REGISTRATION OF INDOOR 3D RANGE IMAGES USING VIRTUAL 2D SCANS

Marco Langerwisch and Bernardo Wagner
Institute for Systems Engineering, Real Time Systems Group
Leibniz Universität Hannover, Appelstraße 9A, D-30167 Hannover, Germany

Keywords: 3D laser range finders, Scan matching, Point cloud registration, Iterative closest point algorithm, Virtual 2D scans, Real time, Mobile robots.

Abstract: For mapping purposes, autonomous robots have to be capable to register 3D range images taken by 3D laser sensors into a common coordinate system. One approach is the Iterative Closest Point (ICP) algorithm. Due to its high computational costs, the ICP algorithm is not suitable for registering 3D range images online. This paper presents a novel approach for registering indoor 3D range images using orthogonal virtual 2D scans, utilizing the typical structure of indoor environments. The 3D registration process is split into three 2D registration steps, and hence the computational effort is reduced. First experiments show that the approach is capable of registering 3D range images much more efficient than ICP algorithm and in real time.

1 INTRODUCTION

Autonomous robots dealing with localization and navigation often do not have a map of their environment available in advance. In this case, the robot has to build a map on its own, based on some perceptual sensing. While conventional approaches use 2 dimensional (2D) mapping (Thrun, 2003), where three degrees of freedom (3DoF) have to be considered, recent approaches use 3 dimensions (3D) with its six degrees of freedom (6DoF), i.e. the x, y and z coordinates, and the roll, pitch, and yaw angles.

One approach used here for sensing the environment in three dimensions are 3D laser range finders. To build a map out of these captured single 3D range images, the rotation and translation have to be calculated to assemble these images into a common global coordinate system.

Lots of research work has been done in mapping two 3D range images into a common coordinate system. One approach is to use the Iterative Closest Point (ICP) algorithm. It calculates the point correspondences between the two images minimizing a mean square error function. Disadvantages of this approach are the computational costs due to the high number of 3D points, and therefore the execution time of the registration. For example, research in accelerating the ICP algorithm for 3D point mapping is done by (Nüchter et al., 2007). They use the algorithm to register 3D range images taken by a robot equipped with a tiltable 2D SICK laser range finder (Nüchter et al., 2005), trying to accelerate the 3D registration by using different representations of the 3D scans, and point reduction during the ICP algorithm.

Based on the typical orthogonality found in indoor environments (Nguyen et al., 2007) presents a lightwight localization and mapping algorithm. The main idea is to reduce complexity by mapping only planes that are parallel or perpendicular to each other. Because objects with different shapes are ignored, the resulting map does not contain any of these features.

A very similar approach to the contribution of our paper is presented in (Jez, 2008). The author first extracts leveled maps of the 3D range images. This is done by aligning the range images with the horizontal plane and then extracting points of vertical structures. These leveled maps are passed to a 2D ICP algorithm to calculate rotation and translation in 3DoF which are applied to the 3D range images. To obtain the final 6DoF transformation, the transformed range images are passed to a full 3D ICP algorithm. Drawbacks are the inability to deal with dynamic obstacles in this approach, and the full ICP at the end, that has high computational costs again.

In contrast, our aim is to avoid a full 3D ICP computation. Moreover, our approach is able to cope with dynamic objects and non-rectangular shaped indoor structures like round columns. It accelerates the range

Langerwisch M. and Wagner B. (2010).
REGISTRATION OF INDOOR 3D RANGE IMAGES USING VIRTUAL 2D SCANS.
In *Proceedings of the 7th International Conference on Informatics in Control, Automation and Robotics*, pages 327-332
DOI: 10.5220/0003003303270332

image registration in indoor enviroments compared to 6DoF ICP to let the robot register range images online. The structure of indoor environments is used to extract virtual 2D scans with a much lower number of points. Having three orthogonal virtual 2D scans, each of these can be registered very fast via 2D ICP to calculate rotation and translation. Finally, we combine the 2D transformations to get full 3D rotation and translation. First experiments show that with our approach, a mobile robot becomes capable to register 3D range images in real time with similar quality to 3D ICP registration.

