


Foot-Eye Calibration of Legged Robot Kinematics

Conference Paper**Author(s):**

Blöchliger, Fabian; Blösch, Michael; Fankhauser, Péter; [Hutter, Marco](#) ; Siegart, Roland

Publication date:

2016-09

Permanent link:

<https://doi.org/10.3929/ethz-a-010655381>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

https://doi.org/10.1142/9789813149137_0050

FOOT-EYE CALIBRATION OF LEGGED ROBOT KINEMATICS

FABIAN BLOEHLIGER, MICHAEL BLOESCH, PÉTER FANKHAUSER,
MARCO HUTTER and ROLAND SIEGWART

*Autonomous Systems Lab (ASL),
ETH Zurich, Switzerland
{blfabian,bloeschm,pfankhauser}@ethz.ch
www.asl.ethz.ch*

Legged robots rely on an accurate calibration of the system's kinematics for reliable motion tracking of dynamic gaits and for precise foot placement when moving in rough terrain. In our automatic foot-eye calibration approach, a monocular camera is attached to the robot and observes the robot's moving feet, which are equipped with Augmented Reality (AR) markers. The measurements are used to formulate a non-linear least squares problem over a fixed time window in order to estimate the 33 unknown parameters. This is efficiently solved with the Levenberg-Marquardt algorithm and we get estimates for both the kinematic and the camera parameters. The approach is successfully evaluated on a real quadruped robot.

Keywords: Calibration, Kinematics, Robot vision, Legged robotics

1. Introduction

Legged locomotion heavily depends on a good knowledge of the robot's kinematic parameters. These include body measures, segment lengths and joint angle offsets. Small errors propagate through the kinematic chain and significantly affect the pose of the foot. It is desirable to reduce these systematic errors, which can be achieved through geometric calibration of the robot's kinematics.¹ In order to avoid tedious and error-prone manual calibration, we can employ an automatic calibration procedure. To this end, a redundant measure of the robot's foot location is required which is then compared to the kinematic model. The resulting discrepancy depends on the calibration parameters which can be optimized such that the associated errors become minimal. An elegant and inexpensive way to get measurements of the robot's foot is a camera which is rigidly attached to the robot's

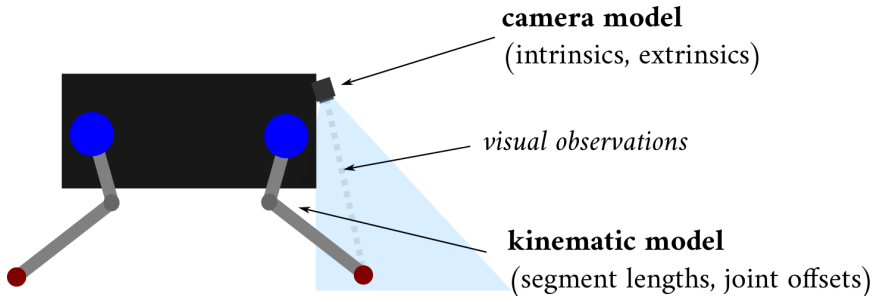


Fig. 1. Overview of the *foot-eye calibration* problem. By minimizing the discrepancy between the two models and the visual observations, the corresponding calibration parameters can be estimated (in brackets).

main body. For industrial robot arms and for humanoid robots, this is commonly referred to as *hand-eye calibration*.^{2,3} Because it is often difficult to directly extract the 3D pose information of the end-effector,⁴ a common practice is to attach checkerboard patterns to the end-effector.⁵ Recent approaches of visual calibration procedures do not only consider geometric parameters, but aim to provide a more general calibration framework with the inclusion of different sensor modalities.^{6,7}

In this paper, we seek to contribute to the calibration of legged robots by adapting the method of Birbach *et al.*⁸ to a quadruped robot. Our framework allows to include an arbitrary number of legs simultaneously in a single calibration run. We include all segment lengths as well as a constant measurement time offset in the calibration process, which we show to significantly improve the result. We achieve a fast optimization runtime (10–20 seconds) by the use of efficient transformation descriptions and a highly optimized solver library.

