INDOOR NAVIGATION BY USING SEGMENTATION OF RANGE IMAGES OBTAINED BY LASER SCANNERS

B. Gorte [a], K. Khoshelham [a], E. Verbree [b]

a Optical and Laser Remote Sensing, Delft University of Technology, The Netherlands
b GIS Technology, Delft University of Technology, The Netherlands
{b.g.h.gorte, k.khoshelham, e.verbree}@tudelft.nl

Commission Ⅰ, ICWG-I/V

KEY WORDS: laser scanning, localization, mapping, navigation, robotics, segmentation, tracking

ABSTRACT:

This paper proposes a method to reconstruct the trajectory of a moving laser scanner on the basis of indoor scans made at different positions. The basic idea is that a position with respect to a measured scene is completely determined by the distances to a minimum of three intersecting planes within that scene, and therefore displacements are defined by changes in those distances between successive scans. A segmentation algorithm based on the gradients of a range image is used to extract planar features from the measured scene. The correspondence establishment between planes of two successive scans is formulated as a combinatorial problem, and a heuristic search method is employed to find combinations that indicate corresponding sets of planes between the two scans. The position of the scanner at each scan is computed by intersecting sets of four planes. An experiment with multiple scans obtained by a Faro laser scanner in an indoor environment is described, and the results show the reliability and accuracy of the proposed indoor navigation technique.

1. INTRODUCTION

Terrestrial laser scanning is a popular method for acquiring 3D data of the interior of buildings and closed environments such as tunnels. Laser range data can also be used for positioning and navigation through sensing features that have known coordinates. The combination of the above tasks, mapping and navigation, is known in the field of robotics as simultaneous localization and mapping (SLAM). The problem of simultaneous localization and mapping concerns the measurement of an environment using a sensor such as a laser scanner, while these measurements are processed to create a (3D) map of the environment and at the same time determine the position of the sensor with respect to the obtained map. Example applications of simultaneous 3D modelling and localization include autonomous vehicles, intelligent wheelchairs and automated data acquisition within indoor environments.

In recent years, simultaneous localization and mapping in both outdoor and indoor environments has been well researched by the scientists in the field of robotics and automation. In indoor environments, localization methods based on a 2D laser range finder have been more popular (Zhang and Ghosh, 2000; Xu et al., 2003; Victorino and Rives, 2004), but these methods provide only a 2D map of the environment. Other sensors that have been used for localization and mapping include sonar (Leonard and Durrant-Whyte, 1991) and cameras with structured light (Kim and Cho, 2006) as well as unstructured light (Diebel et al., 2004). Some of the existing approaches utilize multiple sensors and adopt a sensor fusion strategy. Dedieu et al., (2000) combine a camera and a 2D laser scanner. Newman et al., (2002) and Araneda et al., (2007) opt for the combination of a laser scanner and an odometer. Diosi and Kleeman (2004) fuse laser range data with data acquired from a sonar. Guivant et al., (2000) utilize a laser scanner with a number of beacons for the localization, although natural landmarks are also taken into account to improve the localization accuracy.

In this paper, we describe a method for simultaneous localization and 3D modelling of an indoor environment using a laser scanner. The proposed method is based on extracting planar features from the range images, and positioning the laser scanner with respect to corresponding planes in successive scans. The localization is, therefore, entirely based on the measured scene, and no beacons or known landmarks are required. In addition, the mapping is carried out in 3D since the laser scanner provides a 3D point cloud of the scanned scene.

The paper is structured in five sections. Section 2 describes the algorithm for the segmentation of the range images into planar features. In Section 3, the method for sensor localization using the extracted planar features is discussed. Experimental results of a test with multiple scans in an indoor environment are presented in Section 4. Conclusions appear in Section 5.

2. RANGE IMAGE SEGMENTATION AND EXTRACTION OF PLANAR FEATURES

The fundamental data structure of terrestrial laser scans is the range image (rather than the point cloud), since it is most closely linked to the scanner's operation, measurement of the range (distance between the scanner and a point) as function of an azimuth (horizontal) and a zenith (vertical) angle. The range image is a discrete representation of a point cloud in a spherical coordinate system. With the scanner at the centre of the sphere, the azimuth and zenith angles determine the columns and row

positions within the image, and the ranges are represented by the pixel values.