The following section outlines the ICP algorithm and some needed notations, while section 3 describes the basic concepts of virtual 2D scans. Our new approach itself is presented in section 4, and section 5 shows experimental results that we obtained on one of our service robots. This paper closes with a conclusion and an outlook on future work.

2 ITERATIVE CLOSEST POINT ALGORITHM

To merge two images that have at least partial coverage into one common coordinate system, transformation (R,t) consisting of translation vector t and rotation matrix R have to be calculated. This process is called registration. One approach to register two images is the Iterative Closest Point (ICP) algorithm invented by (Besl and McKay, 1992), and (Chen and Medioni, 1991). It calculates the point correspondences between the two images while trying to minimize a mean square error function. An initial pose estimate is needed, so the ICP is a local registration algorithm (Magnusson et al., 2009).

The ICP algorithm registers a range image, given by a data point set $D = (d_1, ..., d_{N_D})$, into the coordinate system of a range image, given by model point set $M = (m_1, ..., m_{N_M})$, to minimize the error function

$$e(R,t) = \sum_{i=1}^{N_M} \sum_{j=1}^{N_D} w_{i,j} \parallel m_i - (Rd_j + t) \parallel^2 \qquad (1)$$

with $w_{i,j}$ assigned 1 if m_i is the closest point for d_j, otherwise 0.

$$N = \sum_{i=1}^{N_M} \sum_{j=1}^{N_D} w_{i,j} \qquad (2)$$

is the number of points taken into account for its calculation.

The main challenges here are how to compute the closest points and the transformation (R,t) efficiently. In our implementation we use a k-d-tree, a generalization of binary search trees, as point representation for finding the closest point correspondences (Bentley, 1975), and an algorithm based on singular value decomposition (SVD) (Arun et al., 1987) to compute the transformation (R,t).

3 VIRTUAL 2D SCANS

Calculating with full 3D raw point data sets is very computational demanding. The aim of virtual 2D scans is to reduce these 3D point sets to 2D point sets (Wulf et al., 2004). Virtual 2D scans are calculated regarding a special purpose, e.g. localization or navigation. For this special purpose, the raw 3D range images include many redundant and unnecessary information, and points respectively. Extracting points without losing relevant information and projecting them at a 2D plane reduces subsequent computational costs, because 2D algorithms can be used with a relative small number of points instead of 3D algorithms.

According to (Wulf, 2008), virtual 2D scans are calculated in two steps. First, reduce the amount of points of the 3D point set S:

$$f : S \to S_r, \quad \text{having } S_r \subset S. \qquad (3)$$

The size of the new point set S_r is $l \ll |S|$. Function f depends on the special purpose of the virtual 2D scan and will be defined in the next section.

In the second step the virtual 2D scan V_S is calculated from the reduced point set S_r by

$$g : (x,y,z) \to (x,y), \quad x,y,z \in \mathbb{R}, \qquad (4)$$

where x, y and z are the x, y and z coordinates of a 3D point. Equation 4 projects the 3D points of the reduced point set S_r on the 2D plane.

Provided that (3) fits the purpose of this reduction, all relevant information are kept in the virtual 2D scan V_S, and subsequent algorithms will be computationally less demanding.

4 3D RANGE IMAGE REGISTRATION

The following subsection reveals the basic concept underlying our approach, while in section 4.2 the actual algorithm we used for registering 3D range images via virtual 2D scans is described. The special virtual 2D scans used in our algorithm are presented in section 4.3.

4.1 Splitting of 6DoF Transformations

The basic idea behind our approach is that a 6DoF transformation can be split up into several transformations, each having less than six degrees of freedom. Each single transformation can then be computed by an ICP algorithm with much higher efficiency than full 6DoF ICP, resulting in a smaller total computation time. The single transformations are combined subsequently to gain the full 6DoF transformation.

Our approach is to split the 6DoF transformation into three 2D transformations that can each be calculated by a 2D ICP algorithm. For calculating the transformations, we extract three orthogonal virtual 2D scans. In our case, these virtual 2D scans are aligned along the x-y-plane, the x-z-plane, and the y-z-plane respectively. Using virtual 2D scans utilizes the advantages described in the previous section. Virtual 2D scans can cope with every shape in indoor environments, round columns for example. They can be defined insensitive to dynamic objects.

