

# Automatic and Self-Contained Calibration of a Multi-Sensorial Humanoid’s Upper Body

Oliver Birbach

Berthold Bäuml

Udo Frese

**Abstract**—Complex manipulation tasks require an accurate interplay of actuation and sensing. This accuracy can only be achieved by calibrating the relevant components beforehand. Typically calibration procedures are time-consuming and often include subsequent calibration steps, involve multiple people and require external tools. In this paper we alleviate these issues by *auto-calibrating* the different sensors of DLR’s humanoid Rollin’ Justin in a single, completely automatic and self-contained procedure, *i.e.* without calibration plate. By observing a single point feature on each wrist while moving the robot’s head, the stereo cameras’ intrinsic and extrinsic parameters are calibrated together with the arm joint elasticities and joint angle offsets. Additionally, we use the head motion to calibrate an Inertial Measurement Unit (IMU) extrinsically. Parameters are obtained by formulating the calibration problem as a batch-optimization problem that estimates all parameters jointly. A rough initial guess, as is, *e.g.*, available when re-calibrating, is needed for the estimation and to facilitate marker detection. The procedure is validated on real hardware and reduces the effort considerably allowing rapid (5 min movement time), automatic, and accurate calibration by simply “pushing a button”.

## I. INTRODUCTION

Calibration is an inevitable process for determining relevant parameters of complex robotic systems. However, it is considered a necessary evil since it is only a precondition for the tasks the robot was actually designed for. Also, as the robot undergoes maintenance, after collision or simply long usage, the previous calibration parameters often become invalid and the robot has to be re-calibrated. Depending on the routine, performing calibration can be costly in terms of time and human resources. External devices or tools (*e.g.* high-precision measurement systems, checkerboard patterns) might be required. Most importantly, the robot cannot be used for other experiments during the calibration process.

In this paper we alleviate these problems and present an auto-calibration procedure used to determine the necessary parameters of a humanoid robot with a complex kinematic chain and equipped with multiple sensors to accomplish demanding manipulation tasks [1], [2]. Manipulation tasks require accurate interplay between actuation and sensing and therefore a careful calibration of the involved components. In our case these components are a pair of cameras, an Inertial Measurement Unit (IMU) to measure (possibly kinematically unobservable) head movements, and the robot’s actuators to interact with the environment.

O. Birbach and U. Frese are with German Center for Artificial Intelligence (DFKI), 28359 Bremen, Germany [oliver.birbach@dfki.de](mailto:oliver.birbach@dfki.de)

B. Bäuml is with DLR Institute of Robotics and Mechatronics, Münchenerstr. 20, 82234 Wessling, Germany [berthold.baeuml@dlr.de](mailto:berthold.baeuml@dlr.de)

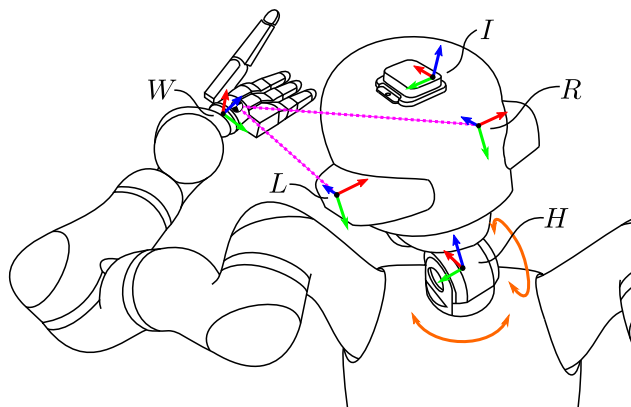


Fig. 1. Sketch of the calibration process showing DLR’s Rolling Justin from the back including the relevant frames. The robot observes (magenta dotted lines) a point-feature attached to its wrist  $W$  (cf. Fig. 3 for a close-up of the marker) with its two cameras while moving the head. Additionally, all measurements from the IMU, mounted in the robot’s head, and the joint angle and torque sensors are recorded. This data together with the corresponding measurement models is fed into a least-squares estimator. The results are the calibrated poses of the left camera ( $L$ ) and the IMU ( $I$ ) relative to the head frame  $H$ , the stereo pair relationship ( $L$  relative to  $R$ ), the cameras’ intrinsic parameters, the joint angle offsets and the arm joint elasticities.

Existing approaches for calibrating these or parts of these parameters are often elaborate procedures. It is common practice to break a complex calibration task for multiple sensors into multiple simpler tasks such as in a pair-wise calibration approach. Due to these individual procedures, inconsistencies in the obtained calibration results might occur (*e.g.* different scale factors). Furthermore, external devices are used, *e.g.* a checkerboard pattern for camera calibration.