In a recent paper we introduced a range image segmentation algorithm (Gorte, 2007), which groups adjacent pixels obtained from co-planar 3D points into segments. The adjacency of pixels can be obtained from a range image, whereas co-planarity is derived from image gradients, taking the scan angles into account. The method is based on a parameterization of a plane with respect to the local coordinate system of the laser scanner, in which the scanner is at the origin (Fig. 1).

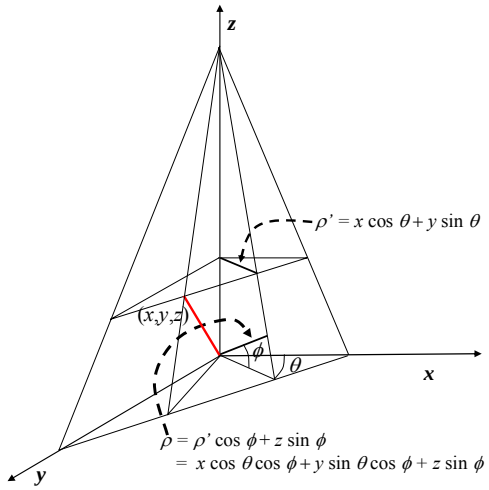

Fig. 1: Parametric form of a plane in 3D.

The segmentation algorithm attempts to group adjacent range image pixels into segments, as far as these pixels belong to the same plane. This is accomplished by estimating in each pixel the parameters of the normal vector of that plane. These parameters are: two angles θ (horizontal) and ϕ (vertical) and the length of the vector ρ. This is the perpendicular distance between the plane and the origin.

The algorithm consists of the following steps:

1. Computing gradient images *gx* and *gy* on the range image. These images denote at each row and column the change that occurs in the image value when moving one pixel horizontally and vertically respectively.
2. Computing the angle $\Delta\theta$ = atan ($gx/R\Delta\alpha$) between the horizontal laser beam angle α and the horizontal normal vector angle θ on the basis of the gradient in *x*-direction *gx*. $\Delta\alpha$ is the angular resolution of the scanner in horizontal direction. Now the first parameter of the normal vector of the plane, the horizontal angle θ, is known.
3. Computing the angle $\Delta\phi'$ = atan ($gy/R\Delta\beta$) on the basis of the gradient in *y*-direction *gy*. $\Delta\beta$ is the angular resolution of the scanner in vertical direction. This yields ϕ' (see Fig. 2). To obtain the second parameter of the normal vector, the vertical angle ϕ, a correction has to be applied given by:

$$\frac{\tan\phi}{\tan\phi'} = \frac{u}{u'} = \cos(\alpha-\theta)$$

The computation is illustrated in Fig. 3.

4. Computing the third parameter ρ of the normal vector using $\rho = x\cos\theta\cos\phi + y\sin\theta\cos\phi + z\sin\phi$ (see Fig. 1).
5. Image segmentation: On the basis of the three features from steps 2, 3 and 4, a quadtree based region-merging image segmentation (Gorte, 1998) is carried out to group adjacent pixels with similar feature values, *i.e.* pixels that are now expected to belong to the same plane in 3D, into segments.

The entire method consists of 2d image processing operations: gradient filtering, image arithmetic and image segmentation, which makes the algorithm extremely fast compared to point cloud segmentations working in 3D.

