

# TOWARDS GLOBALLY CONSISTENT SCAN MATCHING WITH GROUND TRUTH INTEGRATION

J. Gailis, A. Nüchter

Informatics VII – Robotics and Telematics  
Julius-Maximilians University Würzburg, Germany  
andreas@nuechti.de

**KEY WORDS:** scan matching, SLAM, ground control point

## ABSTRACT:

The scan matching based simultaneous localization and mapping method with six dimensional poses is capable of creating a three dimensional point cloud map of the environment, as well as estimating the six dimensional path that the vehicle has travelled. The essence of it is the registering and matching of sequentially acquired 3D laser scans, while moving along a path, in a common coordinate frame in order to provide 6D pose estimations at the respective positions, as well as create a three dimensional map of the environment. An approach that could drastically improve the reliability of acquired data is to integrate available ground truth information. This paper is about implementing such functionality as a contribution to 6D SLAM (simultaneous localization and mapping with 6 DoF) in the 3DTK – The 3D Toolkit software (Nüchter and Lingemann, 2011), as well as test the functionality of the implementation using real world datasets.

## 1. INTRODUCTION

Before the problem and the proposed solutions are discussed in depth, it is necessary to provide the background information on the involved technologies and current research efforts that have enabled this work to be carried out.

The concept of using laser technology for distance measuring was developed in the early 1960s. These efforts eventually resulted in the creation of LiDAR (an acronym for Light Detection And Ranging or a portmanteau of "light" and "radar"), which can be used as a term referring to the general technology, as well as an umbrella term that refers to technological solutions realizing different distance measuring methods that employ a laser. The LIDAR technology is currently quite established and has found many uses ranging from geomatics to remote sensing. Laser scanners mounted on aerial or orbital platforms are used to create terrestrial and planetary digital surface models (Smith et al., 2001), deduce atmospheric properties (Fernald, 1984), gather data for use in forestry (Dubayah and Drake, 2000), agriculture, environmental sciences and more (Campbell, 2002). Apart from one and two dimensional laser scanners used in the mentioned scenarios, three dimensional laser scanners with a near spherical field of view have been developed and are being increasingly used in the recent decades. Such laser scanners as well as carefully arranged two dimensional laser scanner configurations are capable of creating a precise three dimensional point cloud of the surrounding environment.

Although scans acquired by using a robotic platform can nowadays be registered reliably in a common reference frame there is no guarantee of this generated point cloud being in any way attached to ground truth even if it is known. This effectively renders the results of using a robotic platform in these scenarios substandard. Another scenario that could benefit from integrating the ground truth information in the 6D SLAM is the field of car mounted road surveying. Such systems usually use GPS signal to generate reasonably reliable pose information to aid mapping. However, in important infrastructure objects, such as highway tunnels, the GPS signal can be unreliable. Integrating ground

truth data obtained by different surveying methods could improve the overall result. Thus, the topic of implementing known poses or ground truth points along the path of the vehicle carrying out 6D SLAM is a worthwhile effort that would advance the method and widen its fields of application.

The paper is structured as follows: First, we recapitulate the basics of 6D SLAM, i.e., of globally consistent scan matching. Afterwards, we present our ground control point inclusion methods as well as experiments and results.

## 2. POINT CLOUD REGISTRATION AND SCAN MATCHING BASED SLAM

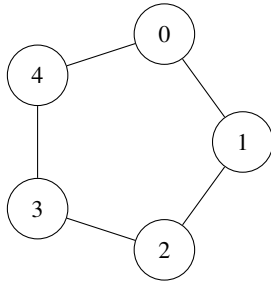
### 2.1 Pre-alignment of scans with ICP

We use the well-known ICP algorithm (Besl and McKay, 1992) to calculate the transformation while the vehicle is acquiring a sequence of 3D scans. ICP calculates point correspondences iteratively. In each iteration step, the algorithm selects the closest points as correspondences and calculates the transformation  $(\mathbf{R}, \mathbf{t})$  for minimizing the equation

$$E(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^m \|\mathbf{m}_i - (\mathbf{R}\mathbf{d}_i + \mathbf{t})\|^2, \quad (1)$$

where the tuples  $(\mathbf{m}_i, \mathbf{d}_i)$  of corresponding model and data points are given by minimal distance, i.e.,  $\mathbf{m}_i$  is the closest point to  $\mathbf{d}_i$  within a close limit (Besl and McKay, 1992). The underlying assumption of the ICP algorithm is that the point correspondences are correct in the last iteration. In each iteration, the transformation is calculated by the quaternion based method of Horn (Horn, 1987).

