Calibration of multiple 3D LiDAR sensors to a common vehicle frame

Nina Heide, Thomas Emter, and Janko Petereit
Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Fraunhoferstr. 1, 76131 Karlsruhe, Germany, {nina.heide, thomas.emter, janko.petereit}@iosb.fraunhofer.de

Abstract

A calibration procedure for multiple 3D LiDAR sensors to a vehicle coordinate system is presented. Calibration consists of several preprocessing steps and point cloud registration by applying the Generalized Iterative Closest Point (GICP) algorithm modified with regard to the particular calibration scenario. After the extrinsic calibration of the sensors, the point cloud composed of all LiDAR measurements is calibrated to the vehicle coordinate system. The calibration reference is provided by a 3D model of the vehicle. Assuming single digit centimeter precision of the sensor data, calibration can be accomplished with single digit centimeter precision.

1 Introduction

Robots can relieve humans from many tasks like supporting people in their everyday life or in industrial production sites. Especially tasks in hazardous environment present danger or risks to the health and safety of humans. Hence they are suitable for execution by robots. To operate autonomously, perception of the environment is needed at first. By using multiple sensors, a large-scale coverage of the surrounding area can be accomplished. To fuse the sensor information into a single coordinate system, all sensors have to be calibrated relative to the chosen common frame. After preprocessing of the LiDAR point clouds, extrinsic calibration is executed at first. The calibration to the vehicle frame fuses the measurements of all extrinsically calibrated sensors and the geometric 3D model of the vehicle. The presented calibration procedure operates semi-automatically only needing user intervention to obtain an initial guess of orientation and rotation. Highly precise calibration results are achieved, reaching an accuracy of single digit centimeters within the noise level of the utilized LiDAR sensors. As technology demonstrator a commercially available excavator platform shown in Figure 1 is used. The platform has i.a. been especially equipped with LiDAR sensors for visual perception and rotatry encoders at each joint [1].

2 Related Work

Maye et al. [2] divide calibration methods into online methods optimizing the state vector of a bayes filter and offline methods minimizing a cost function. Iterative registration methods for point clouds minimize cost functions and are therefore classified as offline methods. The Iterative Closest Point (ICP) algorithm forms the basis for several advanced registration algorithms [3, 4]. Instead of minimizing the error between the distance of single corresponding points as in ICP, Chen and Medoni [5] present point-to-plane-ICP calculating the error between single points and estimated surfaces. The GICP algorithm of Segal et al. [7] optimizes an error metric obtained from the distances between estimated surfaces around the corresponding single points. They calculate the covariance of a fixed number of points around the respective corresponding point and extract the orientation of the estimated surface via singular value decomposition (SVD) of the covariance matrix [6]. The utilization of a fixed number of neighboring points for the calculation of the combined covariance matrix (see eq. 3) yields a surface estimation within a small surrounding sphere in dense areas and includes a significantly larger surrounding sphere in sparse areas of point clouds. Compared to ICP or rather point-to-plane ICP, the root mean square (RMS) error is reduced by $\frac{2}{3}$ or rather $\frac{1}{2}$ applying GICP registration. ICP and its further developments are local methods carrying the risk of convergence to a local minimum if the initialization is not determined correctly. Fitzgibbon [8] proposes an alternative procedure for point cloud registra-
tion, using the non-linear Levenberg-Marquardt algorithm for error minimization. No initial value is needed, as a global optimum is calculated. He presents quantitative results only for 2D point clouds without clearly recognizable geometric structures and not for significantly more complex 3D point clouds, which need determination of two additional degrees of freedom (DOF).

Many calibration approaches between a LiDAR sensor and a camera or a stereo camera system were published in recent years [9, 10, 11]. Schneider et al. [12] present an odometry-based online approach using an Unscented Kalman Filter. Either the extrinsic calibration of sensors providing 6 DOF information or the estimation of a planar offset and 1D orientation for sensors providing planar movement and orientation as single-beam LiDAR can be determined. The requirement of odometry measurements at two different points cannot be fulfilled for all autonomous vehicles. For vehicles exhibiting slippage such as track vehicles or even wheeled vehicles on slippery ground, odometry measurements are not reliable. Furthermore, a calibration for 3D LiDAR sensors as required in many robotic applications cannot be performed using the approach of Schneider et al. [12].