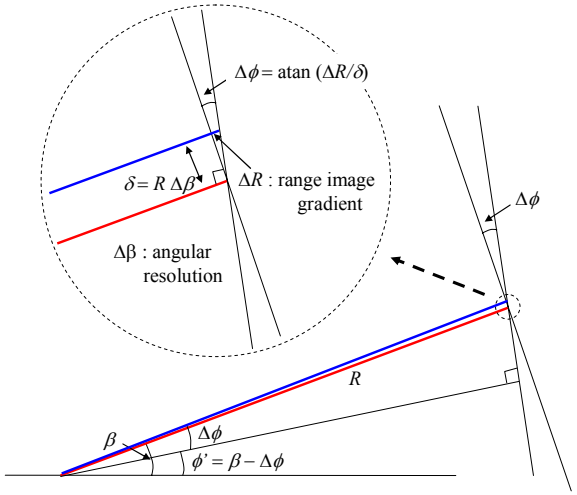

Fig. 2: The gradient determines the angle between the laser beam and the normal vector of a plane.

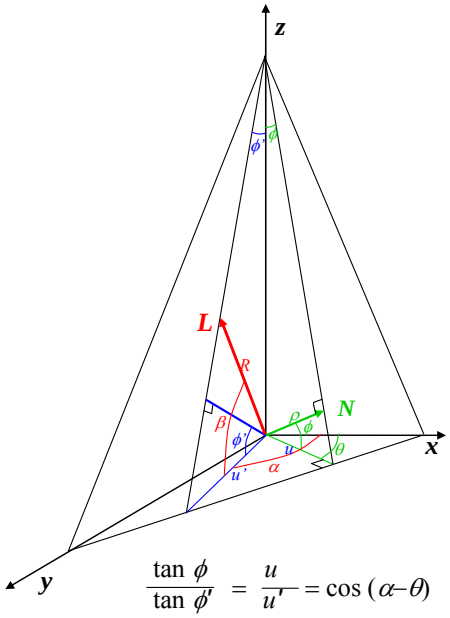

Fig. 3: A laser beam *L* with a direction given by α and β hitting a plane at range *R*, and the plane's normal vector *N* with parameters θ, ϕ and ρ.

The 3 parameters of the 3d planes corresponding to the segments, i.e. the two angles that define the direction of the normal vector of the plane, and the distance between the plane and the origin, are computed on-the-fly as the averages of values computed in the participating pixels. They are expressed in the same polar coordinate system that the range image is based on.

3. SENSOR LOCALIZATION USING EXTRACTED PLANES

The localization of the laser scanner is based on measuring distances to a minimum of three non-parallel planar surfaces (that intersect in non-parallel lines) in an already mapped indoor scene. The first scan maps the scene with respect to the scanner coordinate system; i.e., the parameters of the extracted planes are computed in the scanner coordinate system at its initial position. In the next scan, new ranges to these planes with known parameters are measured. The position of the scanner with respect to the previous coordinate system is then determined by intersecting the measured ranges.

The determination of the scanner position with respect to the previous position requires a minimum of three corresponding planes in the two successive scans. Generally a larger number of planes are extracted by the segmentation algorithm, and a search procedure for finding corresponding planes is needed. The following sections describe the plane intersection and the search procedure for corresponding planes.

1.1 Scanner position determination

The position of a point can be computed by measuring its distance to a minimum of three non-parallel planes with known parameters. Fig. 4 shows three non-parallel planar surfaces scanned from two positions S_1 and S_2. As can be seen, if planes P1, P2, and P3 are shifted by $-\boldsymbol{\rho}^{11}$, $-\boldsymbol{\rho}^{12}$, and $-\boldsymbol{\rho}^{13}$ respectively, they intersect at point S_2. The equation of a plane *i* in the coordinate system *j* of the scanner at position s_j is written as:

$$\rho^{ij} = n_1^{ij} x^{ij} + n_2^{ij} y^{ij} + n_3^{ij} z^{ij} \tag{1}$$

where ρ^{ij} is the distance of the plane *i* from the origin of the coordinate system *j*, $n_1^{ij}, n_2^{ij}, n_3^{ij}$ are the coordinates of the normal vector **n** of the plane *i* in the coordinate system *j*, and x^{ij}, y^{ij}, z^{ij} are the coordinates of a point **x** on the plane *i* in the coordinate system *j*. The short form of the equation can be written as:

$$\rho^{ij} = \mathbf{n}^{ij} \cdot \mathbf{x}^{ij} \tag{2}$$

where **.** denotes the dot product of the two vectors. To compute the position of the scanner at position S_2 we take the parameters of three non-parallel planes in the coordinate system of S_1, shift them according to their distances to S_2, and find the intersection. This can be expressed as:

$$\begin{cases} \rho^{11} - \rho^{12} = \mathbf{n}^{11} \cdot \mathbf{s}_2^1 \\ \rho^{21} - \rho^{22} = \mathbf{n}^{21} \cdot \mathbf{s}_2^1 \\ \rho^{31} - \rho^{32} = \mathbf{n}^{31} \cdot \mathbf{s}_2^1 \end{cases} \tag{3}$$

The solution of the set of equations given in (3) provides the coordinates of S_2 in the coordinate system of S_1. By repeating this procedure at each consecutive scan, the position of the scanner with respect to the coordinate system of the scanner at the previous position can be computed.

To transform the scanner positions to a single coordinate system, e.g. the coordinate system of the first scan, rotations between the coordinate systems of the scanners at successive scans should also be taken into account. If the positions of the scanners are required in a reference coordinate system, then the position and orientation of the scanner at the first position with respect to the reference coordinate system should be known. Assume that a third scan is performed, and the position of the scanner at S_3 is to be determined in a reference coordinate system O; we have:

$$\begin{cases} \rho^{12} - \rho^{13} = \mathbf{n}^{12} \cdot \mathbf{s}_3^2 = \mathbf{n}^{12} \cdot \mathbf{R}_{12}\mathbf{R}_{01}\mathbf{s}_3^0 \\ \rho^{22} - \rho^{23} = \mathbf{n}^{22} \cdot \mathbf{s}_3^2 = \mathbf{n}^{12} \cdot \mathbf{R}_{12}\mathbf{R}_{01}\mathbf{s}_3^0 \\ \rho^{32} - \rho^{33} = \mathbf{n}^{32} \cdot \mathbf{s}_3^2 = \mathbf{n}^{12} \cdot \mathbf{R}_{12}\mathbf{R}_{01}\mathbf{s}_3^0 \end{cases} \tag{4}$$

where $\mathbf{R}_{ij}$ denotes a 3D rotation from the coordinate system *i* to the coordinate system *j*. The rotation matrices can be derived from the parameters of corresponding planes in every pair of successive scans, except for $\mathbf{R}_{01}$, which is the orientation of the scanner at its initial position with respect to the reference coordinate system, and has to be measured in advance (if the navigation is to be computed in a reference coordinate system). In practice, the scanner can be levelled by using the bubble level and levelling screws. In such case, the z axis of the scanner will always remain in the upward direction. Consequently, the 3D rotation matrices will be simplified to rotations around z axis only, which can be easily derived from the differences in the parameter θ of the corresponding planes.

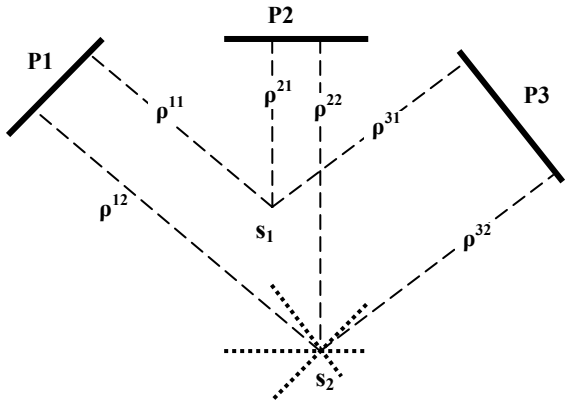

Fig. 4: Determining the position of S_2 by measuring distances to three planes.