To digitalize environments without occlusions, multiple 3D scans have to be registered. Consider a vehicle traveling along a path, and traversing  $(n+1)$  3D scan poses  $\mathbf{X}_0, \dots, \mathbf{X}_n$ . A straightforward method for aligning several 3D scans taken from the poses  $\mathbf{X}_0, \dots, \mathbf{X}_n$  is *pairwise ICP*, i.e., matching the scan taken from



$$\mathbf{B} = \begin{pmatrix} \mathbf{C}_{0,1}^{-1} + \mathbf{C}_{1,2}^{-1} + \mathbf{C}_{4,0}^{-1} & -\mathbf{C}_{1,2}^{-1} & 0 & -\mathbf{C}_{4,0}^{-1} \\ -\mathbf{C}_{1,2}^{-1} & \mathbf{C}_{1,2}^{-1} + \mathbf{C}_{2,3}^{-1} & -\mathbf{C}_{2,3}^{-1} & 0 \\ 0 & -\mathbf{C}_{2,3}^{-1} & \mathbf{C}_{2,3}^{-1} + \mathbf{C}_{3,4}^{-1} & 0 \\ -\mathbf{C}_{4,0}^{-1} & 0 & \mathbf{C}_{3,4}^{-1} & \mathbf{C}_{3,4}^{-1} + \mathbf{C}_{4,0}^{-1} \end{pmatrix} \quad (2)$$

$$\mathbf{A} = \begin{pmatrix} -\mathbf{C}_{0,1}^{-1} \bar{\mathbf{E}}_{0,1} + \mathbf{C}_{1,2}^{-1} \bar{\mathbf{E}}_{1,2} \\ -\mathbf{C}_{1,2}^{-1} \bar{\mathbf{E}}_{1,2} + \mathbf{C}_{2,3}^{-1} \bar{\mathbf{E}}_{2,3} \\ -\mathbf{C}_{2,3}^{-1} \bar{\mathbf{E}}_{2,3} + \mathbf{C}_{3,4}^{-1} \bar{\mathbf{E}}_{3,4} \\ -\mathbf{C}_{3,4}^{-1} \bar{\mathbf{E}}_{3,4} + \mathbf{C}_{4,0}^{-1} \bar{\mathbf{E}}_{4,0} \end{pmatrix} \quad (3)$$

(4)

Figure 1. Simple loop containing 5 vertices. The corresponding system of linear equations is shown on right side.

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & -\bar{t}_z \cos(\bar{\theta}_x) + \bar{t}_y \sin(\bar{\theta}_x) & \bar{t}_y \cos(\bar{\theta}_x) \cos(\bar{\theta}_y) + \bar{t}_z \cos(\bar{\theta}_y) \sin(\bar{\theta}_x) \\ 0 & 1 & 0 & \bar{t}_z & -\bar{t}_x \sin(\bar{\theta}_x) & -\bar{t}_x \cos(\bar{\theta}_x) \cos(\bar{\theta}_y) + \bar{t}_z \sin(\bar{\theta}_y) \\ 0 & 0 & 1 & -\bar{t}_y & \bar{t}_x \cos(\bar{\theta}_x) & -\bar{t}_x \cos(\bar{\theta}_y) \sin(\bar{\theta}_x) - \bar{t}_y \sin(\bar{\theta}_y) \\ 0 & 0 & 0 & 1 & 0 & \sin(\bar{\theta}_y) \\ 0 & 0 & 0 & 0 & \sin(\bar{\theta}_x) & \cos(\bar{\theta}_x) \cos(\bar{\theta}_y) \\ 0 & 0 & 0 & 0 & \cos(\bar{\theta}_x) & -\cos(\bar{\theta}_y) \sin(\bar{\theta}_x) \end{pmatrix} \quad (5)$$

$$\mathbf{M}_i = \begin{pmatrix} 1 & 0 & 0 & 0 & -d_{y,i} & d_{z,i} \\ 0 & 1 & 0 & -d_{z,i} & d_{x,i} & 0 \\ 0 & 0 & 1 & d_{y,i} & 0 & -d_{x,i} \end{pmatrix}. \quad (6)$$

Figure 2. Definition of the matrix decomposition  $\mathbf{M}$  and  $\mathbf{H}$ .

pose  $\mathbf{X}_1$  against the scan from pose  $\mathbf{X}_0$ , matching the scan from  $\mathbf{X}_2$  against the one from  $\mathbf{X}_1$ , and so on. The detection of closed loops operates on the registered scans.

In case of a mobile mapping vehicle, we exploit the fact that these vehicles typically feature two tilted 3D scanners. The pre-alignment is carried out between the two stripes, which gage surfaces of the environment in a time-delayed manner. The continuous stream of observations is broken into small sections and treated in a semi-rigid manner. This semi-rigid registration is very similar to and extends the scan matching described next.

## 2.2 Globally Consistent Scan Matching

A 6DoF graph optimization algorithm for global relaxation is employed, a variant of GraphSLAM. Our method relies on a notion of the uncertainty of the poses, calculated by the registration algorithm. The following method extends the probabilistic approach first proposed in (Lu and Milios, 1997) to 6DoF. For a more detailed description of the extension refer to (Borrmann et al., 2008a) and (Borrmann et al., 2008b). For each pose  $\mathbf{X}$ , the term  $\bar{\mathbf{X}}$  denotes a pose estimate, and  $\Delta\mathbf{X}$  is the pose error.

