SENSORY SYSTEM CALIBRATION METHOD FOR A WALKING ROBOT

Przemysław Labęcki, Dominik Belter

Abstract:
In this paper, a method for calibrating the sensory system of a walking robot is proposed. The robot is equipped with two exteroceptive sensors: a 2D laser scanner and a stereo camera. When using the CAD model of the robot, the positions of both sensors defined in the robot’s body reference frame can only be determined with limited precision. Our goal is to create a method which allows the robot to find the position of the mounted sensors without the need for human input. The presented results show that the method is not only fast but also more precise than calibration using the CAD model.

Keywords: extrinsic calibration, laser scanner, stereo camera, walking robot

1. Introduction

Autonomous robots should perform desired assignments without human support. The capability to acquire information about the environment is crucial if the robot is to plan its motion and interact with objects. Autonomous robots should be also capable to calibrate and re-calibrate its sensory system as humans and animals do. This task is not trivial, since the mechanical mounting of the sensors is imprecise and it can slightly change during normal operation. Moreover, the exact position of sensors’ reference frames is not always precisely defined by manufacturers.

Our six-legged Messor robot (Fig. 1) is equipped with two sensors for terrain measurement. The data obtained from these sensors is used to create two elevation maps of the environment. The robot is located in the center of these maps, and the maps are shifted while the robot is walking. The elevation representation of the surroundings is then used to plan the motion of the robot.

Fig. 1. Messor robot equipped with stereo camera and laser range finder during calibration experiment

The first sensor used for mapping is the 2D Laser Range Finder (LRF) Hokuyo URG-04LX. It is mounted at the front of the robot and tilted down to acquire the profile of the terrain. As the robot walks, a grid-based elevation map is gradually created [1]. The mapping algorithm (described in details in [10]) takes into account the properties of the URG sensor and the discrete nature of the motion of walking robots. The range of the measurements is about 1.2 m because of the geometrical configuration of the system, and it determines the size of the elevation map as 2 × 2 m. The size of the grid cell is 1.5 × 1.5 cm, which is precise enough to appropriately select footholds and to plan movement of the robot (feet and platform paths). On the other hand, the size of the map is not sufficient to plan movement in a longer horizon.

To obtain information about distant terrain and obstacles, the robot is equipped with the Videre Design STOC stereo camera. This camera provides information about the depth of the observed scene and allows to create a strategic elevation map. The strategic map is also shifted as the robot walks, to ensure central location of the robot in the map. The size of the map is set to 10 × 10 m, and the size of the grid cell is 0.1 × 0.1 m. The low precision of this map allows to only roughly plan the path of the robot.

To obtain a high quality elevation map, the sensory system should be precisely calibrated. The goal of the calibration is to find the position of coordinate frames of the sensors (i.e. stereo camera and laser scanner presented in Fig. 2). The coordinate system located in the center of local maps \( O_M \) moves with the robot and the coordinate system \( O_R \) is attached to the robot’s platform (the axes \( x_M \) and \( y_M \) are the same as \( x_R \) and \( y_R \), but \( z_M \) is always 0). The coordinate system attached to the stereo camera and the Hokuyo LRF are \( O_C \) and \( O_S \), respectively. The sought homogeneous transformation to camera frame \( ^B C \) and transition to LRF frame \( ^B S \) can be found by using protractor and ruler or mechanical (CAD) model of the robot. Such measurements are time consuming, should be done carefully and involve human labor. Our goal is to create a method of finding the relation between the coordinate frames of the sensors that does not involve a human operator.

1.1. Related Work

Sensory system calibration methods for mobile robots were proposed many times in the past. Most of the solutions were proposed for wheeled robots. In our research we want to adapt well known methods to calibrate the sensory system of a walking robot. Walk-
ing robots have greater locomotion capabilities than wheeled robots, and these capabilities can be used to simplify the calibration method. We use the most common structure in an urban area - flat surface - to calibrate the sensory system. No known geometrical markers or checkerboards are required for the calibration procedure.