Our idea is to record all the data without an external calibration device and to calibrate actuation and sensor parameters jointly in a mutually supportive way as explained in Fig. 1. The long term vision is to do the calibration concurrently while the robot is operating, just as living animals continuously adapt to their changing body. The result presented in this paper is a “push-button” calibration based on feeding all this data into a least-squares estimator. The method depends on a given rough initial guess, both to have a search region for the marker detection and as a starting point for least-squares estimation. In the frequent case of re-calibration, such an initial guess is always available and the procedure is fully automatic. Computing an initial guess directly from the recorded data for a first, “factory” calibration is ongoing research. But, at most, the user might be

required to click on the marker in the first image and then the marker can be automatically tracked during the subsequent head motion. The rough calibration obtained from this should be enough to bootstrap the method.

We believe it is an important contribution that the sensor-data is evaluated by textbook-least-squares without any *ad hoc* processing. Even though it has been known for a long time, the robotics community has not fully utilized the capability of probabilistic estimation techniques to compute calibration parameters from basic measurements models and data. In contrast, calibration is often viewed procedurally with a camera calibration, a hand-eye calibration and an inertial calibration. Our paper wants to push the point that a joint calibration is conceptually more elegant, easier to execute, easier to implement, and makes the most out of the data. This perspective is supported by using our previously published MTKM [3] library for rapid prototyping of nonlinear least-squares problems.

The rest of the paper is structured as follows. We discuss related work and our prior calibration approach in the next section. The proposed method is presented in Sec. III and experimental results are given in Sec. IV.

## II. RELATED WORK

Early robotic calibration approaches, such as those summarized in [4], mainly focused on correcting sensor readings and improving the accuracy of kinematic models of robotic manipulators. For example, one calibration approach uses an external measurement device, namely a theodolite, which observes a mirror mounted on the robot’s tool center point.

The emerging use of computer vision in robotics gave rise to hand-eye calibration ([5], [6]). Here, a camera is rigidly mounted at the end-effector of a manipulator with the goal of identifying the pose of the camera frame in the end-effector’s frame. In a similar way, the problem of calibrating a pair of cameras mounted on a platform actuated by a pan-tilt (and sometimes vergence) unit is solved, also known as head-eye or neck-eye calibration. The determination of the cameras extrinsic parameters relative to the actuator has been studied in ([7], [8], [9]) with closed-form solutions as well as non-linear refinement. Intrinsic calibration has been studied [10] and also the combination [11] of extrinsic and intrinsic calibration. An in-depth understanding of the problem is given in [12], where two view transformations of image pairs obtained by controlled rotation about the actuation axes are used to determine the alignment of cameras and a pan-tilt unit in static scenes.

Determining the geometric relation between cameras and an Inertial Measurement Unit (IMU) has been extensively studied. A two step approach is proposed in [13], using a vertically aligned checkerboard serving as a vertical reference and a turntable to estimate the offset between camera and IMU. In a more dynamic approach [14] rotation between sensors is estimated by minimizing corresponding rotation differences. A popular approach to determine this relation online is using an Extended Kalman filter [15] observing a checkerboard pattern or an Unscented Kalman Filter [16]

observing either artificial or scene features. Furthermore, in [17], an observability analysis of the problem is given, concluding that only two rotational degrees of freedom (DOF) need to be excited and no translational motion is required to obtain calibration. Indeed, this result allows us to calibrate the IMU on the pan-tilt unit. Most recently, calibration approaches using a passive complementary filter on  $SO(3)$  [18] or explicitly modeling the trajectory of the sensors to obtain the calibration through batch optimization [19] have emerged.

Over time, approaches for calibrating various aspects of humanoid robots have appeared, *e.g.* for the PR2. In [20], sensor poses, joint angle offsets and joint gear reductions are calibrated in an unified approach using a checkerboard pattern. A spatial and temporal calibration approach between camera, pan-tilt unit, odometry and an IMU is formulated as a registration problem [21]. By filtering for structure, motion and relative positions of sensors, a camera and IMU can be calibrated at the same time [22]. Calibration approaches for other robots include that of the head of humanoid CB-i using a planar checkerboard pattern [23] and iCub’s stereo head [24], specifically calibrating the neck’s pan-tilt-swing and the eyes’ pan-tilt angles using an IMU and cameras.

In [25], a two staged approach for calibrating a two-armed robot with stereo vision is presented. Camera intrinsic parameters and their frames to the head link are calibrated using a known calibration target. In an ensuing step, a marker defining a target frame is used to calibrate the unknown transformation between the arms and the neck assuming perfect forward kinematics of the arms.

Most closely related to our approach is the hand-eye calibration for the Robonaut [26]. The arm and the neck are driven according to a set of prerecorded configurations with a spherical target mounted at the arm. Assuming *a priori* camera calibration, extrinsic parameters relative to the neck are determined as well as either optimizing the joint-angle offset or the involved Denavit-Hartenberg (DH) parameters.