1.2 Heuristic search for corresponding planes

The intersection of planar surfaces using Eq. 3 and Eq. 4 requires that a correspondence between the planes in every pair of successive scans is established. The correspondence is formulated as a combinatorial problem, and solutions to this problem are sought through a heuristic search. Suppose that the segmentation algorithm extracts n planes in the first scan and n' planes in the second scan of a pair of successive scans. The number of possible ways to choose r corresponding planes from the two sets of planes can be expressed as:

$$n_c = C_r^n \cdot P_r^{n'} = \frac{n!}{r!(n-r)!} \times \frac{n'!}{(n'-r)!} \quad (5)$$

where C and P denote combination (unordered selection) and permutation (ordered selection) functions respectively. As mentioned before, the selection of $r = 3$ planes is sufficient for the intersection and position determination. However, to be able to find possible incorrect intersections, it is preferable to choose r = 4, and perform a least squares estimation of the intersection point. To reduce the number of combinations, extracted planes are sorted according to their size, and only the top 10 largest planes are used in the search. With these settings, the search space for finding 4 corresponding planes in two sets of 10 planes will contain 1058400 combinations.

A brute force approach to the intersection of all the possible combinations will impose a great computational cost. To avoid this, it is desirable to perform a heuristic search for correct combinations by exploiting intrinsic constraints of the problem. Such constraints have been successfully used in the past to reduce the computational cost of combinatorial problems (Khoshelham and Li, 2004; Brenner and Dold, 2007). In this paper, we use the relative orientation of the corresponding planes as a heuristic to guide the search. The heuristic is based on the fact that the scanner rotation from one scan to another results in a same change in the orientation of the normal vectors in a set of corresponding planes. The comparison of the orientation of the normal vectors can be performed very efficiently in a reasonably small amount of time. By ruling out the combinations that exhibit an inconsistent change in the direction of the normal vectors, the search space can be greatly reduced to only a few correspondences (often 10~20).

The intersection procedure is performed for all the remaining correspondences. The mean residual of the observations in the least-squares estimation of the intersection point is used as a constraint to reject the false intersections. After false intersections are rejected, the median of the remaining points is taken as the final intersection. The use of median is advantageous because of the possibility of yet having a few outliers in the results. The outliers might be the result of errors in the parameters of the extracted planes, or the existence of nearly parallel planes (especially the floor and the ceiling) in the sets of corresponding planes. Using the median ensures that the final intersection point is not biased by the possible outliers.

EXPERIMENTAL RESULTS

The proposed indoor navigation method was tested on a set of range data collected in the corridor of the OLRS section at the building of the Aerospace Engineering Faculty of TU Delft. The test was carried out in an area of approximately 2x5x3

