Robust Probabilistic Robot Arm Keypoint Detection Exploiting Kinematic Knowledge

Lukas Meyer*, Leonard Klüpfel*, Maximilian Durner, and Rudolph Triebel

Deutsches Zentrum für Luft- und Raumfahrt
German Aerospace Center

Institute of Robotics and Mechatronics,
German Aerospace Center, Germany
*The authors contributed equally to this work

Vision-Based Correction of Robot Kinematics

Robots can suffer from **imprecise forward kinematics**, caused by e.g., elasticities, non-linearies, or external loads. We propose a **correction** of erroneous robot kinematics **using vision**:

- Robust detection of **2D robot keypoints** in images using deep-learning: **PK-ROKED**
- **Robot-centric approach**: Steer the network using **prior kinematic knowledge**
- **Uncertainty estimation**: Enable downstream sensor fusion – e.g., with a probabilistic formulation of robot kinematics, see [1]

Figure 1: Our LRU rover on Mt. Etna. PK-ROKED detects keyoints on the Jaco2 arm: Detected keypoints, corresponding uncertainty ellipses, and detected false positives.

Related works: e.g., [2] & [3], but no uncertainty estimates and no usage of prior kinematic knowledge.

Accuracy Evaluation

We evaluate the performance of PK-ROKED on four different data sets and compare its performance **with** and **without prior kinematic knowledge**.

