Active Robot Learning for Efficient Body-Schema Online Adaptation

Gonçalo Cunha
goncalocarvalhocunha@gmail.com
Alexandre Bernardino
alexandre.bernardino@tecnico.ulisboa.pt
Pedro Vicente
pvicente@isr.tecnico.ulisboa.pt
Ricardo Ribeiro
ribeiro@isr.tecnico.ulisboa.pt
Plínio Moreno
plinio.lopez@tecnico.ulisboa.pt

Instituto de Sistemas e Robótica
Universidade de Lisboa - Instituto Superior Técnico
Lisboa, Portugal

Abstract

This work proposes an active learning approach for estimating the Denavit-Hartenberg parameters of 7 joints of the iCub arm in a simulation environment, using observations of the end-effector's pose and knowing the values from proprioceptive sensors. Cost-sensitive active learning, aims to reduce the number of measurements taken and also reduce the total movement performed by the robot while calibrating, thus reducing energy consumption, along with mechanical fatigue and wear. The estimation of the arm's parameters is done using the Extended Kalman Filter and the active exploration is guided by the A-Optimality criterion. The results show cost-sensitive active learning can perform similarly to the straightforward active learning approach, while reducing significantly the necessary movement.

1 Introduction

Active learning is a sub-field of machine learning which aims to reduce the amount of training data required to build a model, with a certain precision. This is done by having the learning algorithm decide which data it wants to label/sample next. A general introduction for this area of research can be found in [8].

1.1 Related Work

Recent works have succeeded in employing different strategies for body schema adaptation, such as [9], [10] and [7]. All these works show different successful ways of accounting for the robot's body errors, but using active learning to this effect would promote faster adaptation.

Active learning methods have better empirical results, when compared to random sampling. Some of these works are described in [3], [4], [7], and [2]. These works have shown the advantages of using active learning but assume all samples have equal acquisition cost. In [5] a criterion is proposed considering uncertainty and travel cost for the designated task, minimising the accumulated path length needed for accurate estimation, at the cost of an increase of the number of samples.

1.2 Contributions

This work aims to estimate the Denavit-Hartenberg (DH) parameters of 7 rotational joints of the iCub arm in a simulation environment, by acquiring observations from the pose of the end-effector, using active learning to select the best joint configurations for movement and sampling efficiency.

By portraying an arbitrary serial robotic arm as in Figure 1, a calibration routine is proposed to make use of active learning to select the best joint configurations to sample the end-effector pose, in order to estimate the DH parameters with the best possible precision, using the Extended Kalman Filter (EKF).

Similarly to [5], we argue that using active learning to reduce the number of samples taken may not be the best approach, since some of the best samples may require unnecessary long movements, increasing execution time and energy spent. The cost-sensitive active learning approach provides a tunable trade-off between minimising the number of iterations required and minimising the required movement. The proposed calibration routine is composed of the key steps shown in Figure 2.

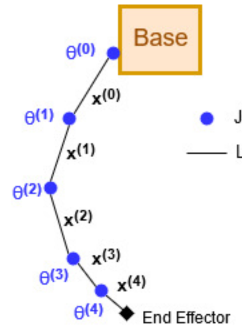

Figure 1: Illustration of a robot's kinematic chain. $\theta^{(i)}$ represents the angle value for joint i and $x^{(i)}$ represents the Denavit-Hartenberg parameters describing the transformation between frames i and $i+1$.

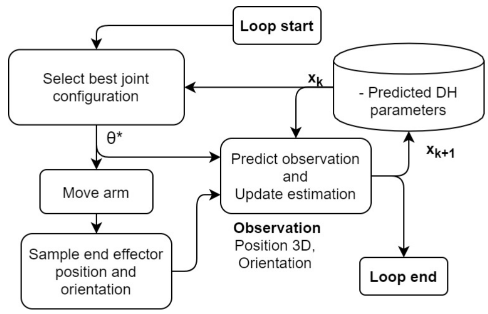

Figure 2: Key steps in the structure of the required program.

