Extrinsic RGB-D camera calibration for legged robots

Conference Paper

Author(s):
Höpflinger, Mark; Remy, Christian D.; Hutter, Marco; Siegwart, Roland

Publication date:
2011

Permanent link:
https://doi.org/10.3929/ethz-a-010027510

EXTRINSIC RGB-D CAMERA CALIBRATION FOR LEGGED ROBOTS

MARK A. HOEPFLINGER, C. DAVID REMY, MARCO HUTTER, ROLAND SIEGWART*

Autonomous Systems Lab, Institute of Robotics and Intelligent Systems, Swiss Federal Institute of Technology (ETHZ), Zurich, Switzerland,

markho@ethz.ch

This paper describes a method to identify the extrinsic parameters of RGB-D cameras mounted on legged robots. Since the calculation of the parameters is based on the detection of the robots feet in the camera images, no special calibration objects are required. Therefore the method is simple to use and can even be applied automatically. Experiments demonstrate the precision and the robustness of the method.

1. Introduction

The main advantages of legged robots is their ability to move in very rough and unstructured terrain where wheeled systems usually would fail. While simple execution of stable walking patterns can already cope with substantial variation in the ground level, careful foothold selection becomes indispensable when it comes to large obstacles or gaps in the ground. Successful step planning in such an environment requires accurate knowledge of the robot position as well as an appropriate map of the environment. Within the DARPA Little Dog Challenge, several groups have impressively pushed the state of the art in traversing highly unstructured terrain [1,2,3]. However, they could rely on accurate maps of the environment that were known before execution as well as on real time state observation by external sensors – information that is not available in most real application.

Along the lines of the little dog challenge, we developed ALoF [4], an improved quadruped robot that has enough payload capability to carry on board sensors for localization and mapping. The appropriate usage of such sensors (for example cameras, laser scanners or inertial measurement units) requires the identification of its mounting position and orientation. This identification is known as extrinsic camera calibration problem. Common methods for extrinsic calibration of RGB cameras are often based on the solution of a point correspondence problem, where the correspondence between points with known

* This research was supported by the Swiss National Science Foundation through the National Centre of Competence in Research Robotics.(NCCR).

3D coordinates and their projection in the image is found. Even if specially designed calibration objects simplify the usage of such methods [5], they are usually still very impractical and involve manual interaction. Other methods, that could in principle be applied to extrinsically calibrate the camera, are based on the detection of geometrical patterns, like lines, circles or ellipses [6, 7], but require multiple views of the pattern. Further methods use motion analysis for the calibration (such as the method developed by F. Pomerleau et al, based on [8]) and are difficult to be applied on legged robots.
Therefore this paper describes a method for the extrinsic calibration of RGBD cameras mounted on legged robots. Because the positions of the front robot feet with known 3D coordinates and their corresponding position in the RBG image are used to identify the mounting pose, no additional calibration objects or sensors and no manual interaction are required to identify the 6 degrees of freedom of the camera.

2. System Description

The RGB-D camera (Microsoft Kinect for Xbox 360) is mounted on ALoF, a four legged robot with a total weight of about 15 Kg and a length of half a meter [4]. Each of the legs has three degrees of freedom. The mechanical design allows hip abduction/adduction, hip flexion/extension as well as knee flexion/extension with a large range of motion (Table 1). This enables the execution of special maneuvers like crawling, recovering from tipping over, but also setting the robot in the calibration configuration (Fig 1) without manual interaction. In this configuration, the front legs are stretched, so that the spherical feet, made of blue rubber material, are visible by the RGB-D camera.
The calibration method has been implemented as a ROS [9] node and applies functions of the OpenCV [10] and ROS PCL libraries.

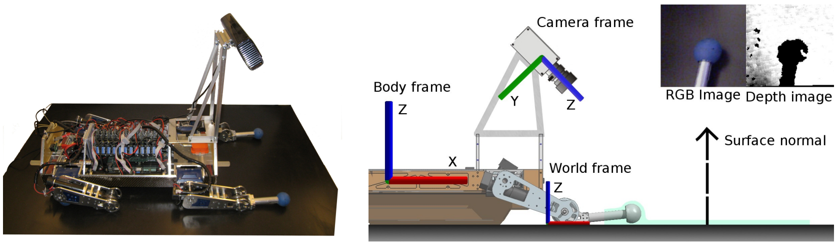

Fig. 1: Left: Picture of ALoF in the calibration configuration (the robot is placed on leveled ground with the front feet in the range of the RGB-D camera).
Right: Introduction of the coordinate frames used for the calibration.

