Gaussian Process Latent Space Policies for Data-efficient Learning of Robotic Clothing Assistance

Nishanth Koganti[1], Tomohiro Shibata[2], Tomoya Tamei[3], and Kazushi Ikeda[1]

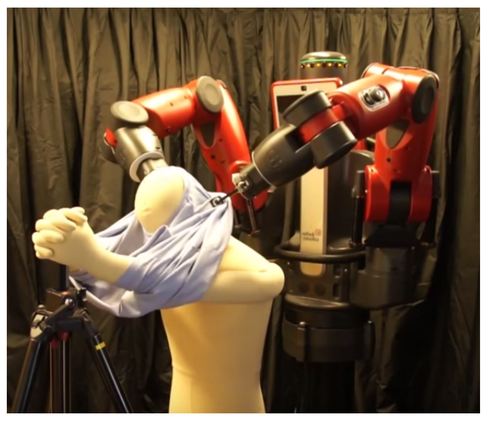
[1]Nara Institute of Science and Technology, Japan,
[2]Kyushu Institute of Technology, Japan, [3]Kobe University, Japan

Introduction

Objective: In this study, we propose the use of Bayesian Gaussian Process Latent Variable Model (BGPLVM) [1] to learn a low-dimensional representation of motor-skills. We implement our framework in a practical setting with a dual-arm robot performing clothing assistance tasks.

Tamei et al. [2] developed clothing assistance robot to perform T-shirt clothing task:

- **Reinforcement Learning** scheme used to acquire motor skills for cloth handling.
- **Via-points** used as policy representation with one via-point as a policy parameter for fast learning.



Problem Description

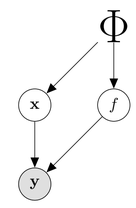
Recent focus has been on LVMs for sample-efficient RL [3, 4].

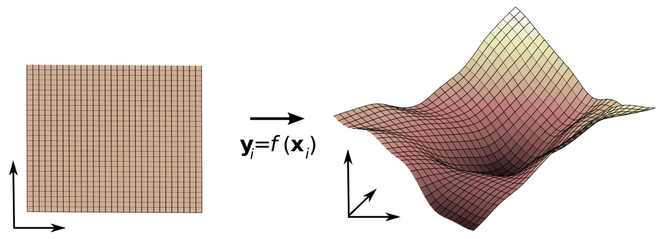
- Latent variable model (Titsias et al., 2010 [1]):

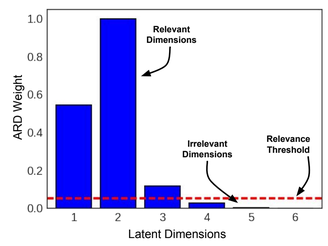
$$\mathbf{y} = f(\mathbf{x}) + \epsilon, \epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$$

- $f : \mathbf{x} \rightarrow \mathbf{y}$: Mapping given by a Gaussian process.
- Automatic dimensionality reduction perfomed using ARD kernel:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{q=1}^{Q} w_q(x_q - x'_q)^2\right)$$