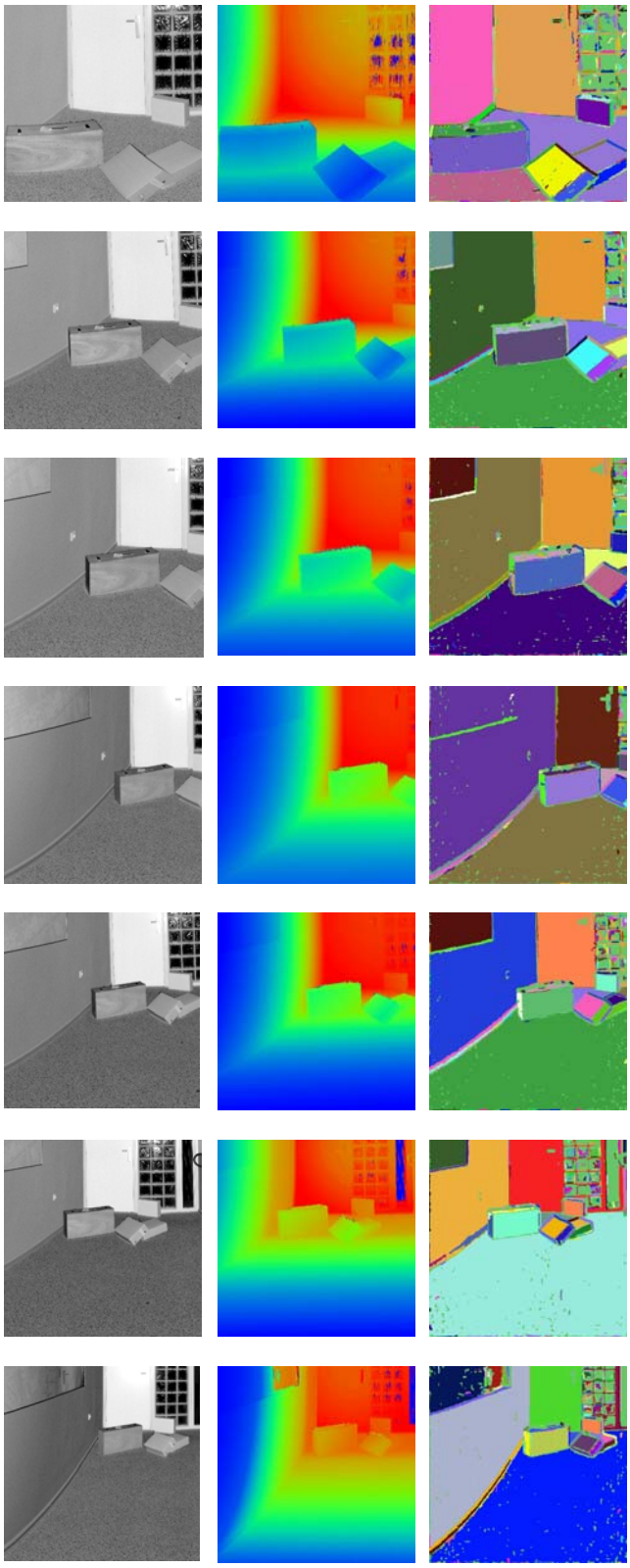

Fig. 5: Data and extracted features of the test scans. Left column: intensity images; middle column: range images; right column; segmented range images.

meters dimension. A total of 7 scans were obtained by a FARO LS 880 laser scanner at a low resolution. The scanner was moved along an arbitrary trajectory that was manually measured later with respect to a reference 2D coordinate system. The performance of the localization algorithm was carried out on the basis of comparing the computed positions with the manually measured reference positions. Only the measured position and orientation of the scanner at the initial starting point was used in the localization algorithm.

The segmentation algorithm was applied to the range images obtained at each scan. Fig. 5 depicts the segmented range images together with the original range images as well as the intensity images obtained at each scan. For each segmented image a list of the extracted planar surfaces with their parameters in the respective coordinate system of the scanner was generated.

The extracted plane parameters were introduced to the localization algorithm. For every pair of successive scans a search space containing all the combinations of four planes in the two scans was formed. Inconsistent combinations were found and ruled out by comparing the orientation parameters of the planes. A threshold of 3 degrees was empirically found suitable for the selection of combinations that showed a consistent change of orientation angles in two successive scans. Table 1 illustrates the changes of parameter θ of three planes in four successive scans. As can be seen, the planes exhibit changes of θ in different scans, which are within the designated 3 degrees of variation. In general, the comparison of parameters θ and φ reduced the search space to less than 0.01% of its original size in the majority of cases.

	θ_1 ($\delta\theta_1$)	θ_2 ($\delta\theta_2$)	θ_3 ($\delta\theta_3$)
Scan 1	149.6	238.4	61.3
Scan 2	157.3 (+7.7)	246.2 (+7.8)	68.5 (+7.2)
Scan 3	154.2 (-3.1)	243.2 (-3.0)	65.6 (-2.9)
Scan 4	153.3 (-0.9)	239.6 (-3.6)	64.3 (-1.3)

Table 1: Changes of parameter θ (in degrees) of three intersecting planes in four successive scans

The position of the scanner at each scan was computed by intersecting the corresponding planes found in the search procedure. The computed positions were transformed to the reference coordinate system by using the initial position and orientation of the scanner and the average rotation between every two successive scans. These were then compared to the manually measured positions to provide a means for assessing the accuracy of the localization algorithm. Fig. 6 shows the computed positions plotted together with the manually measured positions.

Table 2 summarizes the discrepancies between the computed and the manually measured coordinates of the laser scanner positions. As can be seen, the mean and the root mean square error measures are in the order of a few centimetres.