Our previous calibration approach is described in [3]. In this static procedure, measurements are acquired in two steps. First, a leveled checkerboard pattern with known scale is observed by the robot at different (manually positioned) poses, while the inertial sensor measures gravity in its own coordinate system. This defines the cameras’ intrinsic parameters and the relative pose of both. It also defines the orientation of the IMU relative to the cameras but not its translation. In the second step, both cameras observe the checker point feature attached to the robot’s hands in different configurations of arm and head (again, manually positioned), relating the previously calibrated frames to the head-frame. This procedure was quite time-consuming due to manual positioning. Furthermore, as it was a static approach, *i.e.* data was only taken while stationary, the IMU position could not be calibrated because it only influences measurements in motion.

The calibration procedure proposed here differs to all discussed work in that we calibrate all parameters *jointly* and require *no* external calibration target. We model the

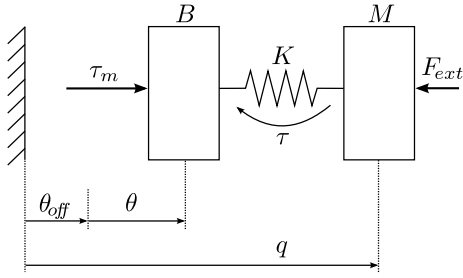


Fig. 2. Mechanical model of a single arm joint. Any torque  $\tau_m$  commanded by the position controller acts on the rotor inertia  $B$ . The elasticity of the transmission between rotor and the following link (with mass  $M$ ) which is subject to a force  $F_{ext}$  (induced by gravity depending on the overall configuration) is modeled as a spring with stiffness  $K$ . The actual link joint angle  $q$  is then composed of a joint angle offset  $\theta_{off}$ , the measured joint angle  $\theta$  and the rotational displacement defined by  $K$  and the torque  $\tau$  as measured by joint torque sensor. Because there is no link-side position sensor,  $q$  cannot be measured directly. The figure depicts a translational analog of the rotational joint.

calibration problem in a consistent and textbook style way, using the robot’s ability to precisely measure the head’s joint angle as the main source of precision in our approach.

### III. PROPOSED CALIBRATION APPROACH

#### A. Problem Formulation

The problem is cross-calibrating a robot’s stereo cameras, IMU, and kinematic chain (Fig. 1), in particular to determine

- both cameras’ intrinsic parameters, including focal length  $f_{L/R}$ , principal point  $C_{L/R}$  and radial distortion ( $\kappa_{L/R}$ ),
- poses of the left camera ( $T_H^L$ ) and the IMU ( $T_H^I$ ) relative to the head frame  $H$ , located after the neck’s pan-tilt-unit, and the left camera relative to the right camera (stereo,  $T_R^L$ ) together with the
- correction parameters (angle offset  $\theta_{off,i}$ , elasticity  $K_i^{-1}$ ) for the joints  $i$  in the kinematic chain.

#### B. Visual Measurement Model

Visual measurements are obtained by running a local checkerboard-corner detector on the images using a search region computed from the initial guess. It returns the image position of the point marker on the robot’s left and right wrist.

The corresponding measurement model needed for least-squares estimation predicts the image position  $u_L$  in the left camera given a hypothetical calibration. It takes the to be estimated marker position  $p_W$  in wrist-coordinates  $W$  and maps it first to head-coordinates using the forward kinematics  $T_H^W(q)$ , then to left camera-coordinates with the to be estimated parameter  $T_H^L$ , and finally, to image-coordinates with the usual pinhole-model  $P(\dots)$  and radial distortion  $d_{\kappa_L}(\dots)$ , where the center point  $C_L$ , the focal length  $f_L$  and the distortion  $\kappa_L$  are also estimated.

$$p_L = T_H^{L-1} T_H^W(q) p_W \quad (1)$$

$$u_L = C_L + f_L d_{\kappa_L}(P(p_L)), \text{ with} \quad (2)$$

$$P(p) = \left( \frac{p_1}{p_3}, \frac{p_2}{p_3} \right), \quad d_{\kappa_L}(u) = \frac{u}{1 + \kappa_L |u|^2} \quad (3)$$

The right camera (subscript  $R$ ) is treated similarly. The only difference is that the point in head coordinates is further transformed into the right camera coordinate system (*i.e.*  $T_R^L T_H^{L-1}$  instead of  $T_H^L$  in Eq. 1).

It turned out that the measured joint angles  $\theta$  are affected by a constant offset  $\theta_{i,off}$  as well as by the joint elasticity  $K_i$  (Fig. 2). We correct these effects using

$$q_{i,t} = \theta_{i,off} + \theta_{i,t} + K_i^{-1} \tau_{i,t} \quad (4)$$

where  $\theta_{i,t}$  and the torque  $\tau_{i,t}$  are measured and  $\theta_{i,off}$  and  $K_i^{-1}$  are estimated for each joint  $i$ . Since there are no torque sensors in the head’s pan-tilt unit, we set  $K_i^{-1} = 0$  for these. Additionally,  $\theta_{i,off} = 0$  for the last arm joint because this offset can be included in  $p_W$ .