4.2 Registration Algorithm

It is assumed that the x-axis points forwards from the robot, the y-axis to the right and the z-axis downwards. The functions according to (3) of the virtual 2D scans will be described in the next subsection. Moreover, we assume that the scans to be registered are coarsely aligned by an initial pose estimate, e.g. by odometry or a gyrometer on a moving platform. Hence, the approach presented is a local registration algorithm just as the ICP. The 3D data point set D is to be registered into the coordinate system of the 3D model point set M.

First, a virtual 2D scan of the x-y-plane is extracted from both 3D point sets. Then, we compute a 2D ICP algorithm as described in section 2. We get the 2D rotation matrix R^{xy} and the 2D translation vector t^{xy}. Apply the rotation matrix

$$R' = \begin{pmatrix} R^{xy} & 0 \\ 0 & 1 \end{pmatrix} \tag{5}$$

and the translation vector

$$t' = \begin{pmatrix} t_1^{xy} & t_2^{xy} & 0 \end{pmatrix}^T \tag{6}$$

to the data point set D.
Next, extract virtual 2D scans of the x-z-plane from M and D'. Compute a 2D ICP algorithm as described in section 2. We get the 2D rotation matrix R^{xz} and the 2D translation vector t^{xz}. Calculate rotation matrix

$$R'' = \begin{pmatrix} R_{11}^{xz} & 0 & R_{12}^{xz} \\ 0 & 1 & 0 \\ R_{21}^{xz} & 0 & R_{22}^{xz} \end{pmatrix}. \tag{7}$$

Because we translated in x- and y-direction in the previous step, we just use the translation in z-direction:

$$t'' = \begin{pmatrix} 0 & 0 & t_2^{xz} \end{pmatrix}^T \tag{8}$$

Apply R'' and t'' to the data point set D' transformed in the previous step.
Finally, extract virtual 2D scans of the y-z-plane from M and D'' and compute a 2D ICP algorithm as described in section 2. We get the 2D rotation matrix R^{yz} and the 2D translation vector t^{yz}. Calculate rotation matrix

$$R''' = \begin{pmatrix} 1 & 0 \\ 0 & R^{yz} \end{pmatrix}. \tag{9}$$

Because translation in all three directions has been calculated before, no more translation is applied in this step.
The global rotation matrix calculates to

$$R = R''' \cdot R'' \cdot R', \tag{10}$$

and the global translation vector calculates to

$$t = R''' \cdot R'' \cdot t' + R''' \cdot t''. \tag{11}$$

Finally, the registered data point set is

$$D_{reg} = R \cdot D + t. \tag{12}$$

A limitation of this algorithm is the assumption of small errors in the initial pose estimate. Having larger errors, the convergence of each single step to an optimal solution does not guarantee global optimality. Other variations of this algorithm are subject of current research, for example the introduction of iteration steps. Nevertheless, the previous algorithm has shown promising results even in case of errors in initial alignment (see section 5).

4.3 Virtual 2D Scans for Registration

Three virtual 2D scans are needed, one for the x-y-plane, one for the x-z-plane and one for the y-z-plane. The calculation is based on the assumption that the 3D range image is given by a point cloud having all scan points aligned in vertical columns and horizontal rows with reference to the center of the scanning device. Here, the assumption is met because we use a yawing 3D scan (Wulf et al., 2004). In case of other scanning devices, the 3D points have to be reordered first.

For the x-y-plane, the indoor landmark scan from (Wulf, 2008) is taken. It extracts landmarks of cluttered indoor scenes like offices or corridors. Because of its robustness against dynamic objects and occluded walls and its ability to cope with non-rectangular shaped environments, it fits very well for our purposes. The idea is to extract one point per vertical column that has the largest radial distance to the

scan root. Function f of (3) is defined accordingly. In indoor environments, we get a kind of ground plan because usually the point with the largest radial distance lies on a wall. If a wall is discontinuous due to doors or gangways, the following wall is extracted to the virtual 2D scan.

When computing the virtual 2D scan for the x-z- and the y-z-planes, the 3D scans have already been rotated around the z-axis and translated in x- and y-direction (see section 4.2). Hence we need to calculate the rotation around the x- and y-axis, and the translation in z-direction using virtual 2D scans. Our approach puts planes orthogonal to x- and y-axis respectively into the coordinate system. Points located near those planes are extracted to virtual 2D scans.