The CPU times required for both the segmentation and the localization phases were measured to provide an estimate of the computational cost of the algorithms. The algorithms were run on two separate computers. The segmentation algorithm was run on a Pentium 4 CPU with 2.4 GHz speed and 1 GB memory. The CPU time required for the segmentation algorithm was in the order of a fraction of a second for each scan. The localization phase was performed on a Pentium D CPU with 3.2 GHz speed and 2.00 GB memory. The CPU time was 21.7 seconds for the entire sets of planes, which is in average 3 seconds for each scan.

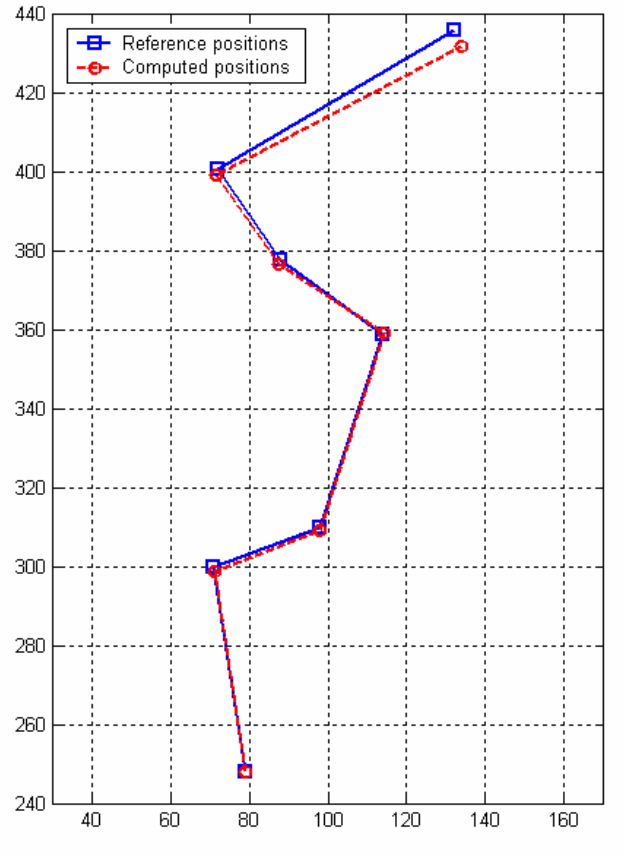

Fig. 6: The computed trajectory of the laser scanner as compared to the manually measured positions.

	X_{ref}	Y_{ref}	Z_{ref}	X_{com}	Y_{com}	Z_{com}	δX	δY	δZ
S_1	79	248	0	79	248	0.0	0	0	0
S_2	71	300	0	71.2	298.7	0.0	0.2	-1.3	0.0
S_3	98	310	0	97.8	309.1	0.2	-0.2	-0.9	0.2
S_4	114	359	0	114.2	359.2	-0.9	0.2	0.2	-0.9
S_5	88	378	0	87.7	376.7	-0.9	-0.3	-1.3	-0.9
S_6	72	401	0	71.7	399.2	-1.8	-0.3	-1.8	-1.8
S_7	132	436	0	133.9	432.1	-1.4	1.9	-3.9	-1.4
Mean Error		-			-		0.2	-1.3	-0.7
RMSE		-			-		1.9	4.7	2.6

Table 2: Discrepancies between computed and manually measured coordinates of the laser scanner position in 7 scans (in centimetres).

CONCLUSIONS

This paper introduced a method to reconstruct the trajectory of a moving laser scanner on the basis of indoor scans made at different positions. The method is composed of two main parts: First, obtaining segmentations of successive scans, second, localizing the scanner with respect to the mapped scene that is composed of the extracted planes in the segmented range images. The segmentation is entirely carried out using fast 2D image processing operations, and can be executed in real-time. The localization is based on keeping track of at least three intersecting planes in successive scans, and measuring distances to these planes. The method was shown to yield a positioning accuracy in the order of a few centimeters within 7 scans in an area of about 5 meters. The processing times also indicated computational efficiency of the method.