2. Modeling

Figure 1 shows the basic components of our investigated *foot-eye calibration*: The *kinematic model* describes the leg’s forward kinematics with respect to the robot’s base frame. The *camera model* is used to express the relation between the detected 3D marker locations and their projections onto the image plane. The parameters of the estimation process are summarized in Table 1. We distinguish between kinematic parameters \mathbf{p}_{kin} and camera parameters \mathbf{p}_{cam} . Additionally, we obtain the joint encoder data $\mathbf{z}_j(t)$ at time t (positions and velocities) and the 2D image marker po-

Table 1. Summary of the calibration parameters. The locomotion relevant kinematic calibration parameters are printed in bold.

Kinematic parameters \mathbf{p}_{kin}		Camera parameters \mathbf{p}_{cam}	
Body measures	b_x, b_y	Position	$\mathbf{t}_C^{\mathcal{B}}$
Segment lengths*	l_h, l_t, l_s	Orientation	$\mathbf{q}_C^{\mathcal{B}}$
Joint offsets*	$\theta_{h,\text{offset}}, \theta_{t,\text{offset}}, \theta_{s,\text{offset}}$	Camera matrix	K
Shank-to-foot angle*	θ_f	Distortion coeff.	D
Foot-to-pattern angle*	θ_p		
Time offset	t_{offset}		

Note: * Estimated separately for each leg.

sitions $\mathbf{z}_m(t)$. We assume that all measurements are affected by additive Gaussian noise. A constant time delay t_{offset} can be observed between the joint encoder data and the marker measurements, which is therefore also included in the calibration. By linking kinematic and camera model, we obtain an estimate of the marker location in the image plane.

2.1. Kinematic model

The kinematic chain from the base frame \mathcal{B} to the foot frame \mathcal{F} is shown in Fig. 2. The transform $T_{\mathcal{F}}^{\mathcal{B}}$ is obtained by concatenating subsequent transforms between the intermediate leg coordinate frames, which are located at the hip (\mathcal{H}), the thigh (\mathcal{T}) and the shank (\mathcal{S}). Each transform is parametrized by a translation vector \mathbf{t} and a subsequent rotation \mathbf{q} (unit quaternion). Taking the transform $T_{\mathcal{H}}^{\mathcal{B}}$ (left leg) as an example, this is a translation by $\mathbf{t}_{\mathcal{H}}^{\mathcal{B}} = (b_x, b_y, 0)^T$ and a rotation by $\mathbf{q}_{\mathcal{H}}^{\mathcal{B}} = (\cos(\theta_h(t)/2), \sin(\theta_h(t)/2) \cdot [1, 0, 0])^T$. In the following, we will only indicate the dependence on model parameters for each transform. The resulting transform from the base frame to the foot frame is

$$T_{\mathcal{F}}^{\mathcal{B}}(\mathbf{p}_{\text{kin}}, \mathbf{z}_j(t)) = T_{\mathcal{H}}^{\mathcal{B}}(\theta_h(t), b_x, b_y) T_{\mathcal{T}}^{\mathcal{H}}(\theta_t(t), l_h) T_{\mathcal{S}}^{\mathcal{T}}(\theta_s(t), l_t) T_{\mathcal{F}}^{\mathcal{S}}(\theta_f, l_s), \quad (1)$$

where an estimate of the j -th joint angle is given by

$$\theta_j(t) = \theta_{j,\text{measured}}(t) + t_{\text{offset}} \dot{\theta}_{j,\text{measured}}(t) + \theta_{j,\text{offset}}. \quad (2)$$