#### C. Inertial Measurement Model

The inertial measurement model predicts acceleration  $a_t$  (with gravity) and angular velocity  $\omega_t$  measured by the IMU at time  $t$  based on a hypothetical calibration. First, the IMU pose  $T_{B_t}^I$  in the base frame is computed using forward kinematics, then  $\omega_t$  is obtained from the rotational difference between successive poses. Next, the geometrical acceleration  $a_t^*$  in the base frame  $B$  is computed from the translation vector of three successive poses  $(T_{B_t}^I)_{\bullet 4}$ . Finally, gravity  $g_B$  is added and the result converted into IMU coordinates to obtain the predicted accelerometer readings  $a_t$ .

$$T_{B_t}^I = T_B^H(q_t) T_H^I \quad (5)$$

$$\omega_t = \frac{\text{arccrot} \left( (T_{B_{t-\delta t}}^I)^{-1} T_{B_t}^I \right)}{\delta T} + b_\omega \quad (6)$$

$$a_t^* = \frac{(T_{B_{t-\delta t}}^I)_{\bullet 4} - 2(T_{B_t}^I)_{\bullet 4} + (T_{B_{t+\delta t}}^I)_{\bullet 4}}{\delta t^2} \quad (7)$$

$$a_t = (T_{B_t}^I)^{-1} (a_t^* + g_B) \quad (8)$$

Here,  $\text{arccrot}$  (“arcus rotation”) is the inverse of the Rodrigues’ rotation formula and maps a rotation matrix to the corresponding rotation axis, with the length of the result giving the angle of rotation. The IMU frame  $T_H^I$ , the gyro bias  $b_\omega$ , and gravity  $g_B$  in the base frame are estimated. The last is required as the robot has a slightly uncertain orientation due to wheel and suspension elasticities. It also incorporates most of the accelerometer bias. During the IMU measurements only the head moves, so there is just a single  $g_B$  to be estimated.

#### D. Structural Vibrations

In experiments, it turned out that the head movements excite vibrations of various frequencies both in the robot’s joints and structure.

The low frequency vibrations originate mainly from joint elasticities and are measured by the torque/position sensors and, hence, are correctly incorporated by (4). Thus it was important to include the torso joints in (5) even though the torso was not actively moved.

The high frequency vibrations have two sources. First, a velocity dependent component due to the motor ripple in the head joints because of the lack of torque sensors in the

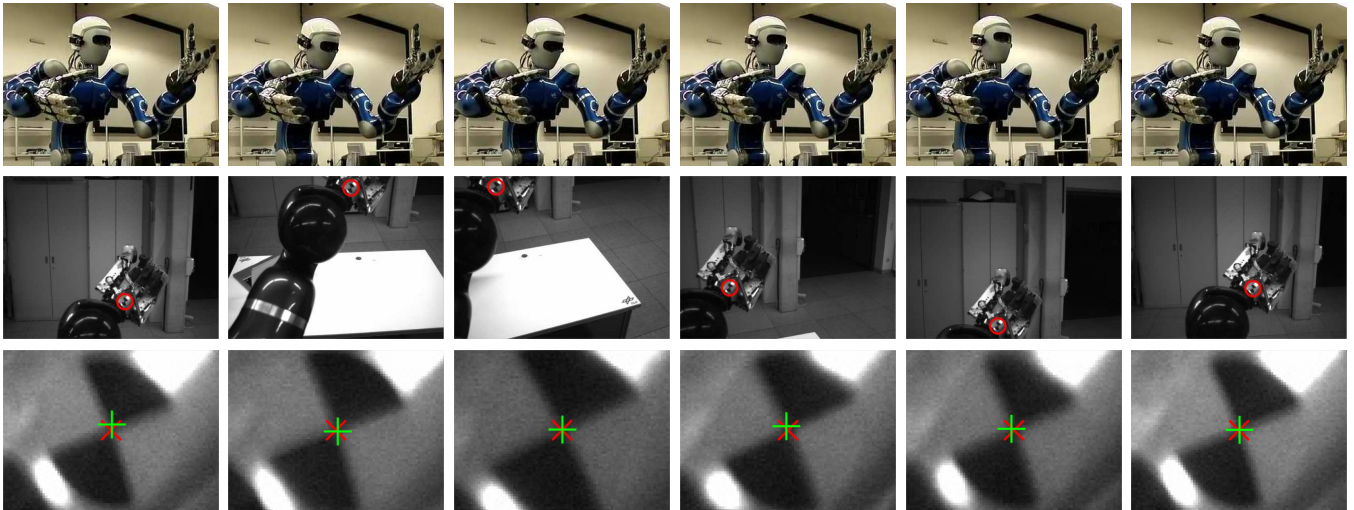


Fig. 3. Snapshots from calibration experiments observing a single arm posture. **(Top)** External view of calibration procedure. **(Middle)** View from the left camera. The attached marker (point feature) is highlighted with a red circle. **(Bottom)** Close-up of corresponding wrist-mounted feature. The detected center is depicted as a red cross ( $\times$ ) and the predicted center as a green cross ( $+$ ). Please see also the supplemental video for a calibration run.