To calibrate two scanners of the six-legged robot Ambler, the calibration targets attached to a single leg are observed [9]. The known position of the markers, computed using the kinematics of the robot, is used to find the transformation to the common coordinate system. In this method the robot should observe its own legs, which is not possible for our Messor robot because of the measurement range of the sensors. The precision of the Hokuyo sensor is very low in ranges lower than 20 cm, and the minimal distance which can be measured using STOC camera is about 40 cm.

Another method used to calibrate a camera and a laser range finder was proposed by Wasielewski and Strauss [16]. A known 3D calibration pattern was observed by two sensors. The resulting measurements of the same object are used to identify transformation between these two sensors. Similarly, an artificial marker like flat checkerboard can be used to calibrate LRF-camera system [18]. Best results are obtained when a triangular checkerboard is used and line features from LRF and a camera are compared to calibrate the sensory system of mobile robot [11]. Both of these methods calculate the transformation between the coordinate systems of the sensors, but do not provide the position of the sensors in an external reference frame attached to the body of the robot.

A minimal closed-form to camera and laser range finder calibration can be found [15, 17]. It is possible to determine a minimal number of poses to properly calibrate the system. However, we acquire many measurements to minimize the role of the measurement errors on the calibration results.

In the proposed method, we take advantage of the properties of the walking robot. The robot can rotate all sensors round each axis of the global coordinate system. Thus, a known geometrical marker is not required during calibration. Similar behavior can be used on humanoid robot with pan-tilt head [3]. Observing a single point feature, the Justin robot calibrates camera extrinsic parameters, elasticities and joint offsets simultaneously.

Calibration process can be also defined as an identification/estimation problem. Then Kalman filtering can be used to find not only sensor-to-sensor transformations but also scene structure and parameters of the Inertial Measurement Unit (IMU) [8]. Here the calibration between camera and IMU is obtained without using a known calibration target. The extrinsic calibration can be also done using Levenberg-Marquardt algorithm [4] or by building nonlinear observer [14].

2. Calibration Method

2.1. Calibration Experiment

In our approach we use a flat floor, the most common and natural object in urban area, to calibrate the sensory system of the robot. As for now, we have not implemented an algorithm that allows the robot to autonomously detect a suitable patch of flat terrain, so the robot is positioned manually. In the future, such detection could be performed using stereo vision data, either directly from the depth image [13] or from the point clouds using a RANSAC method [5]. During the experiment, the flat surface in front of the robot is observed as the position of the robot is modified. 3D data measured in different robot positions by the both LRF and stereo camera are stored.

We tried various combinations of identification motions. Finally, the best results are obtained when the robot performs two simple motions. In the first stage of the calibration experiment (Fig. 3) the robot rotates its body around \( y_R \) axis (angle \( \beta \)). In the second stage the robot rotates around \( x_R \) axis (angle \( \alpha \)). First, the robot looks down (11th sample in Fig. 3), and at the end of the experiment the robot looks up (34th sample in Fig. 3). At each state the robot stores the current position of its body and the data measured by the sensors being calibrated.

The state of the robot during calibration experiment is presented in Fig. 3. The inclination of the robot's platform (\( \alpha, \beta, \gamma \) angles) is measured using an AHRS sensor. The Messor robot is equipped with
2.2. Data Filtering

During the identification experiment the laser scanner returns the position of points that are located on the ground. Position of each point is determined in \( O_S \) coordinate system. Similarly, the stereo camera returns a 3D point cloud. The position of each point is determined in \( O_C \) coordinate system. The data from both sensors are noisy and can be filtered to exclude incorrect measurements.