Parameter	Value	Parameter	Value
Overall Length	557 mm	Total Mass	15 Kg
Typical Height	400 mm	Main Body Mass	10 Kg
Overall Width	386 mm	Range of motion for hip flexion/extension	±180 deg
Length of Thigh Segment	150 mm	Range of motion for knee flexion/extension	+90/-160 deg
Length of Shank Segment	150 mm	Range of motion for hip abduction/adduction	+45/-90 deg
Foot diameter	40 mm		

Table 1: Hardware characteristics of the walking robot ALoF

3. Method

The approach presented in this paper allows to identify the extrinsic camera parameters without manual interaction or the need of any special calibration object. Instead, the robot feet are used as reference point with known 3D position. Together with an estimation of the surface normal of the leveled ground in front of the robot, all the extrinsic camera parameters can be determined.
It is assumed, that the intrinsic camera calibration is done in prior and that the RBG and the depth image are aligned correctly. The current implementation of the method is also based on assumptions about the actual robot (such as the shape and color of the robot feet), but the principle can be applied whenever parts of the robot at known positions can be tracked precisely in an image provided by the RGB camera.
The calibration process can be separated in two parts. In the first part, the camera position and orientation with respect to the world frame is identified. This allows to estimate the camera height above ground and its roll and pitch angle. In a second step, the camera pose in the robot body frame is computed.

The pitch and the roll angle of the camera could have been measured by the three axis accelerometer (Kionix KXSD9) that is incorporated in the Kinect. But, on one hand, the exact pose of the accelerometer inside the Kinect device was unknown and on the other hand, not all RGB-D sensors incorporate additional accelerometers. Therefore, our calibration method does not rely on accelerometer measurements. Instead, estimates of the surface normal of the leveled ground in front of the robot are used to identify the camera roll and pitch angles as well as the height over leveled ground.

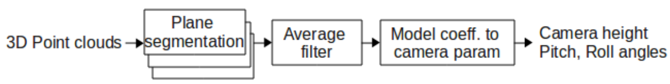

Fig. 2: Diagram illustrating the steps to compute the camera height above ground and its roll and pitch angle

The estimation of the surface normal is done by matching planes (RANSAC plane matching algorithm [9]) to a set of point clouds acquired from the depth sensor. The computed coefficients of the surface normal are then averaged and applied to calculate the camera pose in the world frame.

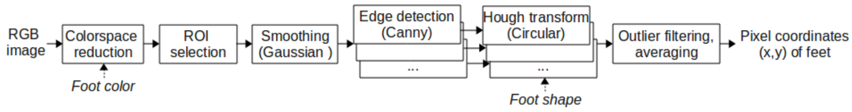

Fig. 3: Diagram illustrating the steps to compute the pixel coordinates of the feet in the RGB image

The assessment of the displacement within the robot body frame is based on the search of the point correspondence of the known robot foot position and its 2D projection in the image. Because of the limited accuracy, precision and eventually sparse depth data at the foot location, precise foot tracking was not possible using only the depth data. Therefore the tracking has been performed on RGB images.

First, under the assumption of having blue feet, the color-space of the RGB images has been reduced. Then the region of interest (ROI) has been set to the lower half of the image where the feet are visible in the calibration configuration of the robot. To lower noise effects, the image has been smoothed with a 5x5 pixel wide Gaussian kernel. Finally, canny edge detection with a set of different thresholds was performed and a circular hough transform was applied to detect the spherical feet in the image. The set of detected circles possibly contained false positives, which, in a next step, are removed by a statistical outlier detection filter. Finally, from an average of the inlier, the estimated foot position in pixel coordinates has been computed (Fig. 3).

Knowing the foot geometry, it would have been possible to compute the remaining camera parameters without using depth image data. In this approach the 2D points from the RGB image are projected to the depth image. Because the depth information in the range of the feet was sparse (Fig. 5, left), the computed 2D position could not directly be mapped to the point cloud data. Instead, linearly distributed sample points through the center of the detected feet in the image have been projected to the point cloud (Fig. 5, right). Using a linear regression (ordinary least squares) on the co-linear sample points, the 3D foot position could be calculated.

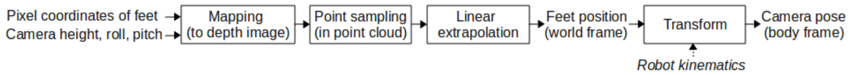

Fig. 4: 3: Diagram illustrating the steps to compute the camera pose in the robot body frame

The foot position served to compute the yaw angle of the camera in the robot body frame and finally, considering the known kinematics of the robot to identify the position of the body frame with respect to the camera frame (Fig. 5).