Gao and Spletzer [13] propose an online approach for calibrating multiple LiDAR sensors on a mobile vehicle platform, reformulating the non-linear calibration model as a second-order cone program. Utilizing retro-reflective tape on vertical poles, they filter the measurements generated by LiDAR sensors implementing a threshold on the remission measurements. The approach works completely autonomous, but poses the disadvantage of attaching retro-reflective tape before calibration.

Frese et al. [14] utilize the manipulator arm of a mobile platform as calibration object of a depth camera to the common sensor frame of the platform. For exact calibration several positions of the arm are recorded by the depth camera and aggregated into one point cloud. The point cloud for registration is extracted from the arm’s kinematic and geometric 3D model in combination with known joint angles. The aggregated camera point cloud is fitted to the 3D model cloud for all arm positions by using the GICP algorithm.

3 Methodology

3.1 Outline

Before applying the GICP algorithm for the registration of point clouds, the clouds are preprocessed with a nearest neighbor (NN) search as explained in section 3.2. Figure 2 explains the entire iterative calibration procedure. It starts with extrinsic calibration of a first pair of LiDAR sensors and concludes in the calibration to the chosen vehicle coordinate system. The preprocessing steps are not included for clarity. The LiDAR sensors colored yellow and green are symbolically showing the first sensor pair selected for extrinsic calibration. The multiple LiDAR sensors are calibrated extrinsically at first. Subsequently, all point clouds generated by LiDAR sensors are merged into one point cloud, using the results of the extrinsic calibration. Hence maximum information for the calibration to the common frame of the vehicle is available. Just as in the first calibration step, preprocessing of the extrinsically calibrated and merged measurement data of all sensors is performed. In the second calibration step, the merged and preprocessed data of all sensors is aligned to a specially created, closely matching 3D model of the vehicle, yielding the relative positions and orientations of all sensors to the vehicle frame.

The initial transformations required for the locally operating GICP algorithm are user-generated and represent the only manual steps in the automated procedure. To avoid the need for measurements by hand, the Robot Operating System (ROS) 3D visualization RViz is supplied with LiDAR data [15]. A 3D interactive marker is bound to the respective sensor position, constituting the origin of its point measurements. By moving and rotating these markers, the initial transformation can easily be generated by visually matching the point clouds inside a GUI [16].

The point cloud registration within the GICP algorithm is conducted by minimizing distances between the estimated surfaces. $T$ symbolizes the transformation from source to target point cloud, $t_i$ a point in the target and $s_i$ a point in the source point cloud. $C_i$ denotes the combined covariance matrix of the surface estimated around the corresponding point pair $(t_i,s_i)$ [7]. The minimization is performed by

\[
T = \arg\min_T \sum_i d_i^* C_i^{-1} d_i, \tag{1}
\]

with * denoting transposition and $d_i$ the Euclidean distance

\[
d_i = t_i - T s_i. \tag{2}
\]

The weighting applied by the combined covariance matrix $C_i$.

\[
C_i = C_i^T + T C_i T^T, \tag{3}
\]

allows high uncertainty inside the estimated surface geometry and low uncertainty into the direction of the normal orientation obtained from SVD. Thereby the covariance matrix of the target is denoted by $C_i^T$ and the covariance of the source point cloud by $C_i^T$. The preprocessing steps described in the following paragraphs and the input parameters to the GICP algorithm are evaluated on the basis of the Euclidean fitness score [17] $e_f$ as well as visual assessment. $e_f$ represents a benchmarking value calculated from the Euclidean distances $d_i$ of target and transformed source points weighted with the number of correspondences $N_C$:

\[
e_f = \frac{\sum_{i=0}^{N_C-1} (d_{x,i} - t_{x,i})^2 + (d_{y,i} - t_{y,i})^2 + (d_{z,i} - t_{z,i})^2}{N_C}. \tag{4}
\]

3.2 Preprocessing for extrinsic calibration

Previous to GICP registration, each point cloud is preprocessed in two steps. For the purpose of filtering outliers a NN search requiring a minimum number of neighbors within a defined surrounding sphere is conducted for each single point. The number of neighboring points depends on the number of reflections. Single points lying in sparse areas containing little or no objects causing reflections do
not contributing to registration and are removed. This is also referred to as separate preprocessing (SPP) as each cloud is filtered separately. Adding a minimum threshold for point distance from the sensor according to the specifications of the respective 3D LiDAR ensures proper measurements and is also imposed during SPP.