The positional error of two poses  $\mathbf{X}_j$  and  $\mathbf{X}_k$  is described by:

$$E_{j,k} = \sum_{i=1}^m \|\mathbf{X}_j \oplus \mathbf{d}_i - \mathbf{X}_k \oplus \mathbf{m}_i\|^2 = \sum_{i=1}^m \|\mathbf{Z}_i(\mathbf{X}_j, \mathbf{X}_k)\|^2. \quad (7)$$

Here,  $\oplus$  is the compounding operation that transforms a point into the global coordinate system. For small pose differences,  $E_{j,k}$  can be linearized by use of a Taylor expansion:

$$\mathbf{Z}_i(\mathbf{X}_j, \mathbf{X}_k) \approx \bar{\mathbf{X}}_j \oplus \mathbf{d}_i - \bar{\mathbf{X}}_k \oplus \mathbf{m}_i - (\nabla_j \mathbf{Z}_i(\bar{\mathbf{X}}_j, \bar{\mathbf{X}}_k) \Delta\mathbf{X}_j - \nabla_k \mathbf{Z}_i(\bar{\mathbf{X}}_j, \bar{\mathbf{X}}_k) \Delta\mathbf{X}_k) \quad (8)$$

where  $\nabla_j$ ,  $\nabla_k$  denote derivatives with respect to  $\mathbf{X}_j$  and  $\mathbf{X}_k$  respectively. Utilizing the matrix decompositions  $\mathbf{M}_i \mathbf{H}_j$  and

$\mathbf{D}_i \mathbf{H}_k$  of the respective derivatives that separate the poses from the associated points gives:

$$\mathbf{Z}_i(\mathbf{X}_j, \mathbf{X}_k) \approx \mathbf{Z}_i(\bar{\mathbf{X}}_j, \bar{\mathbf{X}}_k) - (\mathbf{M}_i \mathbf{H}_j \Delta\mathbf{X}_j - \mathbf{D}_i \mathbf{H}_k \Delta\mathbf{X}_k) \quad (10)$$

$$\approx \mathbf{Z}_i(\bar{\mathbf{X}}_j, \bar{\mathbf{X}}_k) - (\mathbf{M}_i \mathbf{X}'_j - \mathbf{D}_i \mathbf{X}'_k) \quad (11)$$

Appropriate decompositions are given for Euler angles, quaternion representation and the Helix transformation in (Nüchter et al., 2010). In the following, we will work with the pose representation as Euler angles. This matrix decomposition cannot be derived from first principles and was first presented in (Borrmann et al., 2008a). Since  $\mathbf{M}_i$  as well as  $\mathbf{D}_i$  are independent of the pose, the positional error  $E_{j,k}$  is minimized with respect to the new pose difference  $\mathbf{E}'_{j,k}$ :

$$\mathbf{E}'_{j,k} = (\mathbf{H}_j \Delta\mathbf{X}_j - \mathbf{H}_k \Delta\mathbf{X}_k) \quad (12)$$

$$= (\mathbf{X}'_j - \mathbf{X}'_k). \quad (13)$$

$\mathbf{E}'_{j,k}$  is linear in the quantities  $\mathbf{X}'_j$  that will be estimated so that the minimum of  $\mathbf{E}_{j,k}$  and the corresponding covariance are given by

$$\bar{\mathbf{E}}_{j,k} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{Z} \quad (14)$$

$$\mathbf{C}_{j,k} = s^2 (\mathbf{M}^T \mathbf{M}). \quad (15)$$

where  $s^2$  is the unbiased estimate of the covariance of the identically, independently distributed errors of  $\mathbf{Z}$ :

$$s^2 = (\mathbf{Z} - \mathbf{M} \bar{\mathbf{E}})^T (\mathbf{Z} - \mathbf{M} \bar{\mathbf{E}}) / (2m - 3). \quad (16)$$

Here  $\mathbf{Z}$  is the concatenated vector consisting of all  $\mathbf{Z}_i(\bar{\mathbf{X}}_j, \bar{\mathbf{X}}_k)$  and  $\mathbf{M}$  the concatenation of all  $\mathbf{M}_i$ 's.

Up to now all considerations have been on a local scale. With the linearized error metric  $\mathbf{E}'_{j,k}$  and the Gaussian distribution ( $\bar{\mathbf{E}}_{j,k}, \mathbf{C}_{j,k}$ ) a Mahalanobis distance that describes the global er-

ror of all the poses is constructed:

$$\mathbf{W} = \sum_{j \rightarrow k} (\bar{\mathbf{E}}_{j,k} - \mathbf{E}'_{j,k})^T \mathbf{C}_{j,k}^{-1} (\bar{\mathbf{E}}'_{j,k} - \mathbf{E}'_{j,k}) \quad (17)$$

$$= \sum_{j \rightarrow k} (\bar{\mathbf{E}}_{j,k} - (\mathbf{X}'_j - \mathbf{X}'_k)) \mathbf{C}_{j,k}^{-1} (\bar{\mathbf{E}}'_{j,k} - (\mathbf{X}'_j - \mathbf{X}'_k)). \quad (18)$$