4. Experiments

To evaluate the performance of the camera calibration, experiments have been performed on three different artificial terrain types. The robot has been placed on a homogeneous white and black surface and on a floor with a textured wooden surface (Fig. 5, left side). The white surface was selected to test the calibration under conditions where the reflectivity as well as the contrast between the robot feet and the surroundings compared to the other surfaces is large. For each type of surface, the calibration routine has been executed 100 times.
Due to the missing specification of the geometry of the Kinect sensor, the position of the camera within the housing could only be estimated. Therefore, the camera position has been measured with a measuring tape (1 mm resolution), which resulted in a ground truth with low accuracy. The pitch and the roll angle of the camera have been quantified by using the Kinect accelerometers. Again, missing specifications about the mounting position/orientation and the application of a sensor that is not very accurate resulted in a imprecise reference.

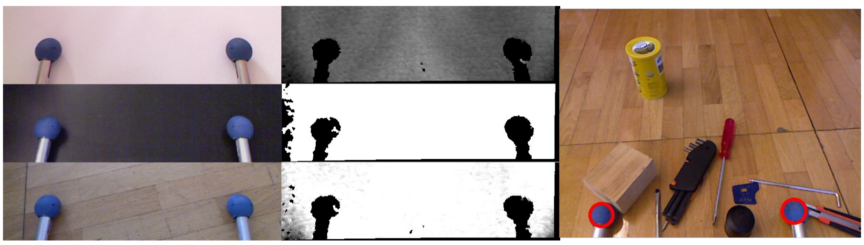

Fig. 5: Left side: Sample images of the different surfaces aquired from the RGB camera.
Middle images: The corresponding depth images (the color has been normalized for visualization).
Right side: Sample image of the foot detection in a disturbed scene.

5. Results:

The following box plots (Fig. 6 and Fig. 7) illustrate the lower/upper quartile (first and third quartile) and the median of the samples. The whiskers of the plot point out the samples that are still within +/- 1.5 times the interquartile range (difference of the third and the first quartile). Samples outside this range are marked as outliers.

Fig. 6 shows plots of the identified camera parameters, that are based on the calculation of the surface normal of the terrain. The computed camera height above ground varies most for the calibrations performed on the white surface (about +/- 1.25 mm). The most precise estimations have been achieved on the black surface (50% of the samples are within about +/- 0.12 mm around the median). Even when considering the outliers, the samples vary by less than approx. +/- 0.75 mm from the median value.

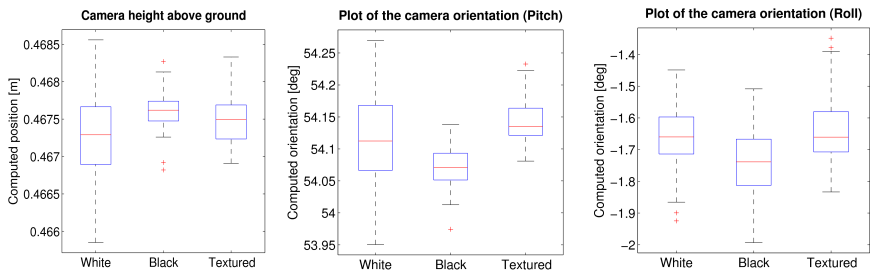

Fig. 6: Box diagram of the identified camera height above ground as well as pitch and roll angle.

The plot in the middle of Fig. 6 presents the computed pitch angle. Again, the estimations based on measurements on the white surface are the most imprecise. Still, 50% of the samples are within +/- approx. 0.05 degrees around the median. While the values measured on the black surface are distributed more symmetrically, the precision is comparable to the one of the measurements performed on the textured surface. The precision of the roll angle measurements is nearly the same for the different surfaces (50% of the samples are within about +/- 0.06 deg around the median, the maximum variation is approx. 0.5 deg).

The median value of the samples differs for the measurements on the discriminative surfaces. The maximal difference is less than 0.1 degrees for the computed angles and 0.5 mm for the height.

The identification of the other extrinsic parameters is based on estimations of the foot position within the RGB image. Therefore it is expected that the optical properties of the surface have an even stronger influence on the quality of the estimation.