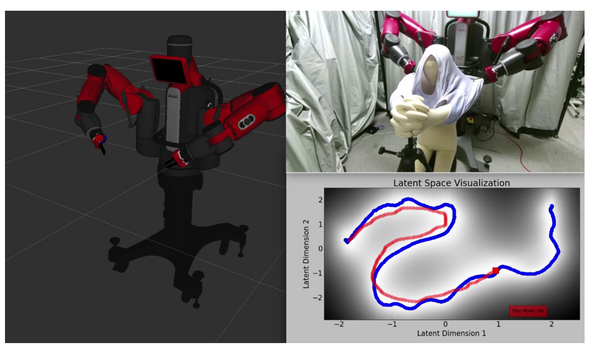




Proposed Method

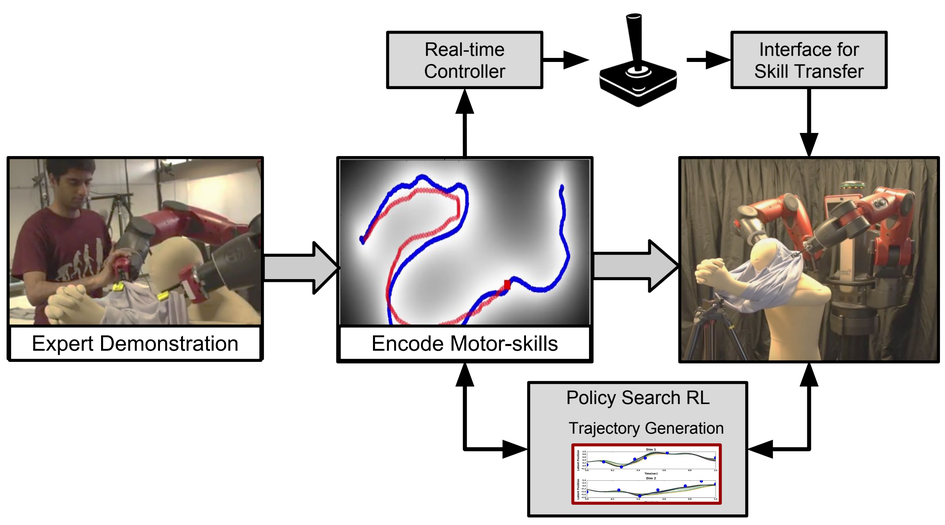
Interface for Latent Space Control

- Inexperienced users can impart noisy demonstrations.
- User-friendly interface: Cursor control in latent space.



User-interface for latent space controller. 2D latent space sufficient to control high DoF dual-arm robot.

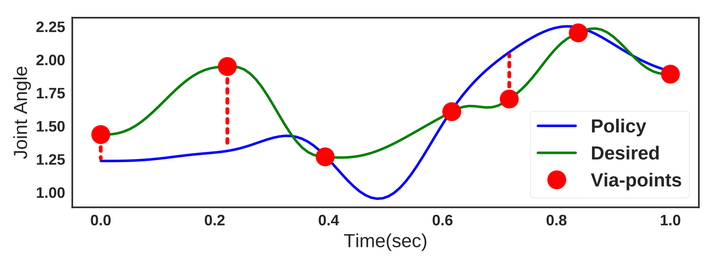
Method Overview



Latent Space Policy Search

Data-efficient Learning: Perform policy search in low-dimensional BGPLVM space.

- Policy Search: PoWER algorithm, Representation: Dynamic movement primitives.



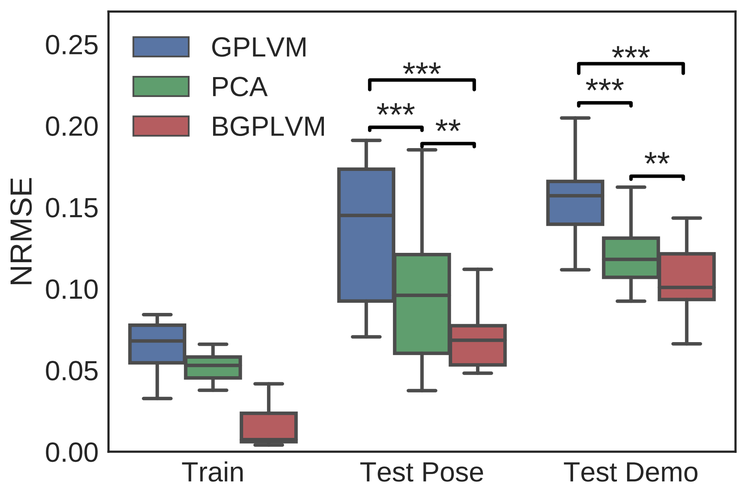
- Represent reward function by distance from desired Via-points to current policy:

$$R(\pi(\theta)) = \sum_{i=1}^{n_{\text{dims}}} \sum_{j=1}^{n_{\text{via}}} \|V_{i,j} - \pi_i(\theta, t_{i,j})\|^2$$