In matrix notation,  $\mathbf{W}$  becomes:

$$\mathbf{W} = (\bar{\mathbf{E}} - \mathbf{H}\mathbf{X})^T \mathbf{C}^{-1} (\bar{\mathbf{E}} - \mathbf{H}\mathbf{X}). \quad (19)$$

Here  $\mathbf{H}$  is the signed incidence matrix of the pose graph,  $\bar{\mathbf{E}}$  is the concatenated vector consisting of all  $\bar{\mathbf{E}}'_{j,k}$  and  $\mathbf{C}$  is a block-diagonal matrix comprised of  $\mathbf{C}_{j,k}^{-1}$  as sub matrices. To derive (19) from (18) we used an incidence matrix and stacked the matrices  $\bar{\mathbf{E}}'_{j,k}$  and  $\mathbf{C}_{j,k}^{-1}$ . For the latter, stacking must proceed in a diagonal fashion. Minimizing function (19) yields new optimal pose estimates. The minimization of  $\mathbf{W}$  is accomplished by the following linear equation system:

$$(\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H}) \mathbf{X} = \mathbf{H}^T \mathbf{C}^{-1} \bar{\mathbf{E}} \quad (20)$$

$$\mathbf{B}\mathbf{X} = \mathbf{A}. \quad (21)$$

The symmetrical matrix  $\mathbf{B}$  consists of the sub matrices

$$\mathbf{B}_{k,j} = \mathbf{B}_{j,k} = \begin{cases} \sum_{l=0}^n \mathbf{C}_{j,l}^{-1} & (j = k) \\ -\mathbf{C}_{j,k}^{-1} & (j \neq k). \end{cases} \quad (22)$$

The entries of  $\mathbf{A}$  are given by:

$$\mathbf{A}_j = \sum_{\substack{k=0 \\ k \neq j}}^n \mathbf{C}_{j,k}^{-1} \bar{\mathbf{E}}_{j,k}. \quad (23)$$

In addition to solving for  $\mathbf{X}$  this allows us to compute the associated covariance  $\mathbf{C}_X$  of  $\mathbf{X}$ :

$$\mathbf{C}_X = \mathbf{B}^{-1}. \quad (24)$$

The results have to be transformed to obtain the optimal pose estimates as follows:

$$\mathbf{X}_j = \bar{\mathbf{X}}_j - \mathbf{H}_j^{-1} \mathbf{X}'_j, \quad (25)$$

$$\mathbf{C}_j = (\mathbf{H}_j^{-1}) \mathbf{C}_j^X (\mathbf{H}_j^{-1})^T. \quad (26)$$

Figure 1 shows a simple graph containing five vertices and five directed edges. Each edge denotes a scan matching, where the model set corresponds to the 3D point cloud with an outgoing edge and the data set corresponds to the point cloud with the incoming edge. For all points in the data set the closest point in the model set is calculated. Based on these point pairs the covariance matrices are estimated as stated above. The matrix  $\mathbf{B}$  features 4 entries, since the first 3D scan, i.e., scan 0, defines the coordinate system and is not transformed. However, the covariance matrices with index 0 appear in the loop closing and at  $\mathbf{B}_{1,1}$

Following the convention in (Borrman et al., 2008a), we repre-

sent a pose  $\mathbf{X}$ , as well as its estimate and error, in Euler angles

$$\mathbf{X} = \begin{pmatrix} t_x \\ t_y \\ t_z \\ \theta_x \\ \theta_y \\ \theta_z \end{pmatrix} \quad \bar{\mathbf{X}} = \begin{pmatrix} \bar{t}_x \\ \bar{t}_y \\ \bar{t}_z \\ \bar{\theta}_x \\ \bar{\theta}_y \\ \bar{\theta}_z \end{pmatrix} \quad \Delta\mathbf{X} = \begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta t_z \\ \Delta \theta_x \\ \Delta \theta_y \\ \Delta \theta_z \end{pmatrix}. \quad (27)$$

The matrix decomposition  $\mathbf{M}_i \mathbf{H} = \nabla \mathbf{Z}_i(\bar{\mathbf{X}})$  is given in Figure 2. As required,  $\mathbf{M}_i$  contains all point information, while  $\mathbf{H}$  expresses the pose information. Thus, this matrix decomposition constitutes a pose linearization similar to that proposed in the preceding section. While the matrix decomposition is arbitrary with respect to the column and row ordering of  $\mathbf{H}$ , this particular description was chosen due to its similarity to the 3D pose solution given in Lu and Milios (1997). Finally, a system of  $6(n-1)$  equations ( $n$  denotes the number of poses to be estimated) has to be solved, but since the pose graph is sparse the resulting equation system is sparse, too. We use a sparse Cholesky decomposition as SLAM back end (Borrman et al., 2008b)

### 3. INCLUDING INFORMATION OF GROUND CONTROL POINTS

Arguably, the most important part of exploiting any available global truth data in static or kinematic 6D SLAM implementations is the ability to fix a pose in the global relaxation step. As described previously, the globally consistent scan registration method with Euler angles fixes the zeroth scan while other scans are movable for registration (cf. Figure 1). This is usually the desired behaviour.