In the second preprocessing step the initial transformation is applied to transform the source point cloud into the coordinate system of the target point cloud (TNN). A minimum number of neighbor points from the target point cloud is required for each point of the transformed source point cloud, which eliminates areas only included in one of the clouds. After preprocessing, each point cloud pair is inserted into the GICP algorithm with the initial transformation as a starting point.

### 3.3 Extrinsic calibration

To extrinsically calibrate all LiDAR sensors, pairs of point clouds are successively selected for registration as explained in Figure 2. A first pair of LiDAR sensors is selected to obtain the largest overlap in their field of view (FOV). After calibration, their point clouds are merged using the resulting transformation. This merged cloud forms one element of the next pair for calibration, together with a point cloud from a single LiDAR. Pairwise calibration is conducted until all LiDAR measurements are merged into one point cloud with a common coordinate system.

Calibration order is defined to achieve a preferably large overlap in the FOV. The larger the overlap in the FOV, the more surfaces can be estimated and used for registration. The registration success also depends on the selection of source and target point cloud, as GICP works asymmetrically. Correspondences for all source points are searched inside the target point cloud. Thus the cloud showing higher density in the overlap of both clouds is chosen as target, consequently maximizing the possibility to find a corresponding point for each source point.

### 3.4 Calibration to vehicle frame

As previous to extrinsic calibration SNN and TNN of source and target point cloud is performed before applying GICP. The calibration of each LiDAR sensor to the coordinate system of the vehicle is done by registering the merged point cloud generated during extrinsic calibration to the preprocessed point cloud of the vehicle model. Naturally, the calibration to the vehicle frame can only be as exact as the 3D model of the vehicle. After determining the respective vehicle parts appearing in the FOV of at least one LiDAR sensor, dense point clouds are produced for those parts of the 3D model. For this purpose, a purpose-built algorithm extracts the geometric information from Collada data by identifying the subdivision of the contained surfaces and fills those surfaces with additional points, calculated on the basis of the planar equation generated from vertices. By applying the described algorithm, a very dense point cloud is created, which is used as target point cloud in the registration process. Consequently, the merged point cloud is used as source point cloud.

Given the 3D model of the vehicle, the surface orientations are exactly known. Instead of estimating the surface orientations via SVD of the covariance matrices as generally done in the GICP algorithm, the point cloud of the 3D model contains a normal for each point. Considering the normal information of vertex points for the generation of additional points, each point inside the model surfaces also includes the orientation of the surface it belongs to. Corner points maintain the normal information set in the 3D model. They do not need to be treated specially, as the points filled inside the surfaces by the purpose-built algorithm prevail significantly.

For registration, the initial transformation, source and 3D model target point cloud as well as adapted configuration parameters are input in the GICP algorithm. The evaluation of algorithm parameters and preprocessing thresholds is described in section 4.3.

---

**Figure 2** Flow diagram of entire calibration procedure: Yellow and green LiDAR form first pair for extrinsic calibration, subsequently they are extrinsically calibrated to the LiDAR represented by the purple circle, etc.
4 Evaluation

The presented calibration method is evaluated on the technology demonstrator shown in Figure 1. The commercially available excavator is equipped with additional sensors as described by Emter et al. [1], including 3D LiDAR sensors mounted horizontally to the sides of cabin and vertically below the boom. Evaluation is conducted on the basis of the Euclidean fitness score $e_f$ exhibited by the registered point clouds as well as on visual assessment. Visual assessment excludes convergence of the GICP algorithm in a local minimum. Additionally it enables the comparison of registration results in the parallelism of large, planar surfaces. The calibration procedure is validated by manually measuring the distance between the first sensor pair, which is rigidly mounted to the excavator cabin and hence easily accessible for measurements. The distance between the two LiDAR sensors is measured to 1.07 m in horizontal direction. The translational accuracy of the Velodyne LiDAR sensors is indicated as ± 0.03 m. The difficulty of the data sets is rated based on visual assessment, considering the registration operations inside the GICP algorithm. A data set rated as more difficult includes less objects with approximately smooth surfaces in the common FOV.

