Robot Skill Adaptation via Soft Actor-Critic Gaussian Mixture Models

Iman Nematollahi*, Erick Rosete-Beas*, Adrian Röfer, Tim Welschehold, Abhinav Valada, Wolfram Burgard

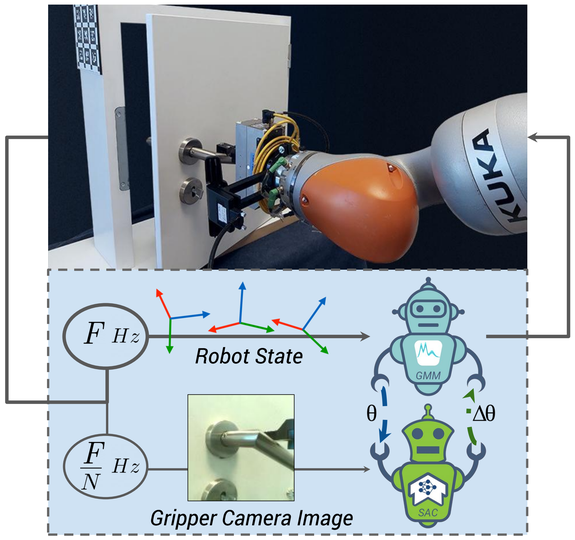
* These authors contributed equally

UNI FREIBURG

AIS Autonomous Intelligent Systems

Introduction

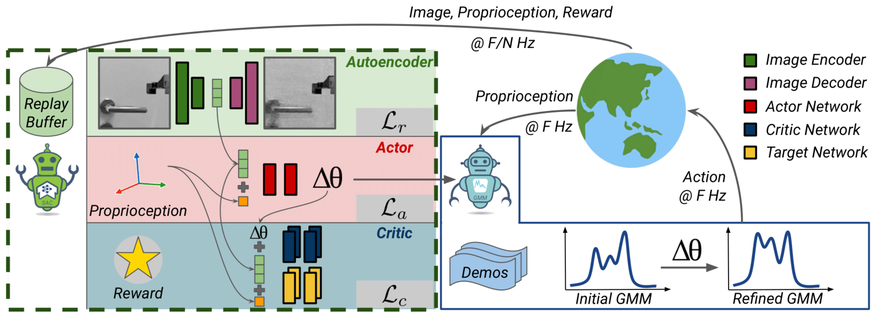
- A Robot acting in the real world needs to think ahead and adapt its repertoire of skills to cope with its noisy perception and dynamics.
- **Goal**: enable robots to learn and refine their skills via trajectories through interactions with the environment.



SAC-GMM

Soft Actor-Critic Gaussian Mixture Models (SAC-GMMs) combine deep reinforcement learning and dynamical systems to:

- Learn a reactive trajectory-based policy from few demos
- Refine this policy through interactions and sparse rewards
- Utilize high-dimensional observation spaces
- Cope with noise in demonstrations and sensory observations
- Mitigate exploration cost and be more sample efficient



Dynamical System: Gaussian Mixture Model Agent

Our Gaussian Mixture Model (GMM) is:

- A trajectory-based policy
- Learned from few demonstrations
- Modeling a first-order ordinary differential equation to map the robot pose to its velocity:

$$\dot{\xi} = f_\theta(\xi) + \epsilon,$$

where ξ defines the robot pose, f_θ is the robot skill model, and ϵ is a zero-mean additive Gaussian noise.

We parametrize the robot skill model f by the parameters of the GMM $\theta = \{\pi_k,\ \mu_k,\ \Sigma_k\}_{k=1}^{K}$, where π_k is the prior, μ_k the mean and Σ_k the covariance of the k-th Gaussian function.

Dynamical System Adaptation: Soft Actor-Critic Agent

Our Soft Actor-Critic (SAC) agent:

- Leverages robot interactions to explore
- Refines the initial GMM policy
- Receives a sparse reward only if the refined policy executes the skill effectively

The state space consists of:

- The robot pose ξ_t
- Latent representations of high-dimensional observations z_t

The action space consists of:

- The desired adaptation in the skill trajectory parameters $\Delta\theta$

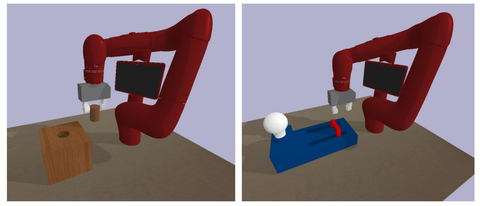
$$\mathbf{s}_t := \{\xi_t, \mathbf{z}_t\}, \quad \mathbf{a}_t := \{\Delta\pi_k, \Delta\boldsymbol{\mu}_k, \Delta\boldsymbol{\Sigma}_k\}_{k=1}^{K}.$$

Our robot skill refinement policy is:

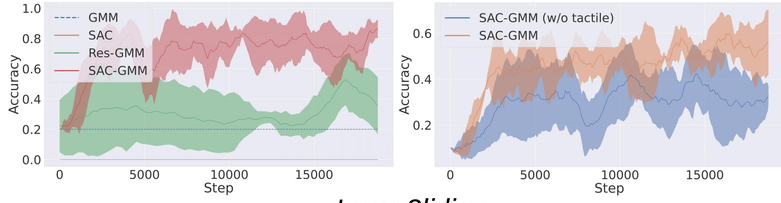
$$\Delta\theta = \pi_\phi(\mathbf{a}_t \mid \mathbf{s}_t).$$

The robot interacts with the world to maximize the expected total reward of the refined skill trajectory.

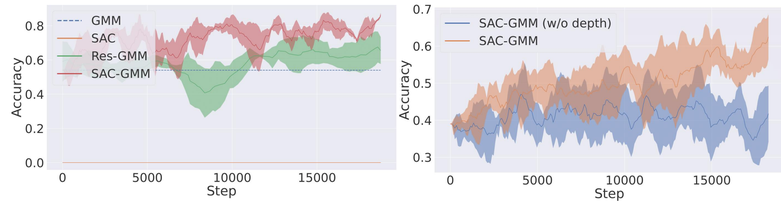
Evaluation and Results



Peg Insertion



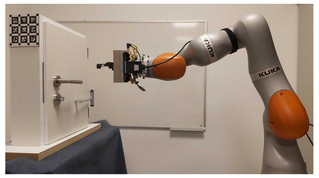
Lever Sliding

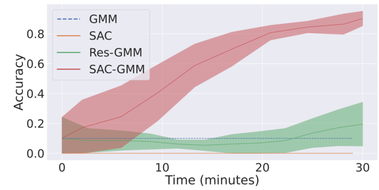


Average success rate over 100 trials per five different random seeds

Model \ Task	Peg Insertion	Lever Sliding
GMM	20%	54%
SAC	0%	0%
Res-GMM	30%	62%
SAC-GMM	**86%**	**81%**

Real-World Door Opening





Conclusions

Our experiments show that SAC-GMMs:

- Refine robot skills through interactions in realistic noisy environments
- Exploit high-dimensional observation spaces to cope better with noise
- Achieve state-of-the-art performance considering accuracy and exploration costs

SCAN ME