The properties of both sensors used on the Messor robot are different so we use two different heuristics to filter the data. We can only guarantee that the terrain directly in front of the robot is flat, so we exclude points that are distant from the current robot position. For the stereo camera, we exclude points which satisfy the condition \( x_C > 1 \text{ m} \) and \( y_C > 1 \text{ m} \). Single measurement using stereo camera returns few hundred thousand 3D points, and they are very often noisy and erroneous. We use a RANSAC method [5] to exclude erroneous measurements. We try to fit a flat surface and when the solution is found we exclude all outliers. The LRF is tilted down so in some configurations of the platform the robot measures its own legs. To prevent using incorrect data we exclude points which are too close to the robot. There is also a risk that the points measured by the scanner don’t lie on the ground. Finally, we accept only points from the LRF which fulfill the following requirements:

\[
|x_S| < 0.2 \text{ m}, \quad |x_M| < 0.5 \text{ m}, \quad y_S > 0.2 \text{ m}, \quad r > 0.4, \quad \text{(1)}
\]

where \( r \) is the distance to the point from the origin of the coordinate system.

2.3. Parameters Identification

To find the unknown homogeneous transformations \( R_C \) and \( R_S \) we compute the position of each measured point \( (M_{p_S}^i) \) and \( (M_{p_C}^i) \) for the scanner and camera, respectively) in the \( O_M \) coordinate system:

\[
M_{p_S}^i = R_M \cdot R_S \cdot p_S^i, \quad \text{(2)}
\]

\[
M_{p_C}^i = R_M \cdot R_C \cdot p_C^i. \quad \text{(3)}
\]

In the above equations, \( R_M \) is the homogeneous transformation from the map coordinate system \( O_M \) to the robot coordinate system \( O_R \), whereas \( p_S^i \) and \( p_C^i \) are the positions of measured points in the LRF and stereo camera coordinate systems, respectively.

Because the terrain is flat the \( z \) coordinate of each \( i \)-th point \( M_{p_S}^i \) and \( M_{p_C}^i \) should be zero. This statement is true only when the sensory system is well calibrated. We use this statement to define the fitness function for optimization (calibration) process:

\[
\varepsilon = \frac{1}{N} \sum_{i=0}^{N} z_{M_{p_S}^i}^2. \quad \text{(4)}
\]

where \( N \) is the number of measurements and \( z_{M_{p_S}^i} \) is the \( z \)-position of the \( i \)-th measured point determined in the \( O_M \) coordinate system.

The calibration process is performed separately for LRF and stereo camera. The calibration is defined here as an optimization process:

\[
\arg \min_{R_C, R_S} \sum_{i=0}^{N} z_{M_{p_S}^i}^2. \quad \text{(5)}
\]

We verified two methods to find the transformations \( R_C \) and \( R_S \): Levenberg-Marquardt algorithm (LMA) and Particle Swarm Optimization.

Particle swarm optimization is a population-based stochastic optimization technique inspired by the behavior of a bird flock or a fish school during food searches [6]. At the beginning of the algorithm a population of random particles is created in the search space. Its size is defined by the user and is fixed during the optimization process. In the main loop of the algorithm the particles explore the search space by moving to new positions. Modification of a particle position during one epoch depends on two positions: the first one \( p_{\text{best}}^i \) is the current position of the most fitted particle in the whole swarm, while the second one \( p_{\text{cur}}^i \) defines the best position of the \( i \)-th particle during its lifetime. The position change \( \Delta p_i \) of the \( i \)-th particle is given as:

\[
\Delta p_i := c_1 \cdot \text{rand} \cdot (p_{\text{best}}^i - p_{\text{cur}}^i) + c_2 \cdot \text{rand} \cdot (p_{\text{glob}}^i - p_{\text{cur}}^i), \quad \text{(6)}
\]

where \( p_{\text{cur}}^i \) is the current position of the \( i \)-th particle. The function \( \text{rand} \) returns a random number from 0 to 1. The constants \( c_1 \) and \( c_2 \) are set to 2 because, according to the results shown by Eberhart and Kennedy [6] who investigated the PSO properties, such a value allows the algorithm to converge quickly to a global optimum. The new position change of a particle depends on the previous position change. This is an analogy to inertia, and it decreases the possibility of getting stuck at a local minima. Additionally, to prevent oscillations near the optimum, the maximal position change of a particle is limited to 25% of the maximal range for the considered parameter. The new position \( p_i \) of the \( i \)-th particle is computed as:

\[
p_i := p_i + \Delta p_i. \quad \text{(7)}
\]