For the virtual 2D scan for the x-z-plane, we compute S_r (see (3)) as follows:
p is a 3D point composed of the coordinate-tuple (p_x, p_y, p_z), and d is a distance threshold.
For each horizontal row in S look for the closest points p' and p'' to the x-z-plane, having $p'_x >= 0$, and $p''_x < 0$ respectively. If p'_y is lower than d, $S_r = S_r \cup p'$.
If p''_y is lower than d, $S_r = S_r \cup p''$.
Equation 4 has to be modified to

$$g : (x,y,z) \rightarrow (x,z), \quad x,y,z \in \mathbb{R}. \qquad (13)$$

Finally, only rotation around x-axis is left that needs to be calculated by the virtual 2D scan for the y-z-plane (see section 4.2). The subset S_r of the last virtual 2D scan is calculated similar to above with respect to the y-z- instead of the x-z-plane.

Unfortunately, the virtual 2D scans for the x-z- and the y-z-plane are sensitive to dynamic objects located in a distance closer than d to the x-z-plane, and y-z-plane respectively. A possible solution could be virtual 2D scans similar to that of the x-y-plane. This is subject of current research, as well as other implementations of the virtual 2D scans capable of dealing with outdoor environments.

5 EXPERIMENTAL RESULTS

This section shows the experimental results of our approach under real time constraints. At first, we will introduce the experimental setup, while the qualitative and computational results of two experiments will be presented in subsections 5.2, 5.3, and 5.4. This section closes with an evaluation of two series of experiments registering range images with simulated error.

5.1 Experimental Setup

The 3D range images were taken by a rotating 3D laser scanner based on SICK's LMS291. An introduction to the scanning device is given in (Wulf et al., 2007). Several 3D scans were taken from different positions on a typical office floor with offices and doors on both sides.

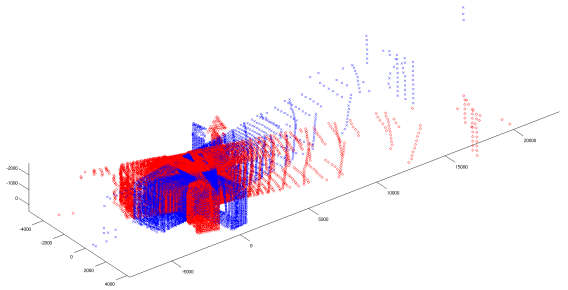


Figure 1: Model 3D scan (blue crosses), data 3D scan (red diamonds), not registered.

Table 1: Results of first experiment.

	Unreg	ICP	RegVirt
$e(R,t)$	29752	3369	3325
N	20483	25877	25870
Time [ms]	-	74050	67

For registering two 3D point clouds we used an embedded PC running our robotic framework RACK on a real time linux operating system. For comparison of the algorithms we used an accelerated ICP implementation, proposed e. g. in (Nüchter et al., 2007), as described in section 2. We conducted our tests by executing both algorithms with identical input data.

5.2 Qualitative Results

The scans we picked for the first evaluation consist of 25858 and 25987 raw 3D points respectively. Both scans in their local coordinate systems are shown in Figure 1. When executing our registration algorithm using virtual 2D scans, three 2D registration steps have to be performed. After 25 iterations of the 2D ICP, the scans were registered successfully in the first step (x-y-plane). The second and third registration steps took 14 and 4 iterations respectively.

For qualitative comparison of our new approach and the 3D ICP we calculated the error function $e(R,t)$ according to (1) before and after full 3D registration. Table 1 shows the results of the error function and the number of points N according to (2) taken into account for its calculation. The translation error between the results of both algorithms is 22 mm.

As can be seen, there are no qualitative differences worth mentioning. The error function reaches a similar value, as well as the number of points taken into account for calculating the error function. This leads to the result that our new approach reaches a similar

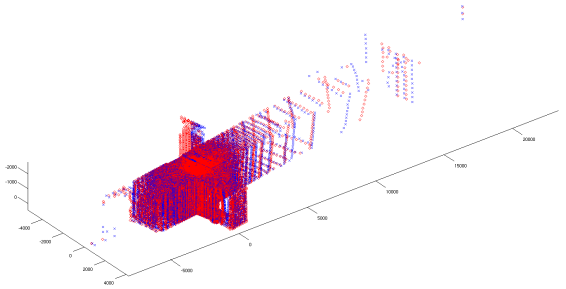


Figure 2: Model 3D scan (blue crosses), data 3D scan (red diamonds), registered.