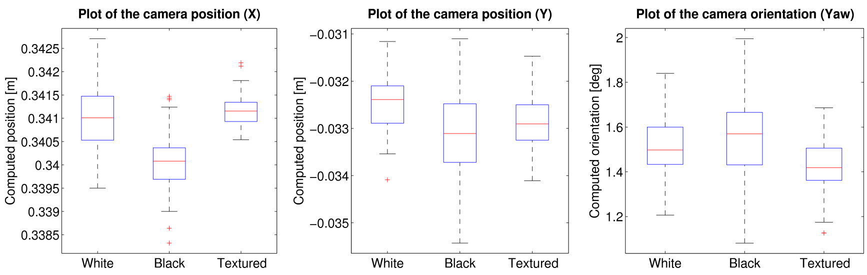

Fig. 7: Box diagram of the identified camera position (X/Y) and yaw angle (robot body frame).

The evaluation of the computed position is visible in Fig. 7, in the left and middle plot. The precision of the measurements performed on the textured surface was the highest. For the X coordinate of the position, the standard deviation is approximately 0.3 mm. Measurements recorded on the black surface where the most imprecise: The biggest variation of all 100 measurements was about 3 mm (and up to 4.3 mm for the estimation of the y component). This also resulted in imprecise yaw estimations (Maximal difference of all samples including outliers: 0.91 degrees for the black surface, 0.56 degrees for the textured surface, and 0.63 degrees for white surface). Still, the standard deviation of the measurements taken on the white or the textured surface is less than 0.13 degrees.

The accuracy of the results seems to be in a reasonable range (Table 2), but due to imprecise and missing reference values, no quantitative comparison was possible.

It can be assumed that the precision of the calibration method could even be improved by the means of applying constant illumination conditions, adding calibration markers to the feet, incorporating additional sensors or simply by processing more RGB-D images per calibration step.

	Mean			**Standard deviation**			**“Ground truth”**
	White	Black	Textured	White	Black	Textured	
X [mm]	341.1	340.1	341.2	0.657	0.575	0.342	340.0
Y [mm]	-32.5	-33.1	-32.8	0.552	0.942	0.587	-30.0
Z [mm]	*402.3*	*402.6*	*402.5*	*0.572*	*0.247*	*0.318*	*400.0*
Roll [deg]	*-1.66*	*-1.74*	*-1.63*	*0.096*	*0.107*	*0.109*	*-1.30*
Pitch [deg]	*54.11*	*54.07*	*54.14*	*0.071*	*0.029*	*0.032*	*54.60*
Yaw [deg]	1.51	1.56	1.42	0.113	0.168	0.126	n.a.

Table 2: Mean and standard deviation of the identified camera parameters, and the 'ground truth'..

6. Conclusion

The paper presented a method to calibrate the extrinsic camera parameters of a RGB-D camera mounted on a legged robot. The calibration is performed while the robot is placed on leveled ground and requires no additional equipment.

Experiments on three different ground surfaces indicate the precision of the method. Standard deviations of around 0.5 mm for the camera position estimations and of about 0.1 degrees for the camera angles indicate a good precision for calibrations performed on the employed white and textured surface.

References

1. J. R. Rebula, P. D. Neuhaus, B. V. Bonnlander, M. J. Johnson, and J. E. Pratt, “A controller for the littledog quadruped walking on rough terrain,” in Proceedings of the IEEE International Conference on Robotics and Automation, 2007
2. J. Buchli, M. Kalakrishnan, M. Mistry, P. Pastor, and S. Schaal, “Compliant quadruped locomotion over rough terrain,” in IEEE International Conference on Intelligent Robots and Systems, 2009
3. J. Zico Kolter, Mike P. Rodgers, and Andrew Y. Ng, “A control architecture for quadruped locomotion over rough terrain”, in IEEE International Conference on Robotics and Automation, 2008, pp. 811-818
4. C. D. Remy, O. Baur, M. Latta, A. Lauber, M. Hutter, M. Hoepflinger, C. Pradalier, R. Siegwart, “Walking and crawling with ALoF – A robot for autonomous locomotion on four legs”, in 12th International Conference on Climbing and Walking Robots , 2010
5. J. A. Hesch, A. I. Mourikis, and S. I. Roumeliotis, “Mirror-Based Extrinsic Camera Calibration”,
6. X. Meng and Z. Hu, “A New Easy Camera Calibration Technique Based on Circular Points”, Pattern Recognition, Vol. 36, pp. 1155-1164, 2003
7. C. Yang, F. Sun and Z. Hu: “Planar Conic Based Camera Calibration”, ICPR, 2000
8. F. Pomerleau, S. Magnenat, F. Colas, M. Liu and R. Siegwart, “Tracking a Depth Camera: Parameter Exploration for Fast ICP”, submitted to International Conference on Robotics and Automation, 2011
9. http://www.ros.org
10. http://sourceforge.net/projects/opencvlibrary/