and 2%, respectively.

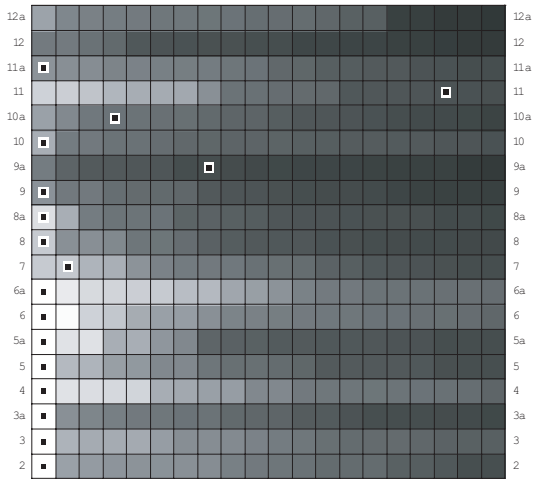

Figure 6: Bins with highest count for all scan combinations. White corresponds to a count of 3,000 or more, black is 0. The small rectangles indicate the bin which corresponds to the reference rotation. For example, the lowest line represents the first 20 bins for scan pair SP01-02 and is the equivalent of Figure 5. It has a clear peak (white) at the first (leftmost) bin, which also represents the true rotation (small rectangle).

6 DETERMINATION OF THE TRANSLATION COMPONENT

The translation is determined according to Eq. 5, using three plane pairs. Note that it is not necessary to actually rotate S_2, because Eq. 5 requires only d_2, e_2, f_2 from S_2, the plane distances from the origin, which are not affected by rotation. Also, instead of picking all triple pairs, one can work on the rotation candidates one after the other, so that not only the rotation matrix is known, but also a set of combinations of two plane pairs which led to this rotation (i.e., a quadruple of plane indexes). For example, for SP01-02, the first rotation considered corresponds to a bin with 8,034 hits, meaning that 8,034 cases of assigned plane pairs are known already. This compares favorably to the 140,000 compatible (and the 3 million total) pairs.

Both pairs, of S_1 and of S_2, need to be extended by a third plane, picked from the remaining $p - 2$ planes. For example, for the mentioned case, this would mean on the order of 8,034·48·48 = 18,510,336 possible picks. However, when imposing angle constraints (of 1°) for the angles between the already picked pair and the newly picked plane, and considering chirality, a much smaller number remains. In the example, only 188,732 picks are left.

However, we chose a conceptually simpler approach. Instead of picking a third plane, we simply pick pairs of quadruples from the bin. Thus, for each pick, we have 4 plane pairs, and solve Eq. 5 for the translation in a least squares manner. Fig. 7 shows the translations corresponding to 100,000 of such picks, where each translation vector is represented by a point in 3D space. The correct translation vector is at the center of the figure, where several 'linear structures' intersect. There are many candidates along the Z axis, indicating a correct lateral position, but a varying height. Perpendicular to this, there are several linear structures which we believe are due to the arrangement of the facades in the 'Holzmarkt' scene: if one moves the point cloud SP02 further apart from SP01, the distance between the right and left building facades increases and there are two choices for the translation, either matching the 'right' or the 'left' facades.

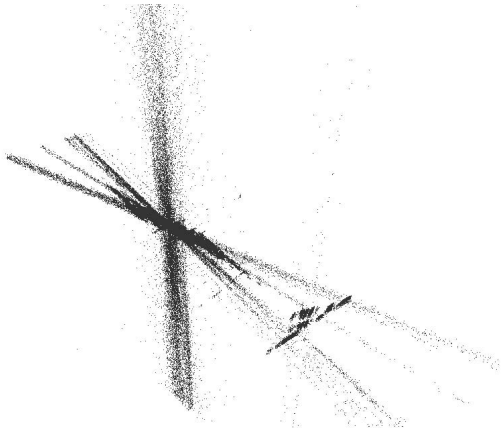

Figure 7: Plot of all translation candidates for the first orientation bin of the scan pair SP01-02. Z axis points upward.

Picking two quadruples from the bin yields 8,034·8,033/2 possible picks for the example bin (way too many). Instead, we apply the RANSAC principle at this point (Fischler and Bolles, 1981). We only pick a subset of m pairs of quadruples. For each pick, we compute the translation and then count the number of planes in S_1 for which a matching plane in S_2 exists. Planes were considered to match if their normal vectors agree within 1° and their distance from the origin agrees within 1 m. Note this comparison is computationally inexpensive, since it uses only the plane parameters, rather than original scan points.

To derive the necessary number of picks m, we picked 10,000 quadruple pairs and determined the percentage of picks which lead to the correct translation (within 1 m along each axis). We found that for close scan positions, such as SP01-02, this is around 20%, decreasing with increasing scan position distance, for a minimum of 3% (not considering SP01-12 and SP01-12a). Following Fischler and Bolles (1981), if we want to ensure

SP01-	02	03	03a	04	05	05a	06	06a	07	08	08a	09	09a	10	10a	11	11a	12	12a
Run 1	1	1	1	1	1	2	3	2	2	15	23	11	5	12	33	14	12	-	-
Run 2	1	1	1	1	1	2	6	5	12	4	20	53	6	35	15	13	12	-	-
Run 3	1	1	1	2	1	2	2	5	12	7	28	13	5	27	13	21	11	-	-

Table 3: Ranking of correct transformations. The value '1' in row 'Run 1' and column '02' means that for the scan pair SP01-02, and the first run, the first transformation returned by the algorithm also was the correct one.

with probability z to find at least one correct solution among m picks, where the probability to draw a correct solution is b, then $m = \log(1-z)/\log(1-b)$. For $z = 99\%$, $b = 3\%$, it follows that $m \approx 150$ picks are required.