The first approach to pose fixing that was explored is the so called algorithmic approach. The main idea of this approach is to find a way to introduce the pose fixing functionality in the 6D SLAM algorithm without changing the way how the LUM with Euler angles works. In other words, without changing any mathematics behind the global relaxation algorithm.

The fixed scan version of 6D SLAM algorithm works similar to the global relaxation described above. Two additional steps are introduced. First, the graph is transformed to take into account the fixed scans. Namely, the nodes representing the fixed poses are merged with the zeroth node and the required link rearrangement is done.

After the graph is rearranged, the vector containing all scans has to be rearranged as well. Namely, a meta scan containing all fixed scans has to be created and set as the zeroth scan. After that all other scans have to be renumbered in a sequential manner accordingly. This has to be done to ensure proper generation of the matrices in the linear equation system  $\mathbf{B}\mathbf{X} = \mathbf{A}$ .

Figure 3 (above) depicts a typical graph for the SLAM problem. Each node is connected to a few further nodes as it is expected that there is a large enough overlap that lets these point clouds to be matched. The loop is closed, hence the last node is connected to the first node. The grey nodes need to be regarded as fixed. If this graph was passed to an unchanged LUM global relaxation algorithm, the grey nodes would be treated as any other. The way the LUM algorithm is defined, only the zeroth node is treated as fixed. Thus, to treat scans 3 and 8 in Figure 3 as fixed, the following actions have to be carried out:

1. Merge the fixed scans with the zeroth node.

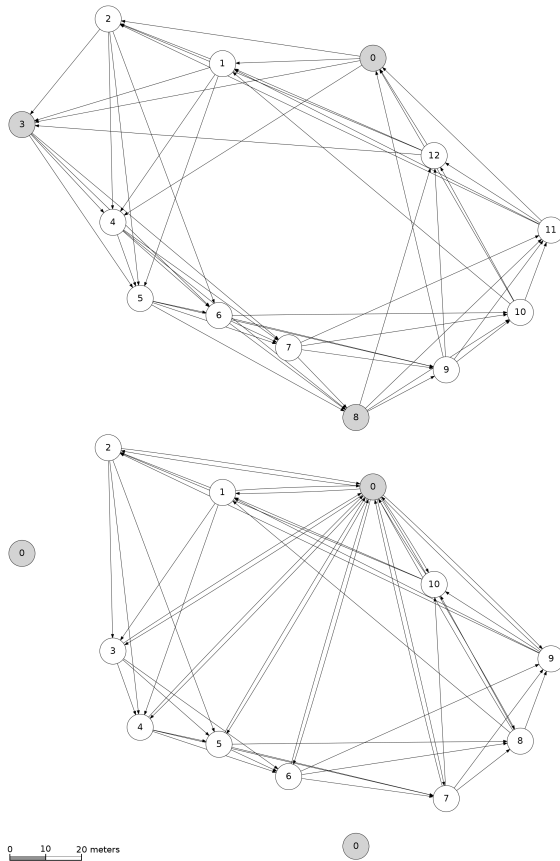


Figure 3. Above: A typical directed graph for a SLAM problem. Each node is connected to a few further nodes that can be matched to it and the loop is closed. The grey nodes have to be regarded as fixed. Below: Structure of the graph after the scan fixing procedure has been applied. The fixed scans are now treated as a single metascan.

2. Add the edges pointing to and from the fixed nodes to the zeroth node.
3. Remove duplicate links.

Figure 3 (below) depicts the example graph discussed previously after the described procedure.

As opposed to the algorithmic approach to pose fixing, the mathematical approach strives to fix the pose by exploiting the way how the global relaxation is being done. In other words, although the graph object has to be changed slightly in order to convey the necessary information about which poses have to be fixed, the graph structure should not be changed. Instead, the global relaxation algorithm should make use of available information regarding fixed scans and construct the linear equation system accordingly.

Given the math in section 2.. It is important to note that  $\mathbf{X}$  represents the concatenation of poses  $\mathbf{X}_1, \dots, \mathbf{X}_n$ . In other words, only the non-fixed poses compose the linear equation system  $\mathbf{B}\mathbf{X} = \mathbf{A}$  and are solved for. The  $j$  and  $k$  in (21) represent columns and rows in the corresponding arrays and hence only include movable nodes, the sums in 21 include the zeroth node as well through the use of index  $l$ . While this is appropriate notation when only the zeroth node is fixed, if a graph with an arbitrary number of fixed nodes is considered, the notation has to be extended. In this case, it is useful to use index  $l$  to denote the set of

fixed nodes. Then equations (21) become:

$$\mathbf{B}_{j,k} = \begin{cases} \sum_{k=0}^n \mathbf{C}_{j,k}^{-1} + \sum_{l=0}^n \mathbf{C}_{j,l}^{-1} & (j = k) \\ \mathbf{C}_{j,k}^{-1} & (j \neq k) \end{cases} \quad (28)$$