For parameterization of the GICP algorithm, an extensive evaluation is performed to achieve an optimal calibration result. The parameter for correspondence randomness (CR) defines the fixed number of neighbors used to calculate the covariance matrix. Rotation epsilon ($e_R$) and transformation epsilon ($e_T$) define the antiproportional weighting for the maximum change to be made in last step previous to convergence ($\Delta$). Thus, changes made to rotation and translation during the last optimization step need to be smaller than $\frac{1}{e_R}$ or rather $\frac{1}{e_T}$. Euclidean fitness epsilon ($e_E$) defined as $e_f$ difference between two consecutive transformations specifies the maximum RMS error before reaching convergence. $d_{MC}$ denotes the maximum distance between points from source and target point cloud to form a correspondence.

4.1 Preprocessing for extrinsic calibration

Calibration results applying SNN compared to GICP registration without preprocessing are shown in Table 1 and Figure 3. The minimum distance threshold for points from their LiDAR of origin is set to 0.85 m, as Velodyne specifics exact measurements from 1 m to 100 m and evaluation yields reliable results from 0.85 m without eliminating parts of the excavator boom, which can be used to calibrate the first sensor pair.

<table>
<thead>
<tr>
<th>Var.</th>
<th>opt.</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>opt.</th>
<th>C7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>nPP</td>
<td>100</td>
<td>100</td>
<td>20</td>
<td>1000</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$e_E$</td>
<td>nPP</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.1</td>
<td>10</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$d_{MC}$</td>
<td>nPP</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>$SNN_{min}$</td>
<td>nPP</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>$d_{SNN}$</td>
<td>nPP</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>$e_f^{LR}$</td>
<td>SPP</td>
<td>4.629</td>
<td>4.634</td>
<td>4.719</td>
<td>6.876</td>
<td>4.861</td>
<td>6.916</td>
<td>4.658</td>
<td>2.183</td>
</tr>
<tr>
<td>$e_f^{BR}$</td>
<td>SPP</td>
<td>0.139</td>
<td>0.140</td>
<td>0.140</td>
<td>0.137</td>
<td>0.138</td>
<td>0.137</td>
<td>0.138</td>
<td>0.138</td>
</tr>
</tbody>
</table>

Table 1 Evaluation of extrinsic calibration: $e_f^{LR}$ for first sensor pair (left-right, L-R), $e_f^{BR}$ for LiDAR mounted below the boom (source, B) to merged point cloud from L-R; $e_E$ and $e_E$ set to $10^{-4}$ $m^2$. $\Delta$ set to 1; $d_{MC}$ and $d_{SNN}$ in m; $e_f$ in $m^2$; $e_f$ calculated as mean between seven data sets with three walls as shown in Figure 3, 4 and 5.
Given data sets with three walls in FOV as for instance shown in Figure 3, the calibration accuracy and the most difficult data set included in the calibration procedure is considered less important than calibration results using data sets recorded with only two surroundings. As no comparably high values in \( \epsilon_f \) occur in the calibration of the first sensor pair to the third LiDAR mounted on the boom, high \( \epsilon_f \) seems to arise for a large overlap in FOV.

For TNN, the required number of neighbors \( TNN_{\text{min}} \) has been evaluated in combination with the radius \( d_{\text{TNN}} \) corresponding surrounding sphere as shown in Table 2. TNN aims at filtering areas only included in one of the clouds for registration. Hence, \( d_{\text{TNN}} \) is set to a larger value in combination with a larger \( TNN_{\text{min}} \) as for SPP, choosing the configuration with the lowest \( \epsilon_f \) and the most promising visual assessment as depicted in Figure 6. The calibration results using data sets recorded with only two surrounding walls (presented in Table 3) show the influences of different preprocessing steps on the \( \epsilon_f \). Utilizing preprocessing methods, \( \epsilon_f \) reduces by more than 97%.

Considering only \( \epsilon_f \) values, using TNN seems to make SNN unnecessary for data sets including two walls in the common FOV of the LiDAR sensors. As the runtime of the calibration procedure is considered less important than calibration accuracy and the most difficult data set included in the mean calculation of Table 3 shows a lower \( \epsilon_f \) for inclusion of SNN, this step remains in the calibration procedure.

### 4.2 Extrinsic Calibration

The merging of already calibrated LiDAR measurements maximizes the overlap in FOV for registration and remarkably improves calibration accuracy compared to singular calibration of the respective LiDAR clouds. This can be shown exemplarily by comparing the \( \epsilon_f \) values shown in Table 1 to registration of the cloud of the left LiDAR as target to the cloud of the LiDAR mounted on the boom. An \( \epsilon_f \) reduction of more than 60 percent by utilization of merged clouds can be achieved.