pan-tilt unit. Second, vibrations in the head structure itself, including a component perpendicular to the pan and tilt axes and induced by high acceleration movements.

Fortunately, those vibrations had a higher frequency than the useful signal allowing to low-pass filter all input data with a 5<sup>th</sup> order Butterworth filter with 5 Hz cutoff-frequency.

#### E. Joint Least-Squares Estimation

The calibration parameters, as introduced in Sec. III-A, are estimated by feeding all measurements ( $u_{t,L}$ ,  $u_{t,R}$ ,  $\omega_t$ ,  $a_t$ ) together with the measurement functions defined by (2), (6), and (8) into our general purpose least-squares library MTKM [3]. The library is designed for rapid prototyping and takes care of all the bookkeeping, numerical calculation of Jacobians, and the actual optimization using the Levenberg-Marquardt algorithm. In particular, it takes care of the parametrization issues and handles 3-D rotations without singularities. The user only has to define the variables to be estimated, supply measurements, measurement functions and uncertainties and define which measurement involves which variables. From the latter information the library automatically derives the problem's sparsity pattern and exploits it for Jacobian computation and linear equation solving.

The uncertainties of all measurements are estimated from their respective residuals with one standard deviation parameter for vision, accelerometer, and gyroscope each. The joint angles  $\theta_{i,t}$  and torques  $\tau_{i,t}$  are used in the measurement models (via (4) and the kinematics) as given parameters, not as measurements with uncertainty.

### IV. EXPERIMENTAL RESULTS

#### A. Prerequisites

For our experiments we used a set of 27 taught robot configurations (13 left, 14 right), which were chosen heuristically with different kinematic configurations and distances to the cameras in mind. Head movements were then generated for each of these configurations requiring the feature to be

at least visible in one camera and covering most of the pan-tilt unit's joint space. The maximum angular velocity for these trajectories was limited ( $24^\circ/s$ ) to reduce vibrations and motion blur. For robust IMU-head calibration  $60^\circ/s$  was commanded. Stereo images ( $1616 \times 1220$  px, 1 ms exposure) were recorded at 12.5 Hz, IMU measurements at 512 Hz and kinematic data at 1 kHz. The total movement time was 302 s with 4737 recorded feature points and 13045 IMU measurements (see Fig. 3).

The time delay between actual measurement and arrival at the recording host is not deterministic, *e.g.* due to transmission delays. We therefore de-jitter all host timestamps using a linear Kalman filter that assumes constant sensor rates. The constant latencies between the physical event and the arrival at the host have been determined in a prior calibration routine [2].

#### B. Calibration Result

Table I presents the estimates  $\mu$ , their  $\sigma$ -bound and the residual root mean square (rms). All sensors are estimated closely to their physical position which is also suggested by the low  $\sigma$ . Although not mentioned in the table, kinematic parameters such as joint angle offset and joint stiffness also have plausible estimates. The residuals are low, in particular for the cameras, suggesting that the models fit well. Their distribution is presented in Fig. 4a) showing a consistently sampled Gaussian distribution. The cameras residual rms corresponds to a location error of roughly 0.7 mm at 1 m object distance.

Another indicator of consistency is depicted in Fig. 4b) showing the residual over time and for different arm poses. For the clear majority of the measurements, the residuals are inside the  $3\sigma$  bound.

An overlay of measured and predicted gyroscope and accelerometer measurements is given in Fig. 4c) and d), respectively. Although the torso kinematics are included in the model, the IMU motion can not be fully explained