$$\mathbf{A}_j = \sum_{k=0, k \neq j}^n \mathbf{C}_{j,k}^{-1} \bar{\mathbf{E}}_{j,k} + \sum_{l=0}^n \mathbf{C}_{j,l}^{-1} \bar{\mathbf{E}}_{j,l} \quad (29)$$

where  $\mathbf{C}_{j,l}^{-1}$  represent the covariance between the  $j$ -th movable and  $l$ -th fixed scans and  $\bar{\mathbf{E}}_{j,l}$  is the corresponding observation. Then, the linear equation system (21) consists of  $n$  equations and can be solved for  $n$  variables, where  $n$  is the amount of movable nodes. The graph edge information, as well as the neighbour relations between all nodes, including the fixed ones, is taken into consideration while building the matrix  $\mathbf{B}$  and  $\mathbf{A}$  according to (28). To build the linear equation system, a fully connected graph is considered where the covariance between two nodes that are not connected with an edge is set to zero.

#### 4. EXPERIMENTS AND RESULTS WITH FIXED POSES

To test the functionality of the pose fixing methods, two experiments were carried out. In both cases data were gathered by a robot traversing the environment and stopping to make 3D point clouds. However, both cases have distinct environmental properties allowing to highlight the differences between used methods.

The first dataset used to verify the developed methods is a dataset recorded in Downtown Bremen, Germany. It contains a total of thirteen three dimensional point clouds generated using a RIEGL VZ-400 3D laser scanner. Each point cloud in turn contains up to 22.5 million points. The initial pose estimations needed for point cloud matching are provided by exploiting the odometry data of the robot. Figure 5. shows the result. One can see that the system cannot complete converge as the scans 0,3, and 8 cannot be moved.

Apparently, the difference that sets these two scan fixing algorithms apart in this situation has to be the final positions of the movable scans. This is easily explainable by using the graphs that have been used for both methods. The first difference is that the zeroth node in the graph used for algorithmic scan fixing contains all point clouds that have to be fixed as opposed to mathematical approach where the point clouds are treated as fixed yet separate. The second difference is the amount of links which is heavily reduced in algorithmic scan fixing as opposed to mathematical scan fixing.

The second dataset used to verify and test the developed scan fixing methods is the Hannover dataset. The dataset consists of 468 3D scans with around twenty thousand data points in each scan. The environment recorded in this dataset differs considerably from the one described by the Bremen dataset. The Hannover dataset considers of data taken by a mobile vehicle as it has travelled through the campus of University of Hannover. It is a relatively vast environment with low to medium rise buildings and park-like features such as trees and other vegetation. Most scans are matchable only to its direct spatial neighbours as no prominent features that can be observed from many vantage points around the campus are present in the dataset. The result is given as video at the following URLs: <http://youtu.be/c9ULBq9CyEQ> (original algorithm as given in (Borrmann et al., 2008a)), <http://youtu.be/aph7fFm8H94> (one scan fixed in addition to the origin), and <http://youtu.be/CFoTtDNaZuQ> (two scans fixed).

## 5. FROM FIXING SCAN POSES TOWARDS INCORPORATING GROUND CONTROL POINTS

In previous parts methods for doing simultaneous localization and mapping, as well as methods for denoting certain poses as fixed while doing SLAM have been described.

Let us consider a situation where a dataset has been acquired in a static terrestrial manner using a 3D scanner and there are ground truth points along the path that can readily be identified in the dataset. In this situation each scan provides a 3D representation of the surrounding environment in which zero to  $n$  ground truth points can be identified. The problem of aligning a pose with the underlying ground truth eventually boils down to finding a transformation between the set of the known ground truth points and the set of the corresponding points of the dataset. The criterion for reliable pose determination is to have at least three non-singular, non-linear ground truth points that can be identified in a single scan if current methods of determining a transformation between two matchable point clouds are used.

Kinematic SLAM adds its own corrections both to pose fixing and the incorporation of ground truth data. While the underlying algorithm for the pre-processing step is the same as for static SLAM, eventually the whole dataset is split up in smallest possible slices where each slice has its own modifiable pose as described in Elseberg et al. (2013). More often than not it will be the case that it is not possible to identify three ground truth positions in a single slice to perform unambiguous data association.

Experiments have been carried out on a tunnel dataset that contains 3D point cloud data taken using a LIDAR system consisting of two 2D laser scanners arranged in the typical "X" formation and positioned on top of a car. Apart from point and trajectory information, the dataset comes with ground truth information. Ground truth points have been acquired using independent surveying methods. The ground truth coordinates point to a bolt head in the centre of a painted circle on the tunnel's wall. Due to paint having different reflectance than the surrounding wall, the painted circle is an identifiable feature in the dataset. Due to separate point clouds being acquired by separate scanners and in separate runs, it is possible to compare datasets that should yield the same result and identify any discrepancies. Such comparison has been done and can be examined in Figure 4.

In future work, we aim at constraining the semi-rigid SLAM correction, with ground control points. In typical setups there are merely three control points per segment visible, thus fixing scans at ground truth poses is not sufficient. A possible solution would be to include the ground control points as a fixed "scan" and to incorporate a high weight of a matching between the extracted marker and these points.