The number of corresponding plane pairs is also used to rank the entire transformation (rotation and translation). Table 3 shows the results obtained for three separate runs of the algorithm. The rankings indicate at which position in the result list the algorithm returned a correct transformation (defined by as most 2° off in rotation and 1 m off in translation, for each axis). As one can see, for most of the close scan pairs, the algorithm returned the correct solution in the first place or within the first few ranks. For all runs except SP01-12 and SP01-12a, the solution was ranked among the first 53. For SP01-12 (overlap 3.6%) and SP01-12a (overlap 2.3%), we obtained no solution. However, for those cases, we were even unable to manually select suitable plane pairs.

7 CONCLUSIONS AND OUTLOOK

In this paper, we addressed the problem of finding good initial values for the relative orientation of two laser scans when no previous information is available. Our method is based on the automatic extraction and assignment of planar patches. For a set of terrestrial laser scans, with 50 extracted planar patches per scan, we showed that there is a large number of 1.15 billion possible assignments, however only 0.2‰ or less (one in 5000) of them lead to a correct transformation. Thus, it was our goal to devise an efficient method which cuts down search space and produces a ranked list of possible transformations, where the correct transformation is among the top entries. The general idea behind this is to built a hierarchy of tests, where the most elaborate test (the score function) needs only to be performed for very few cases.

We showed that the relative angles between patch normal vectors are a good (and scene independent) criterion to eliminate false assignments. For the determination of the rotation matrix, we started from the assignment of two patch pairs. Using a clustering of orientations by way of bins, we obtained a ranking, where the correct solution is at the top for the majority of scan pairs and ranked among the first 18 in all cases. As for the translation, we used a RANSAC based approach, where the sampling consists of picking two patch pairs, and the consensus set is the total number of compatible patch pairs. Overall, we obtained an efficient algorithm which computes a ranked list of transformation candidates, where the correct transformation is at rank one for scans with a high overlap, and ranked among the first 53 for all scan pairs with an overlap larger than 3.6%. We conclude that the number of candidates for which a more elaborate score function needs to be evaluated is on the order of 100, which is, compared to a total of 1.15 billion possible cases, a massive reduction by a factor of 10^7.

In the future, we plan to test the algorithm on other scenes as well, and to work on an efficient yet selective score function.

ACKNOWLEDGEMENTS

This work has been supported by the VolkswagenStiftung, Germany.

References

Besl, P. J. and McKay, N. D., 1992. A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2), pp. 239–256.

Brenner, C., Dold, C. and Ripperda, N., 2007. Coarse Orientation of Terrestrial Laser Scans. ISPRS Journal of Photogrammetry and Remote Sensing (in press, doi:10.1016/j.isprsjprs.2007.05.002).

Chen, Y. and Medioni, G., 1991. Object modeling by registration of multiple range images. In: International Conference on Robotics and Automation, pp. 2724–2729.

Faugeras, O. D. and Hebert, M., 1986. The representation, recognition, and locating of 3-D objects. International Journal of Robotics Research 5(3), pp. 27–52.

Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Comm. ACM 24(6), pp. 381–395.

Greenspan, M. and Godin, G., 2001. A nearest neighbor method for efficient ICP. In: Proceedings of the Third International Conference on 3D Digital Imaging and Modeling, Quebec City, Canada, pp. 161–168.

Gruen, A. and Akca, D., 2005. Least squares 3D surface and curve matching. ISPRS Journal of Photogrammetry and Remote Sensing 59(3), pp. 151–174.

Horn, B. K. P., 1987. Closed-form solution of absolute orientation using unit quaternions. Optical Society of America 4(4), pp. 629–642.

Johnson, A. E. and Hebert, M., 1999. Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Transactions on Pattern Analysis and Machine Intelligence 21(5), pp. 433–449.

Jost, T. and Hügli, H., 2002. A multi-resolution scheme ICP algorithm for fast shape representation. In: Proceedings of 1st International Symposium on 3D Data Processing, Visualization and Transmission, pp. 540–543.

Kapoutsis, C., Vavoulidis, C. and Pitas, I., 1999. Morphological iterative closest point algorithm. IEEE Transactions on Image Processing 8(11), pp. 1644–1646.

Rusinkiewicz, S. and Levoy, M., 2001. Efficient variants of the ICP algorithm. In: Proceedings of the Third Intl. Conf. on 3D Digital Imaging and Modeling, pp. 142–152.

Sansò, F., 1973. An exact solution of the roto-translation problem. Photogrammetria 29(6), pp. 203–216.

Sharp, G., Lee, S. and Wehe, D., 2002. ICP registration using invariant features. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24number 1, pp. 90–102.

Zhang, Z., 1994. Iterative point matching for registration of free-form curves and surfaces. International Journal of Computer Vision 13(2), pp. 119–152.