Table 2: Results of second experiment.

	Unreg	ICP	RegVirt
$e(R,t)$	33908	5048	6376
N	23564	26131	26157
Time [ms]	-	36950	94

quality to a conventional ICP algorithm when registering two indoor 3D range images with small translation in z-direction and rotation around x- and y-axes.

Figure 2 shows both 3D point clouds registered into the model coordinate system by our new registration approach.

5.3 Non-zero Pitch and Roll Angles

To evaluate how well our approach copes with significantly non-zero pitch and roll angles, we tilted our experimental platform for appr. 4 cm to the top while taking the 3D data point set. Together with the yaw motion, this results in non-zero pitch and roll angles with reference to the coordinate system of the 3D model point set. Figure 3a shows the unregistered point clouds. The scans consist of 26088 and 26311 raw 3D points respectively. The remaining setup is the same as presented in section 5.1.

Table 2 shows the results of the error function and the number of points N according to (2) taken into account for its calculation. Computing the registration took 24, 18, and 9 iterations respectively with our new approach. The translation error between both results is 62 mm. Using a gyrometer on a moving platform could increase accuracy, i.e. decreasing the slight difference in the error function resulting from the coarse alignment before the registration. However, the minor difference in the quality can be neglected when achieving real-time computing capabilities in return, as presented in the next subsection.

Figure 3b shows both 3D point clouds registered into the model coordinate system by our new registration approach.

5.4 Computational Requirements

Because of the reduction from 3D to 2D ICP algorithm, our new approach needs much less computational effort for registering the 3D range images. As can be seen in Table 1, the calculation time reduces from appr. 74 seconds in the case of full 3D ICP to 67 milliseconds in the case of our new approach. In the second experiment, the reduction is from appr. 37 seconds to 94 milliseconds (see Table 2). When a full 3D scan takes 1.2 seconds as in our case (Wulf et al., 2007), our algorithm is capable of registering subsequent taken 3D range images online, i.e. in real time.

5.5 Simulated Error

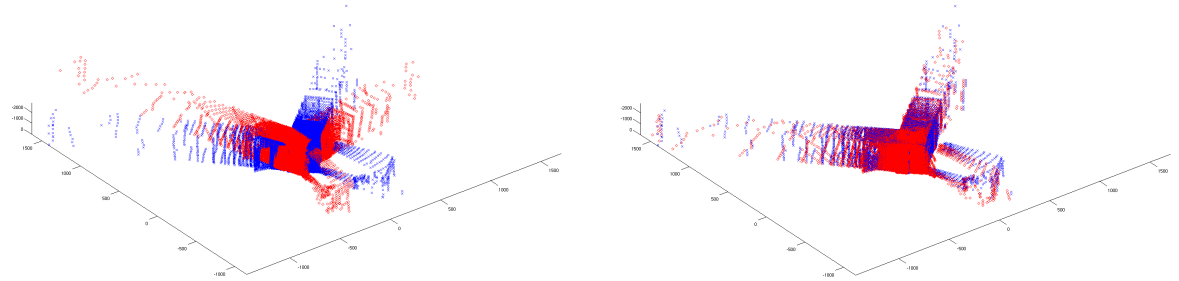
Finally, we evaluated two series of registration with simulated errors. Therefore, we transformed the model 3D scan of Figure 3a and registered it to the untransformed model scan. In a first series, we translated the scan in each direction, ranging from -30 cm to +30 cm, and taking incremental steps of 10 cm. This resulted in a series of 343 registrations for both algorithms. Surprisingly, our new approach had a mean translation error of 0.38 mm, compared to 0.49 mm using the ICP algorithm. The computation time was 25 ms (min. 17 ms, max. 31 ms) in average, compared to 19.7 s (min. 253 ms, max. 42.5 s).