The Levenberg-Marquardt algorithm (LMA) [12] is a nonlinear optimization algorithm that interpolates between the Gauss-Newton algorithm and a modified gradient descent method. In this iterative algorithm, in each step the new position vector \( w_{i+1} \) in the search space is calculated as:

\[
w_{i+1} = w_i - (H - \lambda \text{diag}(H))^{-1} \cdot \mathbf{d}. \quad \text{(8)}
\]

In the above equation, \( w_{i+1} \) is the new position vector, \( w_i \) is the current position vector, \( \lambda \) is a weighting factor, \( \mathbf{d} \) is the gradient vector and \( H \) is an approximation of the Hessian matrix, based on the sum of outer products of the gradients. The method seamlessly switches between the Gauss-Newton and the gradient descent methods.
algorithm by changing the $\lambda$ value. If the error has been reduced in the evaluated step, $\lambda$ is decreased and the algorithm is more like the Gauss-Newton method. If, on the other hand, the error increased, $\lambda$ is also increased and the LMA gets more similar to the gradient descent method. Such behavior greatly increases the convergence rate of the algorithm. However, just as any gradient method, LMA is prone to finding local minima.

3. Results

![Fig. 4. Scanner calibration results -- measurements in local coordinate system (A), measurements for initial $R^S$ transition (B), measurements for optimal $R^S$ transition (C)](image)

**Tab. 1. Estimation results for the laser scanner**

<table>
<thead>
<tr>
<th>model</th>
<th>LMA</th>
<th>PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$ [mm]</td>
<td>-7.47</td>
<td>-3.77</td>
</tr>
<tr>
<td>$y$ [mm]</td>
<td>199.14</td>
<td>197.59</td>
</tr>
<tr>
<td>$z$ [mm]</td>
<td>168.83</td>
<td>168.83</td>
</tr>
<tr>
<td>$\alpha$ [$^\circ$]</td>
<td>-44.44</td>
<td>-44.44</td>
</tr>
<tr>
<td>$\beta$ [$^\circ$]</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>$\gamma$ [$^\circ$]</td>
<td>-2.18</td>
<td>-2.18</td>
</tr>
<tr>
<td>$\varepsilon$ [mm]</td>
<td>7.492</td>
<td>7.492</td>
</tr>
</tbody>
</table>

The calibration results for the Hokuyo scanner are presented in Fig. 4 and in Tab. 1. In Fig. 4A one can see the measured points drawn in the reference frame of the laser scanner $O_S$. The measured points transferred to the $O_M$ coordinate system using (2) and the transition $R^S$ defined according to the distances measured using the CAD model of the robot are presented in Fig. 4B. Because the mechanical parts are made with limited precision the measured points don't lie at the level $z_M = 0$. The error $\varepsilon$ for the system calibrated according to the CAD model of the robot is 8.566 mm. When we use the method proposed in this paper the error $\varepsilon$ decreases (Tab. 1). It can be seen in Fig. 4C that after calibration the measured points create more consistent groups at $z_M = 0$. Similarly, the points measured by the stereo camera, defined in the local coordinate system $O_S$ are presented in Fig. 5A. The measured points obtained using (3) before (Fig. 5B) and after calibration (Fig. 5C) are compared. Before the calibration, the measured points are scattered, and the sensor returns erroneous values because the expected position of the camera (measured using the CAD model of the robot) differs from the real position of the camera. After calibration, the measured points create a consistent and horizontal point cloud (Fig. 5C). The error $\varepsilon$ decreases from 4.145 mm for...
a system without calibration to 2.085 mm for system calibrated using the proposed method.