In the registration process of two point clouds with a considerably different number of points, the cloud containing more information is chosen as target. Taking the point cloud from the LiDAR mounted to the boom as target and the, significantly larger, cloud of the LiDAR mounted on the right of the cabin as source, \( \epsilon_f \) results to more than 0.316 m². Interchanging source and target cloud as proposed and hence choosing the dense cloud as target, \( \epsilon_f \) reduces to only 0.064 m² and thus less than 0.7 percent of the former \( \epsilon_f \) value. Thus, selecting the cloud containing more information as target significantly improves registration accuracy.

Figure 4 and 5 show the results of extrinsic calibration of the 3D LiDAR sensors mounted on the excavator. The distance between the LiDAR pair extracted from the calibration.

### Table 2: Combination of SPP and TPP for extrinsic calibration

<table>
<thead>
<tr>
<th>( d_{\text{MC}} )</th>
<th>( d_{\text{TNN}} )</th>
<th>( TNN_{\text{min}} )</th>
<th>( \epsilon_f )</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>C2</td>
<td>C8</td>
<td>C9</td>
</tr>
<tr>
<td>1.5</td>
<td>1.0</td>
<td>2.0</td>
<td>1.5</td>
</tr>
<tr>
<td>400</td>
<td>400</td>
<td>400</td>
<td>50</td>
</tr>
<tr>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>0.25</td>
</tr>
<tr>
<td>2.157</td>
<td>2.186</td>
<td>2.185</td>
<td>2.179</td>
</tr>
</tbody>
</table>

### Table 3: Evaluation of SPP, TPP and combination of both (STPP) in extrinsic calibration

<table>
<thead>
<tr>
<th>( \epsilon_f )</th>
<th>( e_{\text{fs}} )</th>
<th>( e_{\text{fs}} )</th>
<th>( e_{\text{fs}} )</th>
<th>( e_{\text{fs}} )</th>
<th>( e_{\text{fs}} )</th>
<th>( e_{\text{fs}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.660</td>
<td>2.599</td>
<td>0.062</td>
<td>0.060</td>
<td>( e_{\text{fs}} )</td>
<td>( e_{\text{fs}} )</td>
<td>( e_{\text{fs}} )</td>
</tr>
<tr>
<td>( e_{\text{fs}} \times 100 % )</td>
<td>–</td>
<td>71.0</td>
<td>1.7</td>
<td>1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.912</td>
<td>0.160</td>
<td>0.114</td>
<td>0.123</td>
<td>( e_{\text{fs}} )</td>
<td>( e_{\text{fs}} )</td>
<td>( e_{\text{fs}} )</td>
</tr>
<tr>
<td>( e_{\text{fs}} \times 100 % )</td>
<td>–</td>
<td>2.3</td>
<td>1.7</td>
<td>1.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.3 Calibration to vehicle frame

The merged cloud is formed from all LiDAR measurements applying the results from extrinsic calibration. The origin of the merged point cloud is equivalent to the mounting point of the last LiDAR calibrated onto. Hence for the application on the technology demonstrator shown in Figure 1 the origin of the merged cloud equals the frame of the LiDAR on the excavator boom.

As for extrinsic calibration, preprocessing is done for the merged cloud of all LiDAR sensors before performing registration. The evaluation of SNN yields the most accurate calibration results to 1.04 m, compared to 1.07 m measured by hand. Hence, the validation using measurement of the distance between the first LiDAR pair for calibration yields a difference of 0.03 m. As the accuracy of the Velodyne LiDARs is specified to ±0.03 m, the difference lies in the range of LiDAR noise.

Resuming the results of the error metric and visual assessment, it can be deduced that in addition to the ground plane at least two planar surfaces need to be included in the FOV from all 3D LiDAR sensors to get an extrinsic calibration with an accuracy in the range of single digit centimeters. As the technology demonstrator can be moved next to a building and on planar ground surface, the calibration procedure can be conducted flexibly in standard environment of a mobile platform.