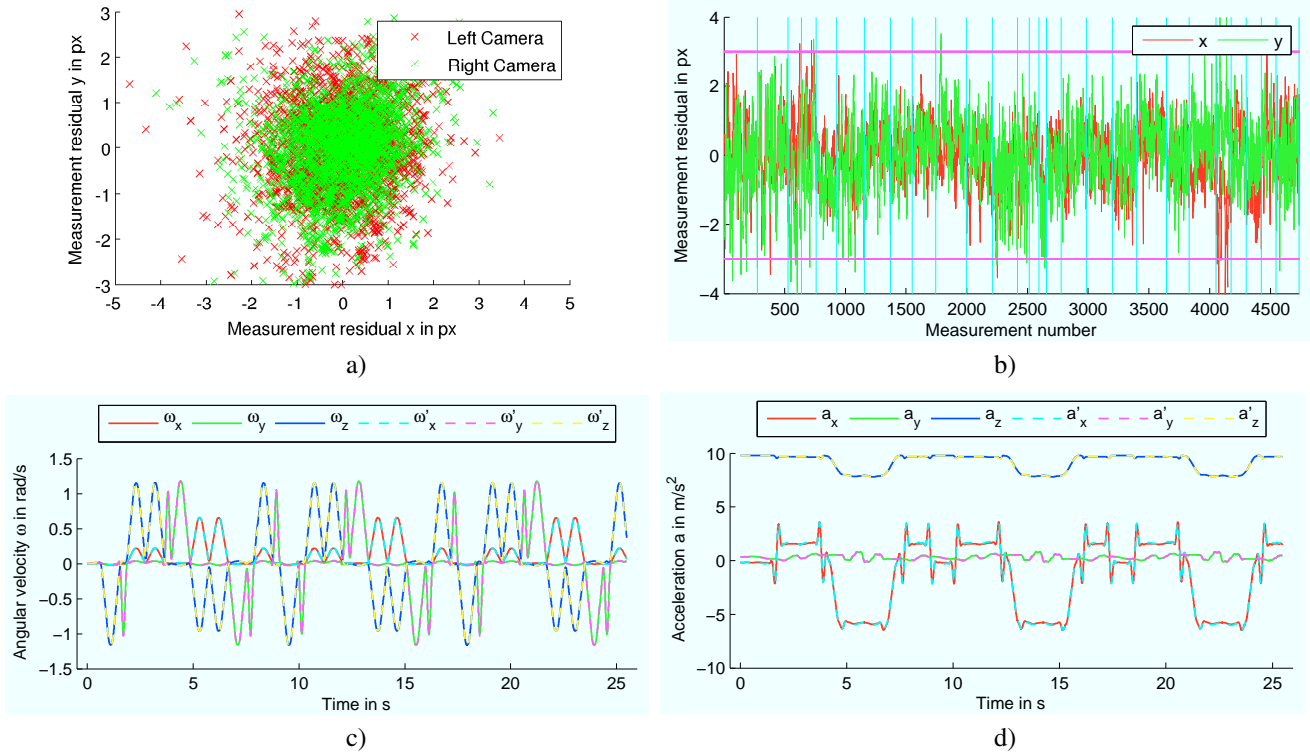


Fig. 4. **a)** Measurement residual of left and right camera measurements. **b)** Measurement residual of both cameras over time. Vertical lines indicate a change in the arm pose. The  $3\sigma$  bound is given as horizontal lines. Angular velocity **c)** and linear acceleration **d)** as measured by the IMU ( $\omega, a$ ) and predicted by the model ( $\omega', a'$ ) given the head joint angles over time.

TABLE I  
CALIBRATION RESULTS (EXCERPT)

	Intrinsic left			Intrinsic right		
	$f_L$ (px)	$C_L$ (px)	$\kappa_L$	$f_R$ (px)	$C_R$ (px)	$\kappa_R$
$\mu$	1869.4	839.7, 619.5	0.10	1860.8	817.3, 619.3	0.10
$\sigma$	0.53	0.54, 0.66	$7.2 \cdot 10^{-4}$	0.52	0.58, 0.66	$7.6 \cdot 10^{-4}$

	Transformation $T_R^L$					
	Translation (m)			Rotation (axis angle)		
	$x$	$y$	$z$	$x$	$y$	$z$
$\mu$	-0.201	0.001	-0.002	0.012	-0.006	-0.009
$\sigma$	0.0001	0.0001	0.0003	0.0005	0.0004	0.0001

	Transformation $T_H^L$					
	Translation (m)			Rotation (axis angle)		
	$x$	$y$	$z$	$x$	$y$	$z$
$\mu$	0.066	0.100	0.130	-1.329	1.316	-1.127
$\sigma$	0.0002	0.0001	0.0001	0.0001	0.0003	0.0002

	Transformation $T_H^R$					
	Translation (m)			Rotation (axis angle)		
	$x$	$y$	$z$	$x$	$y$	$z$
$\mu$	-0.002	0.006	0.236	0.026	-0.014	0.006
$\sigma$	0.0007	0.0012	0.0009	0.0006	0.0005	0.0007

	Residual rms					
	Left camera (px)			Right camera (px)		
	$x$	$y$		$x$	$y$	
	0.90	1.01		0.88	0.95	
	Accelerometer ( $m/s^2$ )			Gyroscope (rad/s)		
	$a_x$	$a_y$	$a_z$	$\omega_x$	$\omega_y$	$\omega_z$
	0.079	0.019	0.019	0.004	0.009	0.006

by the kinematic model. This is indicated in the increased accelerometer residual in  $x$ -direction. It is suspected that there is still unmodeled shaking (back and forth) due to elasticities which will be investigated in the future.