1.3 Extended Kalman Filter

The EKF, explained in detail in [6], allows recursive parameter estimation of systems represented by a nonlinear model, which is the case for the relation between the DH parameters, x, and the end-effector pose, z, given by the function

$$z = h(x, \theta), \tag{1}$$

where θ represents the known joint angles. Since the DH parameters are constant in time, the EKF can be summarized in the following 3 equations. The predicted co-variance, P, of the DH parameters, x, is given by

$$P_{k+1|k} = P_{k|k} + Q_k, \tag{2}$$

where Q_k is the co-variance matrix of the Gaussian noise associated with slow changes in the DH parameters, e.g. due to temperature. The update of the prediction, $\hat{x}$, after obtaining a measurement, z_k, is given by

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}[z_k - h(\hat{x}_{k+1|k}, \theta_k)] \tag{3}$$

and the update of the co-variance, P, is given by

$$P_{k+1|k+1} = P_{k+1|k} - K_{k+1}\left[H_k P_{k+1|k} H_k^T + R_{k+1}\right] K_{k+1}^T, \tag{4}$$

where

$$K_{k+1} = P_{k+1|k} H_{k+1}^T \left[H_k P_{k+1|k} H_k^T + R_{k+1}\right]^{-1}, \tag{5}$$

H is the jacobian matrix of the observation function in (1), with respect to x, $\frac{\partial h}{\partial x}$, and R is the co-variance matrix of the Gaussian noise present in the measurements.

Link	0	1	2	3	4	5	6
$\boldsymbol{a}$ [mm]	0	0	-15	15	0	0	62.5
$\boldsymbol{d}$ [mm]	-107.74	0	-152.28	0	-137.4	0	16
$\boldsymbol{\alpha}$ [rad]	$\frac{\pi}{2}$	$-\frac{\pi}{2}$	$-\frac{\pi}{2}$	$\frac{\pi}{2}$	$\frac{\pi}{2}$	$\frac{\pi}{2}$	0
$\boldsymbol{\theta}$ [rad]	$-\frac{\pi}{2}$	$-\frac{\pi}{2}$	$-\frac{7\pi}{12}$	0	$-\frac{\pi}{2}$	$\frac{\pi}{2}$	π

Table 1: Actual DH parameters of the iCub arm in the iCub simulator.

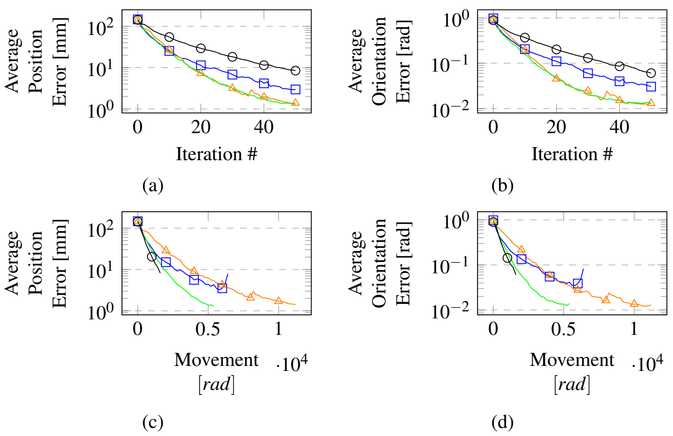

Figure 3: Mean error evolution while performing the calibration routine for different values of δ: (a) Average position error evolution in millimetres at each iteration; (b) Average orientation error evolution in radians at each iteration; (c) Average position error evolution in millimetres with respect to joint movement; (d) Average orientation error evolution in radians with respect to joint movement. Legend: Square - Random; Triangle - $\delta = 1$; No marker - $\delta = 0.4$; Circle - $\delta = 0.1$.

1.4 Cost-sensitive Active Learning

This work aims to choose the best joint configurations to sample the end-effector pose, at each iteration of the calibration routine, to reduce both the body-schema error and movement performed while calibrating. Martinez-Cantin *et al.* [2] successfully used the A-optimality criterion to reduce the number of samples taken. It consists in choosing the joint angles θ which minimise the expected mean squared error of the robot parameters, x, which approximates to minimising the expected trace of the co-variance matrix, P. As described, the cost function is given by $C(\theta) = \mathbb{E}\left[(\hat{x}_{k+1} - x)^T(\hat{x}_{k+1} - x)|z_{1:k}, \theta_{1:k}\right] \approx \mathbb{E}\left[tr(P_{k+1})|z_{1:k}, \theta_{1:k}\right]$.

This work proposes adding constraints to the optimisation problem as in

$$\theta_k^* = \underset{\theta \in \left[\theta_{k-1}^* - \Delta, \theta_{k-1}^* + \Delta\right]}{\operatorname{argmin}} C(\theta), \tag{6}$$

where θ_{k-1} is the previous joint configuration selected and Δ is a vector of size n (number of joints), defining the boundaries of the search space. Considering normalised joint values in the interval $[0,1]$, Δ is defined as $\Delta = \delta \cdot \mathbf{1}_{\boldsymbol{n}}$, where $\mathbf{1}_{\boldsymbol{n}}$ is a unit vector of size n and δ is a tuning parameter that defines the relative movement every joint can perform around the current arm configuration. Since the problem defined in (6) is not convex, it must be solved using a global optimisation method, for which it is used the DIRECT algorithm, proposed in [1].