For this approach to be useful, for example in autonomous robot navigation, fast, affordable, and light equipment would be required that can be easily handled. This is not fulfilled when using the kind of terrestrial laser scanners presented in the experiment (FARO 880). Also, the time these scanners need for a single scan at each position does not favor real-time navigation. Alternative devices, such as SwissRanger by MESA, that can make 3d scans at video rates, are increasingly available, and are already being proposed as robot navigation sensors by various authors. The point density of such scanners is much lower, as is the signal-to-noise ratio of the distance measurements. However, this would only influence the positioning accuracy, and should have a minor impact on the navigation of the laser scanner within the mapped scene.

ACKNOWLEDGMENTS

The research was executed within a project called RGI-150: *Indoor positioning and navigation*, sponsored by the Dutch Ministry of Economic Affairs within the BSIK program.

REFERENCES

Araneda, A., Fienberg, S.E. and Soto, A., 2007. A statistical approach to simultaneous mapping and localization for mobile robots. *The Annals of Applied Statistics*, 1(1): 66-84.

Brenner, C. and Dold, C., 2007. Automatic relative orientation of terrestrial laser scans using planar structures and angle constraints, ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, Espoo, Finland, pp. 84-89.

Dedieu, D., Cadenat, V. and Soueres, P., 2000. Mixed camera-laser based control for mobile robot navigation, Proceedings of the 2000 *IEEE/RSJ lnternational Conference on Intelligent Robots and Systems*, pp. 1081-1086.

Diebel, J., Reutersward, K., Thrun, S., Davis, J. and Gupta, R., 2004. Simultaneous localization and mapping with active stereo vision, Proceedings of 2004 *IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan*, pp. 3436-3443.

Diosi, A. and Kleeman, L., 2004. Advanced sonar and laser range finder fusion for simultaneous localization and mapping, Proceedings of the 2004 *IEEE/RSJ lnternational Conference on Intelligent Robots and Systems, Sendai, Japan*, pp. 1854-1859.

Gorte, B., 1998. Probabilistic segmentation of remotely sensed images. PhD Thesis, ITC publication number 63, Enschede.

Gorte, B., 2007. Planar feature extraction in terrestrial laser scans using gradient based range image segmentation, ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, Espoo, Finland, pp. 173-177.

Guivant, J., Nebot, E. and Baiker, S., 2000. Autonomous navigation and map building using laser range sensors in outdoor applications. *Journal of Robotic Systems*, 17(10): 565-583.

Khoshelham, K. and Li, Z.L., 2004. A model-based approach to semi-automated reconstruction of buildings from aerial images. *Photogrammetric Record*, 19(108): 342-359.

Kim, M.Y. and Cho, H., 2006. An active trinocular vision system of sensing indoor navigation environment for mobile robots. Sensors and Actuators A: Physical, 125(2): 192-209.

Leonard, J.J. and Durrant-Whyte, H.F., 1991. Simultaneous map building and localization for an autonomous mobile robot, *IEEE/RSJ lnternational Workshop on Intelligent Robots and Systems IROS '91, Osaka, Japan*, pp. 1442-1447.

Newman, P., Leonard, J., Tardbs, J.D. and Neira, J., 2002. Explore and return: experimental validation of real-time concurrent mapping and localization, Proceedings of the 2002 *IEEE International Conference on Robotics & Automation, Washington DC*, pp. 1802-1809.

Victorino, A.C. and Rives, P., 2004. Bayesian segmentation of laser range scan for indoor navigation, Proceedings of 2004 *IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan*, pp. 2731-2736.

Xu, Z., Liu, J. and Xiang, Z., 2003. Map building and localization using 2D range scanner, Proceedings 2003 *IEEE International Symposium on Computational Intelligence in Robotics and Automation, Kobe, Japan,* pp. 848-853.

Zhang, L. and Ghosh, B.K., 2000. Line segment based map building and localization using 2D laser rangefinder, Proceedings of the 2000 *IEEE International Conference on Robotics & Automation, San Francisco, CA*, pp. 2538-2543.