This includes two correction terms, taking into account the time offset t_{offset} and the joint angle offset $\theta_{j,\text{offset}}$. We rigidly attach circularly symmetric marker patterns to the robot to get visual observations of its feet (see Fig. 2). The transform $T_{\mathcal{P}}^{\mathcal{F}}$ describes the rotation of the circular pattern around the foot's z -axis, and the translation vector $\mathbf{t}_{\mathcal{M}}^{\mathcal{P}}$ is the center of

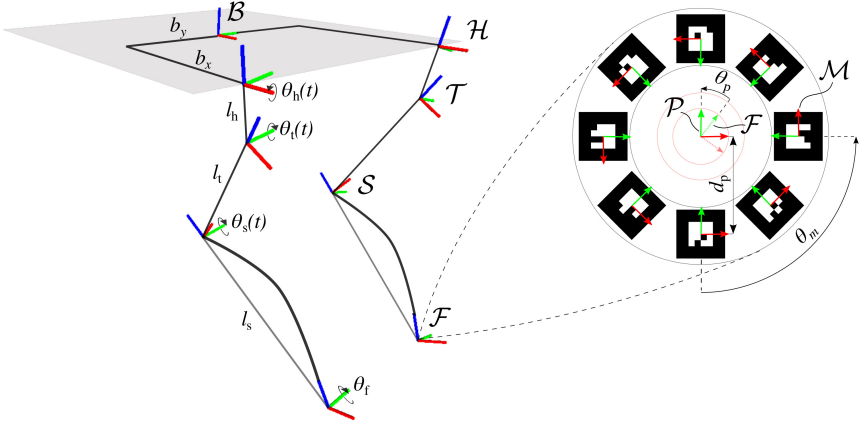


Fig. 2. The coordinate frames of the legged robot and of the circular marker pattern (red: x -axis, green: y -axis, blue: z -axis). The location of a specific marker m is determined by its known rotation θ_m and the radial offset d_p relative to the circular marker pattern center. The circular marker pattern is not required to be aligned with the foot coordinate frame since the angular offset θ_p is co-estimated.

a specific marker frame \mathcal{M} with respect to the circular pattern frame \mathcal{P} . Finally, with respect to the robot's base frame, this translation vector is

$$\mathbf{t}_{\mathcal{M}}^{\mathcal{B}}(\mathbf{p}_{\text{kin}}, \mathbf{z}_j(t)) = T_{\mathcal{F}}^{\mathcal{B}}(\mathbf{p}_{\text{kin}}, \mathbf{z}_j(t)) T_{\mathcal{P}}^{\mathcal{F}}(\theta_p) \mathbf{t}_{\mathcal{M}}^{\mathcal{P}}(\theta_m, d_p). \quad (3)$$

2.2. Camera model

In order to describe the location of a marker in the image plane, we first express the translation vector from (3) in the camera frame \mathcal{C} , using the extrinsic camera parameters. This gives us

$$\mathbf{t}_{\mathcal{M}}^{\mathcal{C}}(\mathbf{p}_{\text{kin}}, \mathbf{p}_{\text{cam}}, \mathbf{z}_j(t)) = T_{\mathcal{C}}^{\mathcal{B}}(\mathbf{t}_{\mathcal{C}}^{\mathcal{B}}, \mathbf{q}_{\mathcal{C}}^{\mathcal{B}})^{-1} \mathbf{t}_{\mathcal{M}}^{\mathcal{B}}(\mathbf{p}_{\text{kin}}, \mathbf{z}_j(t)). \quad (4)$$