**Table 2. Estimation results for the stereo camera**

<table>
<thead>
<tr>
<th>model</th>
<th>LMA</th>
<th>PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>x[mm]</td>
<td>-30.00</td>
<td>-29.15</td>
</tr>
<tr>
<td>y[mm]</td>
<td>180.00</td>
<td>179.63</td>
</tr>
<tr>
<td>z[mm]</td>
<td>230.00</td>
<td>223.19</td>
</tr>
<tr>
<td>α[°]</td>
<td>-115.00</td>
<td>-114.49</td>
</tr>
<tr>
<td>β[°]</td>
<td>0.00</td>
<td>1.04</td>
</tr>
<tr>
<td>γ[°]</td>
<td>0.00</td>
<td>-2.28</td>
</tr>
<tr>
<td>ε[mm]</td>
<td>4.145</td>
<td>2.085</td>
</tr>
</tbody>
</table>

### 3.1. Mapping Experiment

The proposed calibration method was tested in a terrain map building experiment. Two maps were simultaneously built from the LRF and the stereo vision data. The $R_S$ and $R_C$ transformations were used to calculate the position of the measured points in the map coordinate system, according to equations (2) and (3). As the robot moved, new measurements were added to the point clouds from both the LRF and stereo vision camera. The motion of the robot between measurements was estimated using a Parallel Tracking and Mapping (PTAM) method, originally described in [7], enhanced with the information from the Inertial Measurement Unit installed on the robot [2].

The results are shown in Fig. 6. When the system was calibrated using the CAD model, the point cloud from the LRF (shown in light gray) was poorly aligned to the point cloud acquired with stereo vision (shown in dark gray), which is particularly visible in the area corresponding to the boxes (marked as 1 in Fig. 6A) and in the border of the terrain mockup (marked as 2). The proposed calibration method provided more precise alignment, as presented in Fig. 6B. To assess the improvement, a coefficient that represents the alignment in a quantitative way was used. Separate elevation raster maps were built from the LRF and stereo vision point clouds. In each cell of the maps, the elevation value was determined using the points with the highest $z_M$ value. The resulting maps are presented in Fig. 7. For the overlapping area of both maps, an error value was computed using the following equation:

$$\text{err} = \frac{\sum_{i=0}^{N_x} \sum_{j=0}^{N_y} \left| z_{S}^{i,j} - z_{C}^{i,j} \right|}{N_o},$$  

where $i$ and $j$ determine the cell in the raster elevation maps, $z_{S}^{i,j}$ and $z_{C}^{i,j}$ are the elevation values in the $i,j$ cell of the scanner and stereo map, respectively, and $N_o$ is the number of cells in the overlapping area. The proposed calibration method reduced this average error to 0.014 m, from the value of 0.041 m that was calculated for the system calibrated using the CAD model.

### 3.2. Stability of the Calibration Results

The LMA, being a deterministic algorithm, always converges to the same result given the same starting point. However, as every gradient optimization method, it is prone to getting stuck in local minima. As the distance from the starting point to the true parameters increases, the risk of terminating at a local minimum also increases. Because in the presented calibration task the starting point for the LMA may not always be determined accurately, some experiments...
were performed. The starting point was randomly selected by adding noise to the estimated values shown in tables 1 and 2. The added noise value was drawn from a normal distribution with zero mean and a standard deviation \( \sigma_d \) for distances and \( \sigma_a \) for angles. The calibration was then performed 50 times, each time from a different starting point. The results of three experiments, with increasing standard deviations, are shown in tables 3 and 4. For both the laser scanner and stereo camera calibration, the algorithm reached local minima only with starting points located far from the true parameters. Within reasonable error boundaries (\( \sigma_d = 0.1 \) and \( \sigma_a = 10^\circ \)), the algorithm still managed to find the parameters every time.