Table 4 Evaluation of $d_{MC}$ and CR for calibration to the vehicle frame: SPP requiring 1000 neighboring points inside a sphere with radius of 0.15 m and maximum point distance from origin set to 2.5 m; $\Delta$, $e_{g}$, $e_{f}$ and $e_{e}$, as optimized for extrinsic calibration; $e_{fs}$ in m$^2$, $d_{MC}$ in m.

<table>
<thead>
<tr>
<th>Var. opt.</th>
<th>CV1</th>
<th>CV2</th>
<th>CV3</th>
<th>CV4</th>
<th>CV5</th>
<th>CV6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_{MC}$</td>
<td>1.0</td>
<td>1.5</td>
<td>1.0</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>CR</td>
<td>200</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>$e_{fs}$</td>
<td>0.027</td>
<td>0.034</td>
<td>0.030</td>
<td>0.033</td>
<td>0.029</td>
<td>0.029</td>
</tr>
</tbody>
</table>

results in $e_{fs}$ and visual assessment requiring at least 1000 neighboring points inside a surrounding sphere with a radius of 0.15 m. Additionally, the maximum distance from the origin of the merged cloud is set to 2.5 m according to the dimensions of the vehicle calibrated onto. For TNN we require a high number of nearest neighbors in a considerably larger search area compared to extrinsic calibration. Hence nearly all points remaining inside the merged cloud are depicting parts of the vehicle. The merged point cloud before and after preprocessing is depicted in Figure 7.

As for extrinsic calibration the point cloud containing more information is set as target. This leads to choosing the cloud constructed from the 3D model of the vehicle as target (see section 3 and Figure 8).

With some GICP configurations as well as with some of the assessed data sets not all 6 DOF can be determined correctly. For instance, all data sets with lifted or rather outstretched excavator arm cannot be used to calculate all 6 DOF as LiDAR measurements capture too less of the excavator arm.

For selection of the evaluated GICP parameters, the results from extrinsic calibration are analyzed. The parameters whose variations cause significant changes in calibration are $d_{MC}$ and CR, which is why they are assessed again for vehicle calibration. The results are given in Table 4. It is deduced that for calibration to the vehicle coordinate system the most accurate results applying the standard GICP can be achieved by setting $d_{MC}$ to 1.0 m and CR to 200.

The orientation of all surfaces inside the target cloud are known from the 3D model of the vehicle. Instead of estimating the surface orientation for the target point cloud, the normal information from the 3D model is integrated. Hence determination of the surface orientations using SVD of the covariance matrix is only performed for the preprocessed source cloud. Additionally, the fixed number of corresponding points included in the calculation of the covariance matrix is replaced with specification of a sphere (with radius $r_{N}$) including the points for covariance calculation. The comparison of the results with integration of the normal information from the 3D model and an estimation sphere instead of a fixed number of points is shown in Table 5.

The ground surface is not included in the 3D model of the vehicle but in the merged cloud of the LiDAR sensor data. Either the points representing the ground surface can be eliminated from the merged cloud or a ground surface can
be included into the cloud of the 3D model. For the elimination of the points forming the ground surface the Random Sample Consensus (RANSAC) algorithm can be used to determine points forming surfaces in point clouds after specifying the maximum distance of a point from the estimated model surface \( r_g \). A detailed description of the RANSAC algorithm can be found in Schnabel et al. [18]. The insertion of the ground surface into the cloud of the 3D model rests on the assumption of a smooth ground surface with the same orientation as the tracks of the vehicle. The two approaches are compared with variable \( r_g \) values for the RANSAC algorithm as shown in Table 5. For clarity with 0.01 m and 0.04 m only the \( r_g \) values yielding the most accurate calibration results for the RANSAC approach are given.

For further improvement of the calibration result, the point cloud from the 3D model can be filtered by normal orientation. As normal information extracted from a Collada file contains direction and sign of the normal vector, points with surface normals pointing away from all LiDAR sensors cannot be detected and are removed.

The results given in Table 5 show the most accurate results for the approach including the ground surface in the point cloud deduced from the 3D model, considering the normal information and utilizing an \( r_g \) of 0.40 m in the source cloud. In addition to \( e_f \), the mahalanobis distance \( d_M \) is given for the final calibration results to the vehicle frame. \( d_M \) is calculated from

\[
d_M = \frac{\sum_{i=0}^{N_c-1} d_i^r C_i^{-1} d_i}{N_c}
\]

and constitutes the error metric in the optimization step of the GICP algorithm (eq. 1.). Compared to \( e_f \), \( d_M \) allows assessment of weighted errors between points via the covariance matrix \( C_i \) (eq. 3). For instance, distances between corresponding points inside the same estimated surface are hardly weighted whereas point distances in the direction of normal orientation are weighted strongly, according to the optimization criteria of the GICP algorithm.