We use the radial-tangential distortion model with three radial distortion parameters to simulate the distortion of the lens, and the pinhole camera model for the projection of the marker center location onto the image plane. The normalized image projection coordinates of $\mathbf{t}_{\mathcal{M}}^{\mathcal{C}}$ are

$$\mathbf{x}_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} \mathbf{t}_{\mathcal{M},x}^{\mathcal{C}} / \mathbf{t}_{\mathcal{M},z}^{\mathcal{C}} \\ \mathbf{t}_{\mathcal{M},y}^{\mathcal{C}} / \mathbf{t}_{\mathcal{M},z}^{\mathcal{C}} \end{pmatrix}, \quad (5)$$

and applying the distortion vector $D = (k_1, k_2, p_1, p_2, k_3)$ yields

$$\mathbf{x}_d = \begin{pmatrix} u_d \\ v_d \end{pmatrix} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \mathbf{x}_n + \begin{pmatrix} 2p_1 u_n v_n + p_2 (r^2 + 2u_n^2) \\ p_1 (r^2 + 2v_n^2) + 2p_2 u_n v_n \end{pmatrix}, \quad (6)$$

with $r^2 = u_n^2 + v_n^2$. Finally, we obtain the modeled pixel position of a marker on the image plane as

$$\mathbf{M}(\mathbf{p}_{\text{kin}}, \mathbf{p}_{\text{cam}}, \mathbf{z}_j(t)) = \begin{pmatrix} f_x u_d + c_x \\ f_y v_d + c_y \end{pmatrix}, \quad (7)$$

with the principal point coordinates $(c_x, c_y)^T$ and the focal lengths (including the pixel spacings) f_x and f_y from the camera matrix K .

3. Calibration

The whole calibration framework is implemented within ROS (Robotic Operating System). We perform the calibration by moving the feet of the robot within the field of view of the camera. For the image acquisition, we use a Point Grey Chameleon 3 camera with a wide-angle lens. The ARToolKit marker detector⁹ finds a certain marker m in the image. Due to the unique ID of this marker, the foot allocation is trivial. We get the corresponding joint encoder data according to the marker’s time stamp and ID and estimate the modeled marker position according to (7). The calibration is done by finding the model parameters which maximize the likeliness of the observations, i.e. we try to minimize the error between the modeled and measured image marker position. To this end, the residual

$$\mathbf{f}_{m,k}(\mathbf{p}_{\text{kin}}, \mathbf{p}_{\text{cam}}) = \mathbf{z}_m(t_k) - \mathbf{M}(\mathbf{p}_{\text{kin}}, \mathbf{p}_{\text{cam}}, \mathbf{z}_j(t_k)), \quad (8)$$

describes the difference between the detected and the modeled position of marker m at time t_k . After collecting the residuals for all detected markers over a finite time horizon, we get the calibrated parameters $(\mathbf{p}_{\text{kin}}^*, \mathbf{p}_{\text{cam}}^*)$ as the minimum argument of the cost function

$$(\mathbf{p}_{\text{kin}}^*, \mathbf{p}_{\text{cam}}^*) = \arg \min_{(\mathbf{p}_{\text{kin}}, \mathbf{p}_{\text{cam}})} \left[\frac{1}{2} \sum_{m,k} \rho(\|\mathbf{f}_{m,k}(\mathbf{p}_{\text{kin}}, \mathbf{p}_{\text{cam}})\|^2) \right]. \quad (9)$$

The loss function $\rho(\bullet)$ is used to reduce the influence of large residuals, which are mainly caused by false marker detections. We model and solve this non-linear least squares problem with the C++ library Google Ceres.¹⁰ Concerning the initial guesses for the solver iterations, we use the available CAD values for the length measurements, and we set the angle offsets, the shank-to-foot angle and the time offset to zero. For the camera intrinsics, we use a calibration file obtained with the ROS camera calibrator. In order to get initial values for the camera extrinsics and the foot-to-pattern angle, we run a less expensive pre-calibration with only these values as free parameters.