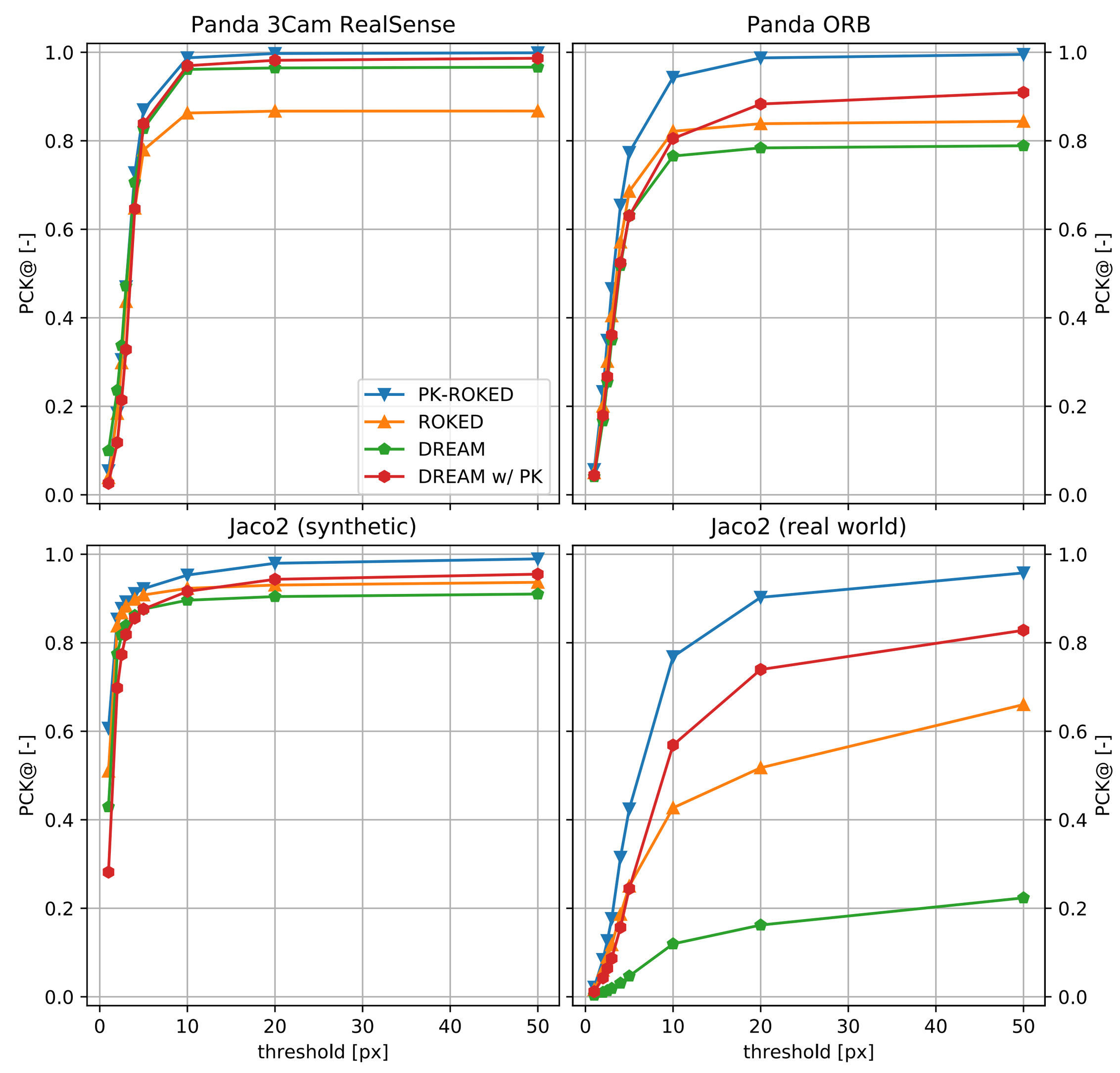

Figure 3: Accuracy evaluation.

- DREAM network [2] as baseline
- We evaluate the performance of DREAM and PK-ROKED both with and without prior kinematic knowledge as input channels
- Evaluation Data:
  - two data sets from [2]
  - synthetic and real data from our Jaco2 arm
- Metric: **percentage of correct keypoints (PCK)** at pixel thresholds w.r.t. the ground truth
- The resulting keypoint locations are the mean of the image coordinates $\mathbf{y}_i^* = f^{\hat{W}_i}(\mathbf{x}^*)$ over all t forward passes:

$$\mathrm{E}(\mathbf{y}^*) = \frac{1}{t}\sum_{i=1}^{t} \mathbf{y}_i^*$$

Network Architecture

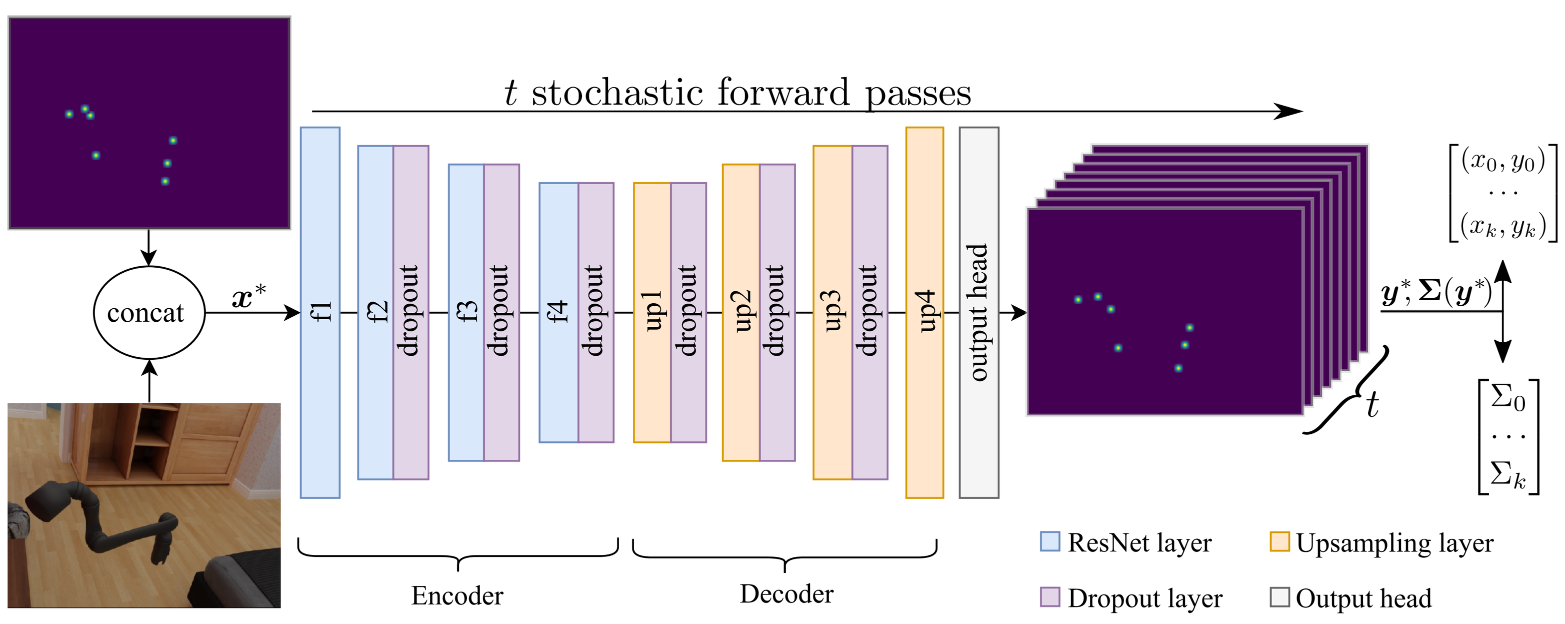

Figure 2: PK-ROKED architecture.

Prior Knowledge - RObot KEypoint Detection **(PK-ROKED)** network – hourglass architecture; training on synthetic data with domain randomization techniques.

- input: RGB image of robot + heatmaps of k keypoints (predicted with robot kinematics)
- output: k **heatmaps** of predicted **keypoint locations**
- dimensions: $640\times480\times(3+k) \rightarrow 40\times30\times2048 \rightarrow 640\times480\times k$
- **active dropout layers** [4] around the bottleneck with $t = 20$ forward passes

Uncertainty by Monte Carlo Dropout

We evaluate several approaches for uncertainty estimation for PK-ROKED.

Explicit Uncertainty Computation

According to [4], the covariance matrix $\boldsymbol{\Sigma}$ can be approximated as:

$$\boldsymbol{\Sigma}(\mathbf{y}^*) \approx \tau^{-1}\mathbf{I}_D + \frac{1}{t}\sum_{i=1}^{t} \mathbf{y}_i^{*T}\mathbf{y}_i^* - \mathrm{E}(\mathbf{y}^*)^T\mathrm{E}(\mathbf{y}^*)$$

- hyperparameter τ encodes the **aleatoric uncertainty (homoscedastic)**
- alternative: instead of τ, learn **aleatoric uncertainty (heteroscedastic)** [5] – additional network output head
- other terms: **epistemic uncertainty**

Implicit Uncertainty Computation from Belief Maps

The **heatmap** of the predicted keypoint locations can also be viewed as a **belief map approximating the probability of keypoint locations**. We stack the heatmaps for all forward passes and binarize the image. We use **image moments** to compute the resulting covariance matrix.