## 6. CONCLUSIONS

This paper has presented methods for fixing scan poses in our globally consistent scan matching, i.e., our 6D SLAM optimization framework. Experiments confirm the applicability in several scenarios. Furthermore, we describe the real-world scenario of using ground control points in a tunnel mapping task.

### References

Besl, P. and McKay, N., 1992. A method for Registration of 3-D Shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)* 14(2), pp. 239–256.

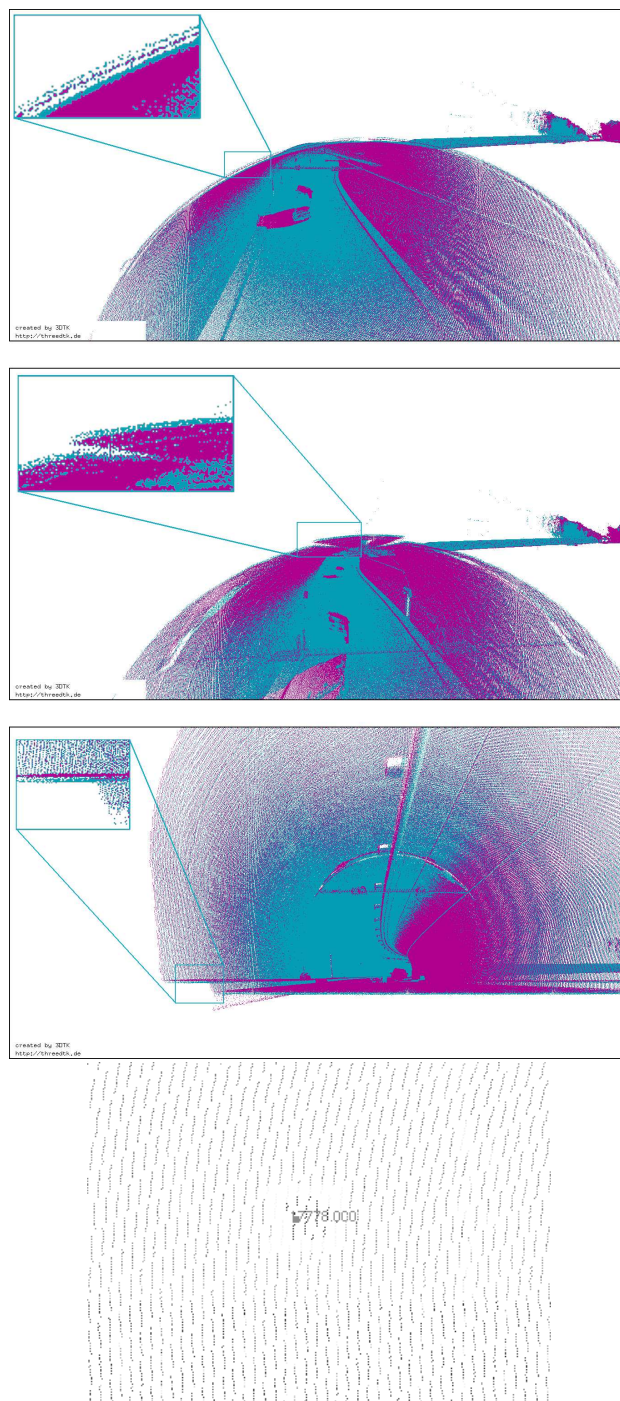


Figure 4. Above: The two images from the top depict discrepancies found between point clouds acquired using different scanners on the same run. The third image shows discrepancies between point clouds acquired using the same setup on different runs. Different colors denote different point clouds. Below: A ground truth point plotted together with data. The point's coordinates point to a bolt head in the middle of the painted circle.

Borrmann, D., Elseberg, J., Lingemann, K., Nüchter, A. and Hertzberg, J., 2008a. Globally Consistent 3D Mapping with Scan Matching. *Journal Robotics and Autonomous Systems (JRAS)* 56(2), pp. 130–142.

Borrmann, D., Elseberg, J., Lingemann, K., Nüchter, A. and Hertzberg, J., 2008b. The Efficient Extension of Globally Consistent Scan Matching to 6 DoF. In: *Proceedings of the 4th International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT '08)*, Atlanta, USA, pp. 29–36.