**Tab. 3. Standard deviation of the estimated parameters for the laser scanner**

<table>
<thead>
<tr>
<th></th>
<th>( \sigma_d = 0.01 \text{ m} )</th>
<th>( \sigma_d = 0.1 \text{ m} )</th>
<th>( \sigma_d = 0.2 \text{ m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_a = 5^\circ )</td>
<td>0.000004</td>
<td>0.000004</td>
<td>0.018</td>
</tr>
<tr>
<td>( \sigma_a = 10^\circ )</td>
<td>0.000007</td>
<td>0.000005</td>
<td>0.16</td>
</tr>
<tr>
<td>( \sigma_a = 50^\circ )</td>
<td>0.000003</td>
<td>0.000002</td>
<td>0.41</td>
</tr>
<tr>
<td>( \sigma_x [\text{m}] )</td>
<td>0.000036</td>
<td>0.00024</td>
<td>134.61</td>
</tr>
<tr>
<td>( \sigma_y [\text{m}] )</td>
<td>0.000053</td>
<td>0.000057</td>
<td>134.61</td>
</tr>
<tr>
<td>( \sigma_z [\text{m}] )</td>
<td>0.00035</td>
<td>0.00044</td>
<td>108.88</td>
</tr>
</tbody>
</table>

**Tab. 4. Standard deviation of the estimated parameters for the stereo camera**

<table>
<thead>
<tr>
<th></th>
<th>( \sigma_d = 0.01 \text{ m} )</th>
<th>( \sigma_d = 0.1 \text{ m} )</th>
<th>( \sigma_d = 0.2 \text{ m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_a = 1^\circ )</td>
<td>0.000001</td>
<td>0.000007</td>
<td>0.083</td>
</tr>
<tr>
<td>( \sigma_a = 10^\circ )</td>
<td>0.000005</td>
<td>0.000004</td>
<td>0.55</td>
</tr>
<tr>
<td>( \sigma_a = 50^\circ )</td>
<td>0.000004</td>
<td>0.000004</td>
<td>0.44</td>
</tr>
<tr>
<td>( \sigma_x [\text{m}] )</td>
<td>0.0004</td>
<td>0.00040</td>
<td>38.16</td>
</tr>
<tr>
<td>( \sigma_y [\text{m}] )</td>
<td>0.00014</td>
<td>0.00013</td>
<td>72.33</td>
</tr>
<tr>
<td>( \sigma_z [\text{m}] )</td>
<td>0.00094</td>
<td>0.00058</td>
<td>34.11</td>
</tr>
</tbody>
</table>

4. Conclusions

In this paper, we propose a new method for calibrating the sensory system of the robot. The calibration methodology is valid for walking robots which can modify the inclination of the body in relation to a flat surface. Using the assumption that the terrain in front of the robot is flat and the position of the robot is known we determine the position of the sensors in the robot’s body reference frame.

We verified two methods to calibrate the sensory system. Both methods find a set of parameters which determine the position of the sensors and reduce the defined calibration error. However, the Levenberg-Marquardt algorithm returns smaller calibration error. Moreover, the optimization time is several minutes when LMA is used. Because the execution time is short, the method can be used whenever the robot starts an operation or there is a risk that sensors changed their position (because of an accident or collision with an obstacle).

The PSO reduces the calibration error but the results are slightly worse. The execution time is about 15 minutes for the laser scanner and about one hour for estimating the parameters of the stereo camera (the population size is 1000 and maximal number of epochs is 10). For that reason the PSO method can be used only off-line.

In future we are going to use the proposed method to calibrate various exteroceptive 3D sensors like Kinect and SwissRanger. We also want to implement a method for detecting flat surfaces on the ground, so that the calibration procedure could be performed fully autonomously.

**AUTHORS**

Przemysław Łabęcki - Poznań University of Technology, Institute of Control and Information Engineering, ul. Piotrowo 3A, 60-965 Poznań, Poland, e-mail: Przemyslaw.Labecki@cie.put.poznan.pl.

Dominik Belter* - Poznań University of Technology, Institute of Control and Information Engineering, ul. Piotrowo 3A, 60-965 Poznań, Poland, e-mail: Dominik.Belter@put.poznan.pl.

*Corresponding author

**ACKNOWLEDGEMENTS**

This work was supported by NCN grant no. 2011/01/N/ST7/02080.

**REFERENCES**