To highlight the importance of integrating joint-angle offsets and joint stiffness into the model, we performed

calibration experiments where these parameters were not part of the estimation. Additionally, an estimation leaving the forward kinematic out and estimating the feature points as 3-D positions relative to the torso was conducted. Measurement residuals of both cameras are shown in Fig. 5 for the different cases. This underlines that consistency and quality is considerably improved by integrating these parameters.

The quality of our calibration approach is also confirmed by the performance that could be reached for a demanding task which depends on the precision of the calibrated parameters: using the automated calibration as the robot's re-calibration routine before ball catching experiments ([1], [2]) the same catch-rate as for the manual calibration procedure has been reached.

## V. CONCLUSION

In this paper we proposed an automated and self-contained calibration approach for multiple sensors of a humanoid's upper body. Using no external target and exploiting the robot's ability to precisely measure the head's motion, we are able to calibrate stereo cameras, an IMU, their relation to the kinematic chain and possible deviations in the kinematic model. We validated our approach in a calibration experiment on DLR's Rollin Justin achieving sound results.

In future work we want to integrate temporal calibration of the sensors to complete the set of desired calibration parameters. Furthermore, a global feature detection approach and computing an initial guess from this data is desired. Also, we will investigate what arm configurations and head trajectories are required for optimal calibration performance.

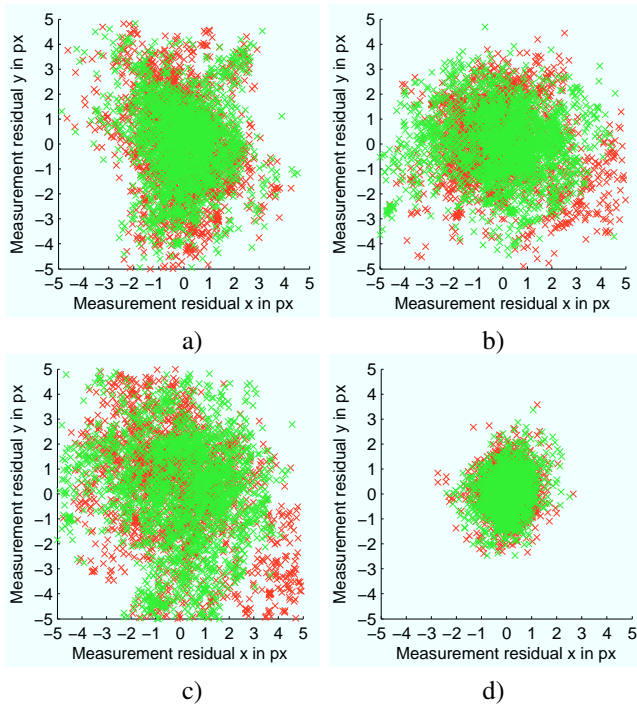


Fig. 5. Vision measurement residuals of the left (red crosses) and right (green crosses) camera from estimations not considering  $K$  (a),  $xy$ -rms 1.21, 1.80 px). It is visible how the wrist is displaced in the  $y$ -axis due gravity dragging the arm down. Not considering  $\theta_{\text{off}}$  leads to inaccurate wrist positions in different arm poses visible as increased noise in both axes (b),  $xy$ -rms 1.71, 1.45 px). Neglecting both, the kinematic error adds up considerably (c),  $xy$ -rms 1.93, 2.53 px). Additionally, not using forward kinematics but only estimating the 3-D position of the feature in torso coordinates leads to (d),  $xy$ -rms 0.55, 0.73 px). Compared to Fig. 4a), there is still some error. But compared to not considering the deviations in the kinematic model our approach considerably increases consistency.

Although the used optimization framework MTKM turned out to be useful for rapid-prototyping this complex calibration problem, computational performance needs to be improved. For this, we plan to port our current implementation to a C++ based optimization framework, preferably SLOM [27]. This will reduce optimization time considerably.

As the next step, we will extend the calibration to the whole kinematic chain including torso and wheels using a marker on the platform and the wheels, which are mounted on extensible legs. This will allow for a full, automatic and self-contained calibration of a whole humanoid robot.

## VI. ACKNOWLEDGEMENTS

This work was supported under DFG grant FR2620/1-1. We thank DLR's *Rollin' Justin* team for their inevitable technical support.