2 Results

Results were obtained for different values of δ. For each different value, the calibration routine runs fifty times and the plots from Figure 3 show the average values of position and orientation error. At each run, the DH parameters are initialised with values from a uniform distribution, where the means are the actual values of the DH parameters, from Table 1 and the width of the distribution is 30% of the highest value from all the linear and angular DH parameters, 46 mm and 0.94 rad, respectively. The position error is given by the euclidean distance between the predicted position and actual position of the end-effector and the orientation error is given by computing $d(R_A, R_B) = \sqrt{\frac{\|logm(R_A^T R_B)\|_F^2}{2}}$ [rad], between the predicted, R_A, and actual, R_B, end-effector rotation matrices, where *logm* is the principal matrix logarithm and $\|\cdot\|_F$ is the Frobenius norm. Gaussian noise is added to the observations with a standard deviation of 2 mm for the position coordinates and 0.08 radians for the orientation.

Looking at Figures 3(a) and 3(b), the advantages of using the active learning method proposed in [2], corresponding to $\delta = 1$, can be observed by comparing it with selecting random joint configurations to sample, instead of solving (6), since there is a more significant reduction in error at each iteration. In Figures 3(c) and 3(d), the same data is represented, but the x axis represents the movement performed by the arm. It is visible the amount of extra movement performed by the active learning method, $\delta = 1$, almost double of the random method. Restricting movement, making $\delta = 0.4$, yields no performance loss, regarding Figures 3(a) and 3(b), and it is more efficient, as Figures 3(c) and 3(d) show.

3 Conclusions

The results show there is an advantage in restricting movement during the optimisation stage. It is possible to reduce the movement performed by roughly half and still maintain the iteration wise performance. If movement efficiency is a priority, one can restrict the movement even more, at the cost of more iterations. It is worth mentioning, more iterations does not mean lower time-efficiency, since reducing the amount of time spent moving may make up for the extra computing time. Indeed, it will depend on the computing power and the speed at which the arm moves.

For future work, it is planned to obtain results using the iCub cameras and fiducial markers placed on its hand. This comes with observation noise dependant on the observed pose and it should be taken into account when selecting optimal joint configurations.

Acknowledgements

This work was supported by FCT with the LARSyS - FCT Project UIDB/50009/2020 and the PhD grant PD/BD/135115/2017.

References

[1] D. R. Jones, C. D. Perttunen, and B. E. Stuckman. Lipschitzian optimization without the Lipschitz constant. *Journal of Optimization Theory and Applications*, 79(1):157–181, 1993.

[2] Ruben Martinez-Cantin, Manuel Lopes, and Luis Montesano. Body schema acquisition through active learning. *Proceedings - IEEE International Conference on Robotics and Automation*, pages 1860–1866, 2010.

[3] Uriel Martinez-Hernandez, Tony J. Dodd, Mathew H. Evans, Tony J. Prescott, and Nathan F. Lepora. Active sensorimotor control for tactile exploration. *Robotics and Autonomous Systems*, 87:15–27, 2017.

[4] Uriel Martinez-Hernandez, Tony J. Dodd, and Tony J. Prescott. Feeling the Shape: Active Exploration Behaviors for Object Recognition with a Robotic Hand. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 48(12):2339–2348, 2018.

[5] Takamitsu Matsubara and Kotaro Shibata. Active tactile exploration with uncertainty and travel cost for fast shape estimation of unknown objects. *Robotics and Autonomous Systems*, 91:314–326, 2017.

[6] Nancy Reid. *Estimation. In: Lovric M. (eds) International Encyclopedia of Statistical Science*, pages 455–459. Springer, Berlin, Heidelberg, 2011.

[7] Arturo Ribes, Jesus Cerquides, Yiannis Demiris, and Ramon Lopez de Mantaras. Active Learning of Object and Body Models with Time Constraints on a Humanoid Robot. *IEEE Transactions on Cognitive and Developmental Systems*, 8(1):26–41, 2015.

[8] Burr Settles. Active Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 6(1):1–114, 6 2012.

[9] Pedro Vicente, Lorenzo Jamone, and Alexandre Bernardino. Online body schema adaptation based on internal mental simulation and multisensory feedback. *Frontiers Robotics AI*, 3(MAR), 2016.

[10] Rodrigo Zenha, Pedro Vicente, Lorenzo Jamone, and Alexandre Bernardino. Incremental adaptation of a robot body schema based on touch events. *2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics, ICDL-EpiRob 2018*, pages 119–124, 2018.