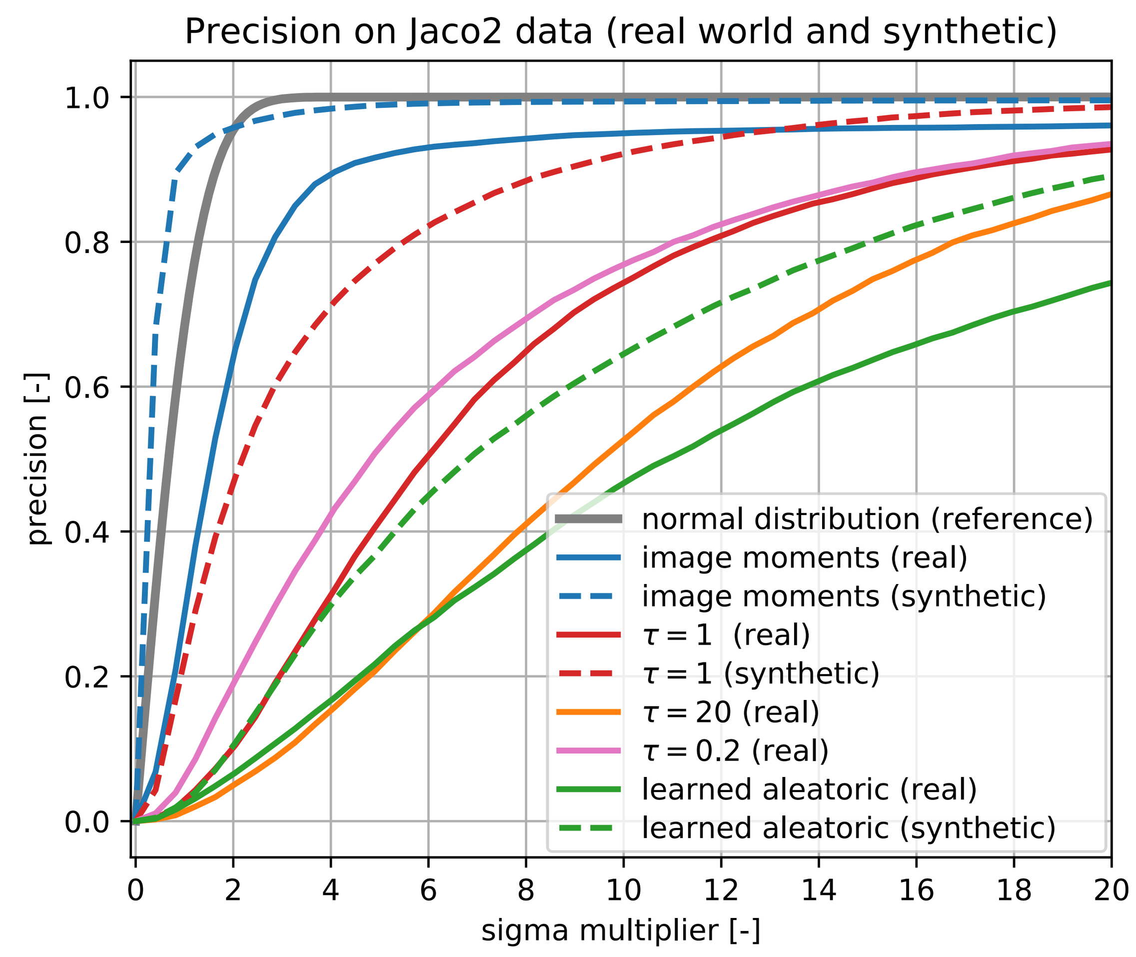

Figure 4: Precision: percentage of keypoints within a $s * \sigma$ boundary for different uncertainty computation approaches.

References

[1] L. Meyer, K. H. Strobl, and R. Triebel, "The Probabilistic Robot Kinematics Model and its Application to Sensor Fusion," in *IEEE/RSJ IROS*, 2022.

[2] T. E. Lee, J. Tremblay, T. To, J. Cheng, T. Mosier, O. Kroemer, D. Fox, and S. Birchfield, "Camera-to-robot pose estimation from a single image," in *IEEE ICRA*, 2020, pp. 9426–9432.

[3] J. Lu, F. Richter, and M. C. Yip, "Pose Estimation for Robot Manipulators via Keypoint Optimization and Sim-to-Real Transfer," in *IEEE RA-L*, vol. 7, Apr. 2022, pp. 4622–4629.

[4] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," in *International Conference on Machine Learning*. PMLR, Jun. 2016, pp. 1050–1059.

[5] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?" in *Advances in Neural Information Processing Systems*, vol. 30. 2017.

Workshop on Probabilistic Robotics in the Age of Deep Learning
IEEE/RSJ International Conference on Intelligent Robots and Systems
Kyoto, Japan, October 27th, 2022