## REFERENCES

- [1] B. Bäuml, F. Schmidt, T. Wimböck, O. Birbach, A. Dietrich, M. Fuchs, W. Friedl, U. Frese, C. Borst, M. Grebenstein, O. Eiberger, and G. Hirzinger, "Catching flying balls and preparing coffee: Mobile humanoid Rollin Justin performs dynamic and sensitive tasks," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2011, pp. 3443–3444.
- [2] B. Bäuml, O. Birbach, T. Wimböck, U. Frese, A. Dietrich, and G. Hirzinger, "Catching flying balls with a mobile humanoid: System overview and design considerations," in *Proc. of the IEEE-RAS Int. Conf. on Humanoid Robotics*, 2011, pp. 513–520.
- [3] R. Wagner, O. Birbach, and U. Frese, "Rapid development of manifold-based graph optimization systems for multi-sensor calibration and SLAM," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2011, pp. 3305–3312.
- [4] Z. Roth, B. Mooring, and B. Ravani, "An overview of robot calibration," *IEEE Journal on Robotics and Automation*, vol. 3, no. 5, pp. 377–385, 1987.
- [5] Y. C. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form  $AX = XB$ ," *Trans. on Robotics and Automation*, vol. 5, no. 1, pp. 16–29, 1989.
- [6] K. H. Strobl and G. Hirzinger, "Optimal hand-eye calibration," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2006, pp. 4647–4653.
- [7] M. Li and D. Betsis, "Head-eye calibration," in *Proc. of the Int. Conf. on Computer Vision*, 1995, pp. 40–45.
- [8] S. D. Ma, "A self-calibration technique for active vision systems," *Trans. on Robotics and Automation*, vol. 12, no. 1, pp. 114–120, 1996.
- [9] M. Li, "Kinematic calibration of an active head-eye system," *IEEE Trans. on Robotics*, vol. 14, no. 1, pp. 153–158, 1998.
- [10] C. Yang and Z. Hu, "An intrinsic parameters self-calibration technique for active vision system," in *Proc. of the Int. Conf. on Pattern Recognition*, 1998, pp. 67–69.
- [11] G. N. DeSouza, A. H. Jones, and A. C. Kak, "A world-independent approach for the calibration of mobile robotics active stereo heads," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2002, pp. 3336–3341.
- [12] J. Knight and I. Reid, "Automated alignment of robotic pan-tilt camera units using vision," *International Journal of Computer Vision*, vol. 68, no. 3, pp. 219–237, 2006.
- [13] J. Lobo and J. Dias, "Relative pose calibration between visual and inertial sensors," *Int. Journal of Robotics Research*, vol. 26, no. 6, pp. 561–575, 2007.
- [14] P. Lang and A. Pinz, "Calibration of hybrid vision / inertial tracking systems," in *Proc. of the InerVis: Workshop on Integration of Vision and Inertial Sensors*, 2005.
- [15] F. M. Mirzaei and S. I. Roumeliotis, "A Kalman filter-based algorithm for IMU-camera calibration," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2007, pp. 2427–2434.
- [16] J. Kelly and G. S. Sukhatme, "Fast relative pose calibration for visual and inertial sensors," in *Proc. of the Int. Symposium on Experimental Robotics*, 2008.
- [17] F. M. Mirzaei and S. I. Roumeliotis, "A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation," *IEEE Trans. on Robotics*, vol. 24, no. 5, pp. 1143–1156, 2008.
- [18] G. G. Scandaroli, P. Morin, and G. Silveira, "A nonlinear observer approach for concurrent estimation of pose, IMU bias and camera-to-IMU rotation," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2011, pp. 3335–3341.
- [19] M. Fleps, E. Mair, O. Ruepp, M. Suppa, and D. Burschka, "Optimization based IMU camera calibration," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2011, pp. 3297–3304.
- [20] V. Pradeep, K. Konolige, and E. Berger, "Calibrating a multi-arm multi-sensor robot: A bundle adjustment approach," in *Proc. of the Int. Symposium on Experimental Robotics*, 2010.
- [21] J. Kelly and G. S. Sukhatme, "A general framework for temporal calibration of multiple proprioceptive and exteroceptive sensors," in *Proc. of Int. Symposium on Experimental Robotics*, 2010.
- [22] J. Kelly and G. S. Sukhatme, "Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration," *Int. Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.
- [23] A. Ude and E. Oztop, "Active 3-d vision on a humanoid head," in *Int. Conf. on Advanced Robotics*, 2009.
- [24] J. Santos, A. Bernardino, and J. Santos-Victor, "Sensor-based self-calibration of the iCub's head," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2010, pp. 5666–5672.
- [25] C. Garcia, "Fully vision-based calibration of a hand-eye robot," *Autonomous Robots*, vol. 6, no. 2, pp. 223–238, 1999.
- [26] K. M. Nickels, "Hand-eye calibration for Robonaut," NASA Summer Faculty Fellowship Program Final Report, Johnson Space Center, Tech. Rep., 2003.
- [27] C. Hertzberg, R. Wagner, U. Frese, and L. Schröder, "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds," *Information Fusion*, (to appear).