Figure 9 depicts the calibration to the excavator frame. As only the excavator arm is captured by the LiDAR sensors, additional information improving the calibration has to be included by inserting a flat ground plane into the 3D model, i.e., a locally flat ground around the excavators position is assumed. Furthermore, an accurate extrinsic calibration, an exact 3D model of the excavator as well as the lowering of the excavator arm, as depicted in Figure 9, are required.

### 5 Conclusion

All calibration steps have been implemented inside the ROS framework and evaluated on the technology demonstrator shown in Figure 1. Preprocessing steps and parameter configurations for the GICP algorithm have been optimized for calibration in unstructured outdoor environment. For extrinsic calibration two surfaces inside the overlap in FOV are required, for instance constituted by the ground surface and an exterior wall of an arbitrary building. Calibration to the vehicle frame works on the basis of an approximately planar ground surface. All 6 DOF can be determined with an accuracy in the range of single digit centimeters, for both extrinsic calibration as well as calibration to the vehicle coordinate system. Compared to calibration via measurements by hand, 3D LiDAR sensors rotated relatively to each other can be calibrated easily and with high accuracy. Besides, a calibration to the vehicle frame is possible, even if its origin lies within the vehicle itself.

---

**Table 5** Calibration to the vehicle frame including preprocessing: Classical GICP surface estimation (SVD of covariance with 100 points) compared to integration of normal information from 3D model; neglecting ground surface (GS) compared to elimination of GS with RANSAC and to inclusion of GS in model cloud; \( r_g \), \( r_N \) in m; \( d_M \) and \( e_f \) in m²; DOF denotes number of approx. correct determined DOF.

<table>
<thead>
<tr>
<th>Approach</th>
<th>neglect GS</th>
<th>RANSAC</th>
<th>include GS</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_g ) for SVD</td>
<td>0.061</td>
<td>0.050</td>
<td>0.044</td>
</tr>
<tr>
<td>DOF</td>
<td>4.7</td>
<td>4.7</td>
<td>5.3</td>
</tr>
<tr>
<td>( d_M )</td>
<td>0.105</td>
<td>0.101</td>
<td>0.089</td>
</tr>
<tr>
<td>( e_f ) for ( r_N = 0.10 )</td>
<td>0.055</td>
<td>0.050</td>
<td>0.040</td>
</tr>
<tr>
<td>DOF</td>
<td>5</td>
<td>5.3</td>
<td>5.3</td>
</tr>
<tr>
<td>( d_M )</td>
<td>0.082</td>
<td>0.086</td>
<td>0.071</td>
</tr>
<tr>
<td>( e_f ) for ( r_N = 0.25 )</td>
<td>0.056</td>
<td>0.046</td>
<td>0.036</td>
</tr>
<tr>
<td>DOF</td>
<td>5</td>
<td>5.3</td>
<td>5.3</td>
</tr>
<tr>
<td>( d_M )</td>
<td>0.084</td>
<td>0.065</td>
<td>0.065</td>
</tr>
<tr>
<td>( e_f ) for ( r_N = 0.40 )</td>
<td>0.059</td>
<td>0.052</td>
<td>0.039</td>
</tr>
<tr>
<td>DOF</td>
<td>4.7</td>
<td>5.3</td>
<td>5.3</td>
</tr>
<tr>
<td>( d_M )</td>
<td>0.097</td>
<td>0.074</td>
<td>0.049</td>
</tr>
</tbody>
</table>

---

**Figure 9** Calibration to the vehicle frame with highest accuracy according to Table 5: Registration of merged cloud from 3D LiDAR sensors (grey, source) to point cloud generated from 3D model including the ground surface (red, target).
6 Acknowledgements

The research has been supported within the project “AKIT – Autonomy KIT for conventional work vehicles to facilitate networked and assisted removal of hazards” (funding code 13N14099), which is being promoted in the course of the announcement “Innovative rescue and security systems” by the Federal Ministry of Education and Research (BMBF) within the scope of the Federal Government’s Research for Civil Security Framework Programme. The authors are grateful for the support.

Literature