In a second series, we rotated the scan around all axes, ranging from -10° to +10° each. The applied increment was 5°, resulting in a total of 125 registrations. The mean rotation error of our new approach was 0.86° compared to 0.01° in case of ICP. The average computation times were 55 ms (min. 16 ms, max. 113 ms), and 30.3 s (min. 254 ms, max. 56.6 s).

Examining these two experimental series with simulated errors, the computation times compared to the ICP algorithm were convincing again. The translation error with our new approach was even lower than with ICP. In case of rotation we evaluated a higher rotation error using our approach, but we are confident to gain an improvement with further research regarding the algorithm (see section 4.2).

6 CONCLUSIONS AND FUTURE WORK

This paper presents an approach for registering indoor 3D range images in real time. In our experiments, it registers two 3D point clouds up to 1100 times faster than the well-known ICP algorithm. This is done by splitting the full 6DoF transformation into three transformations. Each single transformation is calculated



(a) Unregistered 3D scans. (b) Registered 3D scans.

Figure 3: Registration of 3D scans with non-zero roll, pitch, and yaw angles.

by a 2D ICP algorithm with orthogonal virtual 2D scans, utilizing the typical structure of indoor environments. Due to the use of virtual 2D scans, the algorithm can cope with dynamic objects (with small restictions) as well as with non-rectangular environments. The resulting registration leads to a similar quality to the full 3D ICP algorithm.

Limitations and therefore future research work have been identified in sections 4.2 and 4.3. Nevertheless, our experiments showed promising results. A comparison to other approaches like techniques based on the extraction of planes out of the 3D range images and matching them will be necessary to stress the advantages of our algorithm. Further future work will deal with the implementation and evaluation of a full SLAM cycle based on the registration of 3D range images using virtual 2D scans.

REFERENCES

Arun, K. S., Huang, T. S., and Blostein, S. D. (1987). Least-squares fitting of two 3-D point sets. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9(5):698–700.

Bentley, J. L. (1975). Multidimensional binary search trees used for associative searching. *Commun. ACM*, 18(9):509–517.

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

Chen, Y. and Medioni, G. (1991). Object modeling by registration of multiple range images. In *Proc. IEEE International Conference on Robotics and Automation ICRA 1991*, pages 2724–2729 vol.3.

Jez, O. (2008). 3D mapping and localization using leveled map accelerated ICP. In Bruyninckx, H., Preucil, L., and Kulich, M., editors, *EUROS*, volume 44 of *Springer Tracts in Advanced Robotics*, pages 343–353. Springer.

Magnusson, M., Nüchter, A., Lorken, C., Lilienthal, A., and Hertzberg, J. (2009). Evaluation of 3d registration reliability and speed - a comparison of icp and ndt. In *Proc. IEEE International Conference on Robotics and Automation ICRA 2009*, pages 3907–3912.

Nguyen, V., Harati, A., and Siegwart, R. (2007). A lightweight SLAM algorithm using orthogonal planes for indoor mobile robotics. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems IROS 2007*, pages 658–663.

Nüchter, A., Lingemann, K., and Hertzberg, J. (2005). Mapping of rescue environments with Kurt3D. In *Proc. IEEE International Safety, Security and Rescue Robotics, Workshop*, pages 158–163.

Nüchter, A., Lingemann, K., Hertzberg, J., and Surmann, H. (2007). 6D SLAM - 3D mapping outdoor environments. *J. Field Robot.*, 24(8-9):699–722.

Thrun, S. (2003). *Robotic mapping: a survey*, pages 1–35. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Wulf, O. (2008). *Virtuelle 2D-Scans - Ein Verfahren zur echtzeitfähigen Roboternavigation mit dreidimensionalen Umgebungsdaten*. PhD thesis, Leibniz Universität Hannover, Hannover, Germany.

Wulf, O., Arras, K., Christensen, H., and Wagner, B. (2004). 2D mapping of cluttered indoor environments by means of 3D perception. In *Proc. IEEE International Conference on Robotics and Automation ICRA 2004*, volume 4, pages 4204–4209 Vol.4.

Wulf, O., Nüchter, A., Hertzberg, J., and Wagner, B. (2007). Ground truth evaluation of large urban 6D SLAM. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems IROS 2007*, pages 650–657.