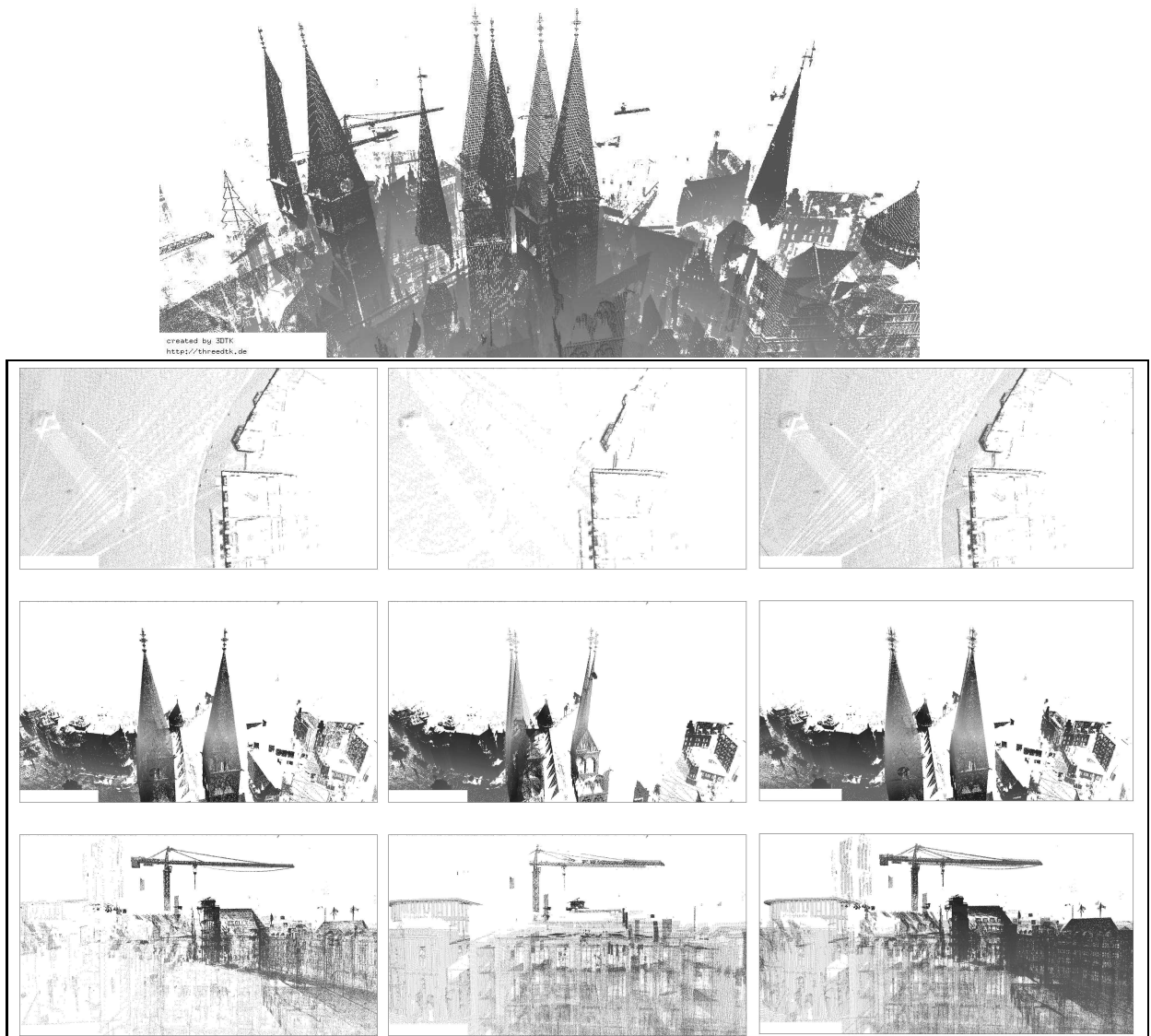


Figure 5. Above: Non-registered input 3D point clouds. Below: Left column contains images created using the original global relaxation without pose fixing, middle column contains images taken from the same position with algorithmic pose fixing, right column contains images with mathematical pose fixing.

Campbell, J. B., 2002. Introduction to remote sensing. CRC Press.

Dubayah, R. O. and Drake, J. B., 2000. Lidar remote sensing for forestry. *Journal of Forestry* 98(6), pp. 44–46.

Elseberg, J., Borrmann, D. and Nüchter, A., 2013. Algorithmic solutions for computing precise maximum likelihood 3d point clouds from mobile laser scanning platforms. *Remote Sensing* 5(11), pp. 5871–5906.

Fernald, F. G., 1984. Analysis of atmospheric lidar observations: some comments. *Applied optics* 23(5), pp. 652–653.

Horn, B. K., 1987. Closed-form solution of absolute orientation using unit quaternions. *JOSA A* 4(4), pp. 629–642.

Lu, F. and Milios, E., 1997. Globally Consistent Range Scan Alignment for Environment Mapping. *Autonomous Robots* 4(4), pp. 333 – 349.

Nüchter, A. and Lingemann, K., 2011. 3DTK – The 3D Toolkit.

Nüchter, A., Elseberg, J., Schneider, P. and Paulus, D., 2010. Study of Parameterizations for the Rigid Body Transformations of The Scan Registration Problem. *Journal Computer Vision and Image Understanding (CVIU)* 114(8), pp. 963–980.

Smith, D. E., Zuber, M. T., Frey, H. V., Garvin, J. B., Head, J. W., Muhleman, D. O., Pettengill, G. H., Phillips, R. J., Solomon, S. C., Zwally, H. J. et al., 2001. Mars orbiter laser altimeter: Experiment summary after the first year of global mapping of mars. *Journal of Geophysical Research: Planets* (1991–2012) 106(E10), pp. 23689–23722.