4. Experimental Results

Three datasets were recorded and evaluated with the presented calibration framework for the two front legs of the quadruped robot StarLETH.¹¹ The robot movements were chosen such that the markers would be appearing in different areas of the image. The approximate duration of each dataset is 30 seconds, and the solver takes 10–20 seconds, including the pre-calibration step. All parameters from Table 1 are included in the calibration, except for the body measure b_x , which is not observable if the camera extrinsics are co-estimated while considering the front legs only. This gives a total of 33 jointly estimated parameters. Figure 3 shows snapshots from the recording process and the visual result of the calibration. Including the camera intrinsics significantly improved the results and lowered the optimization residual. Even at the borders of the image, where higher radial distortions occur, the calibrated model is able to convincingly depict the pose of all markers.

The numerical results of the locomotion relevant kinematic parameters are summarized in Table 2. The deviations of the calibrated length measurements from their corresponding CAD values are around five millimeters (except l_s). The standard deviations are relatively low, indicating a good accuracy and a low dataset dependency of the results. The obtained joint angle offsets are around three degrees with a standard deviation of one degree in average, which is certainly worth dealing with for a better robot

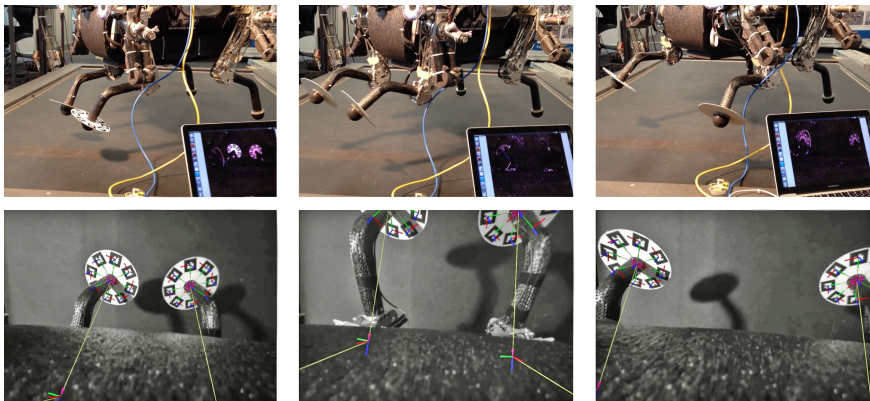


Fig. 3. *Top*: Examples of robot configurations during the dataset recording. *Bottom*: Rectified camera images overlaid with the calibrated kinematic chains of the legs (the calibrated intrinsics were employed). Modeled and observed marker poses fit well.

Table 2. Estimated kinematic parameters of StarLETH. The calibration was run for all three datasets separately, the table shows the means and the standard deviations.

Segment Lengths* (mm)					
<i>left</i>			<i>right</i>		
l_h	l_t	l_s	l_h	l_t	l_s
71.5 ± 0.9	194 ± 2	251 ± 7	62.6 ± 2.6	205 ± 1.9	269 ± 2
Joint Offsets (deg)					
<i>left</i>			<i>right</i>		
$\theta_{h,offset}$	$\theta_{t,offset}$	$\theta_{s,offset}$	$\theta_{h,offset}$	$\theta_{t,offset}$	$\theta_{s,offset}$
-3.8 ± 0.4	-2.5 ± 1.5	1.1 ± 2.5	-2.2 ± 0.3	2.2 ± 0.6	-5.6 ± 0.9
Shank-to-Foot Angle θ_f (deg)			Body Measure* (mm)		
<i>left</i>		<i>right</i>	b_y		
24.8 ± 1.4		30.3 ± 1.2	182 ± 1		

Note: * The corresponding CAD values are $l_h = 68.5$ mm, $l_t = 200$ mm, $l_s = 235$ mm and $b_y = 185$ mm. The value for the shank length l_s corresponds to an earlier shorter shape of the shank segment.

performance. Also the camera calibration parameters yielded reasonable results. The time offset between joint measurements and images is approximately -0.05 s. The importance of including t_{offset} in our calibration procedure can be seen in Fig. 4, where the pixel errors after the calibration are plotted. If not considering the time offset, a motion dependent error can be observed.