Experimental Results

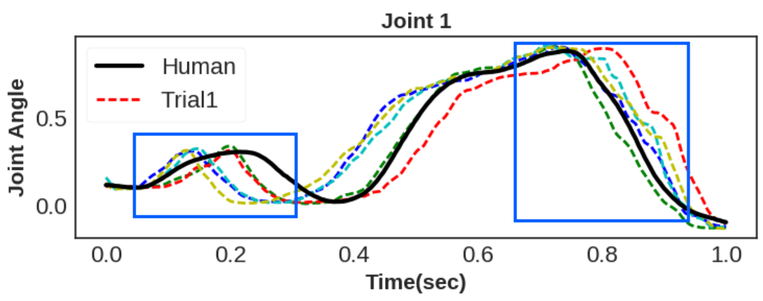
Comparison of Latent Variable Models

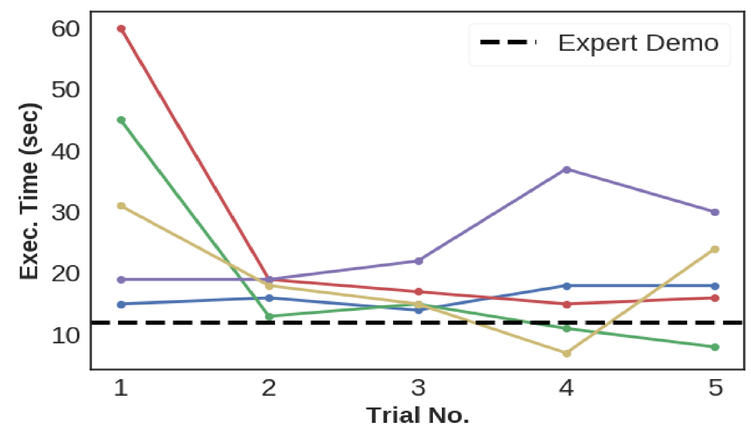
- **Evaluation**: Reconstruction error of LVM with NRMSE.
- **Dataset**: Demonstration of clothing assistance by 3 subjects for 6 postures of mannequin.



Trajectory Generation using Controller

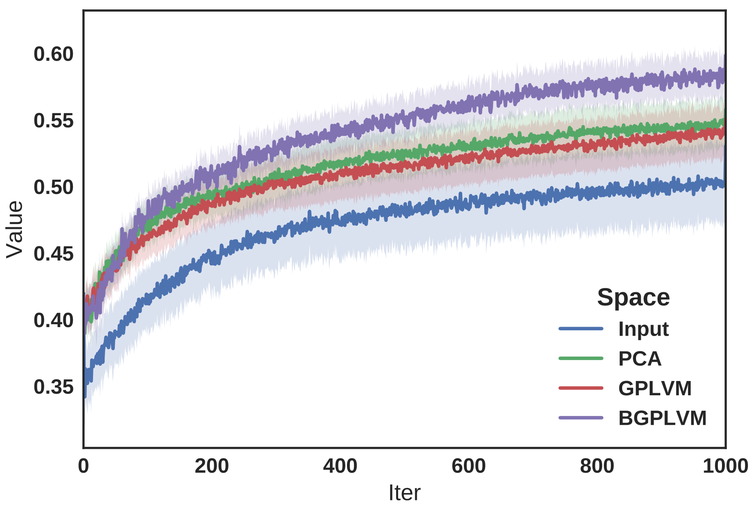
- **Evaluation**: 5 subjects used controller to impart skills.





Evaluation of Latent Space Policy Search

- Apply RL in different latent spaces with same formuation.
- Parameters: $50 \times n_{\text{dims}}$ basis functions.
- PoWER: 10 best iterations used for parameter updates.



Discussion and Future Work

- BGPLVM can encode motor-skills for clothing assistance.
- Future Work: Learn from human preferences using latent space controller.
- Combine dimensionality reduction with policy search.

References

[1] Titsias, Michalis, and Neil D. Lawrence. "Bayesian Gaussian process latent variable model." Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. AISTATS 2010.

[2] Tamei, Tomoya, et al. "Reinforcement learning of clothing assistance with a dual-arm robot." Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on. IEEE, 2011.

[3] SÃęmundsson, SteindÃşr, et al. "Meta Reinforcement Learning with Latent Variable Gaussian Processes." arXiv preprint arXiv:1803.07551 (2018).

[4] Haarnoja, Tuomas, et al. "Latent Space Policies for Hierarchical Reinforcement Learning." arXiv preprint arXiv:1804.02808 (2018).