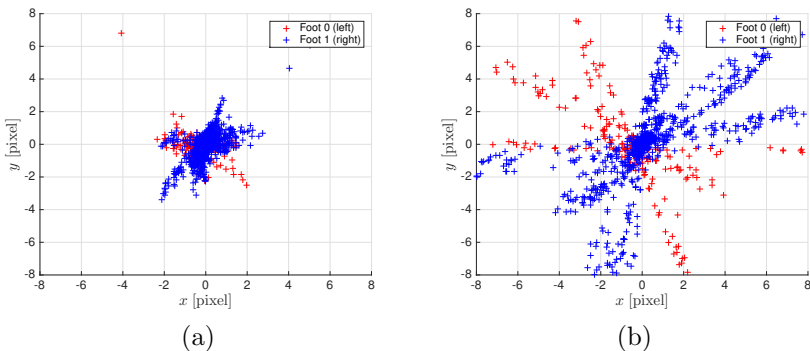


Fig. 4. The resulting pixel errors for the first dataset. (a) Full 33 DOF calibration, RMS error of 2.3. (b) Calibration without time offset (32 DOF), RMS error of 3.7.

5. Conclusion

A fast and automatic foot-eye calibration framework has been presented, which is versatile in the sense that a variable number of kinematic leg chains can be calibrated concurrently. It was successfully evaluated for the quadruped robot StarLETH with an overall runtime of only 40–50 seconds. In order to get redundant measures of the robot’s feet, we temporarily mounted a camera and AR marker patterns to the robot. Subsequently, we estimated the most likely calibration parameters by minimizing the error between the measured and the modeled marker positions. The accuracy of our estimates is in the range of five millimeters for the length measurements of the robot. Future work may include an evaluation of the lens distortion model or a systematic observability analysis.

References

1. W. Khalil, S. Besnard and P. Lemoine, Comparison study of the geometric parameters calibration methods, in *International Journal of Robotics and Automation (IJRA)* **15**, 56 (2000).
2. R. Horaud and F. Dornaika, Hand-eye calibration, in *The International Journal of Robotics Research (IJRR)* **14**, 195 (1995).
3. K. H. Strobl and G. Hirzinger, Optimal hand-eye calibration, in *International Conference on Intelligent Robots and Systems (IROS)*, 2006.
4. K. M. Nickels and K. Baker, Hand-eye calibration for robonaut, in *NASA Summer Faculty Fellowship Program* (2003).
5. T. Kastner, T. Röfer and T. Laue, Automatic robot calibration for the nao, in *RoboCup 2014: Robot World Cup XVIII*, (Springer, 2015) pp. 233–244.
6. V. Pradeep, K. Konolige and E. Berger, Calibrating a multi-arm multi-sensor robot: A bundle adjustment approach, in *International Symposium of Experimental Robotics (ISER)*, 2014.
7. M. Bloesch, M. Hutter, C. Gehring, M. Hoepflinger and R. Siegwart, Kinematic batch calibration for legged robots, in *International Conference on Robotics and Automation (ICRA)*, 2013.
8. O. Birbach, U. Frese and B. Bäuml, Rapid calibration of a multi-sensorial humanoids upper body: An automatic and self-contained approach, in *The International Journal of Robotics Research (IJRR)* **34**, 420 (2015).
9. H. Kato and M. Billinghurst, Marker tracking and hmd calibration for a video-based augmented reality conferencing system, in *International Workshop on Augmented Reality (IWAR)*, 1999.
10. S. Agarwal, K. Mierle and Others, Ceres solver <http://ceres-solver.org>.
11. M. Hutter, C. Gehring, M. Bloesch, M. A. Hoepflinger, C. D. Remy and R. Siegwart, Starleth: A compliant quadrupedal robot for fast, efficient, and versatile locomotion, in *International Conference on Climbing and Walking Robots (CLAWAR)*, 2012.