Bayesian Nonparametric Motor-skill Representations for Efficient Learning of Robotic Clothing Assistance

Nishanth Koganti[1,2]**, Tomoya Tamei**[1]**, Kazushi Ikeda**[1]**, Tomohiro Shibata**[2]
[1] Nara Institute of Science and Technology, Japan [2] Kyushu Institute of Technology, Japan
{nishanth-k,tomo-tam,kazushi}@is.naist.jp, tom@brain.kyutech.ac.jp

Abstract

Motor-skill learning for complex robotic tasks is a challenging problem due to the high task variability. Robotic clothing assistance is one such challenging problem that can greatly improve the quality-of-life for the elderly and disabled. In this study, we propose a data-efficient representation to encode task specific motor-skills of the robot using Bayesian nonparametric latent space learning. We implement our framework in a practical setting with a dual-arm robot performing clothing assistance tasks. We demonstrate that performing motor-skills learning in such task specific latent spaces outperforms learning in the high-dimensional joint configuration space of the robot. Furthermore, our framework can also be used as a tool for learning from demonstration to impart novel skills to the robot.

1 Introduction

Enabling robots to learn motor skills for performing complex tasks is one of the major goals of the AI community. However, existing methods require large number of interactions to learn optimal behavior which is not suitable in practical scenarios such as in the field of Assistive Robotics. Clothing assistance is one such task which is a necessity in the daily life of the elderly and disabled people. Design of a robotic framework involves a motor skills learning framework that can detect and adapt to various failure scenarios. Tamei *et al.* [1] proposed a reinforcement learning (RL) framework for clothing assistance with a dual-arm robot to cloth a soft mannequin with a T-shirt. For tractable learning time, policy update was done using finite difference policy gradient applied to a small subset of policy parameters. This constrained the generalization capability of the algorithm to very different environmental settings such as major changes in the subject's posture or using different clothing articles.

To address this problem, we consider an alternate representation that is more flexible and suitable for robust learning of motor-skills. An efficient approach is the use of dimensionality reduction (DR) along with reinforcement learning (RL) [3], [4]. In this study, we propose an efficient representation of motor skills that relies on the use of Bayesian Gaussian Process Latent Variable Model (BGPLVM) [2]. We hypothesize that the latent space learned using BGPLVM is an efficient search space to learn novel motor-skills. The mapping from latent space to data space is modeled using a Gaussian Process (GP) [5] allowing data-efficient learning. This is crucial to ensure the robot can generalize well to unseen environmental settings within few interactions making it suitable for a feasible social implementation of robotic clothing assistance. We apply our proposed method to a clothing assistance setup as shown in Figure 1. We demonstrate that performing policy search in this latent space outperforms learning in the high-dimensional action space. We further present the design of a real-time controller from the BGPLVM latent space that can be used as a tool for Learning from Demonstration (LfD). The experimental results indicate a promising reinforcement learning framework that can be used for challenging tasks such as robotic clothing assistance.

30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

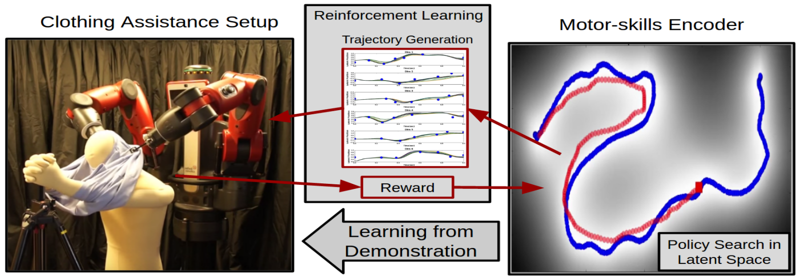


Figure 1: Overview of proposed framework. Figure on the left shows the setup for clothing assistance and the figure on the right shows the latent space encoding the motor skills.

2 Proposed Method

Motor skills for clothing assistance are given by high dimensional joint angle trajectories of the dual-arm robot. The robot also has to maintain several task space constraints such as coupling with a clothing article along with safe human-robot interaction as shown in Figure 1. To address these problems, we propose the use of BGPLVM for learning a low-dimensional latent space through a nonlinear mapping to the kinematic space. There are several motivating factors for this choice. The Bayesian treatment avoids overfitting to the training data and the use of an ARD kernel in the GP mapping leads to inference of the inherent dimensionality to encode motor-skills.

2.1 Bayesian Nonparametric Latent Space Learning

Bayesian Gaussian Process Latent Variable Model (BGPLVM) is a dimensionality reduction technique proposed by Titsias *et al.* [2] derived from the probabilistic model where the observations, $\mathbf{Y} = \{\mathbf{y}_i \in \mathbb{R}^D\}_{n=1}^N$, are assumed to be generated from latent inputs $\mathbf{X} = \{\mathbf{x}_i \in \mathbb{R}^q\}_{n=1}^N$ through a noisy process,

$$\mathbf{y}_i = f(\mathbf{x}_i) + \epsilon,\ \epsilon \sim \mathcal{N}(\mathbf{0}, \beta^{-1}\mathbf{I}) \tag{1}$$

The mapping function f is modeled using a Gaussian Process (GP) suitable for performing non-linear dimensionality reduction with use of a non-linear kernel function. The conditional likelihood for the generative model is given by:

$$p(\mathbf{Y}|\mathbf{X}, \Phi) = \prod_{d=1}^{D} \mathcal{N}(\mathbf{y}_{:,d}|\mathbf{0}, \mathbf{K} + \beta^{-1}I) \tag{2}$$

where $\mathbf{X}, \Phi$ are the unknown latent points and hyperparameters for the GP mapping that need to be inferred. $\mathbf{K}$ is the kernel matrix constructed from the latent points. In the generative model, the latent positions need to be marginalized out for having a purely Bayesian treatment. Titsias *et al.* [2] proposed a variational inference approach to compute a tractable lower bound for the marginalization thereby inferring a posterior distribution on the latent space which avoids the problem of overfitting.

For performing automatic model selection of the latent space dimensionality, the Automatic Relevance Determination (ARD) Kernel [5] can be used in the GP mapping:

$$k_{\text{ard}}(\mathbf{x}_i, \mathbf{x}_j) = \sigma_{ard}^2 \exp\left(-\frac{1}{2}\sum_{k=1}^{q} \alpha_q (x_{i,k} - x_{j,k})^2\right) \tag{3}$$

The ARD weights α_q describe the relevance of each dimension and zero weight indicates complete irrelevance. Maximizing the marginal likelihood w.r.t. these weights allows the inference of latent space dimensionality. The inference for unseen test data can now be performed through a Bayesian formulation instead of relying on a MAP estimation of the latent space. The predictive distribution is given by the ratio of two marginal likelihoods, both of which can be approximated using the variational inference technique:

$$p(\mathbf{y}^*|\mathbf{Y}) = \frac{\int p(\mathbf{y}^*, \mathbf{Y}|\mathbf{x}^*, \mathbf{X}) p(\mathbf{x}^*, \mathbf{X}) d\mathbf{X} d\mathbf{x}^*}{\int p(\mathbf{Y}|\mathbf{X}) p(\mathbf{X}) d\mathbf{X}} \tag{4}$$

2.2 Motor Skills Representation using BGPLVM

In this section, we present the formulation used to apply BGPLVM to clothing assistance skills. We consider the clothing task where a dual-arm robot clothes a soft mannequin with a T-shirt which is initially resting on the mannequin's arms. We represent the robot poses in two ways i.e. 1) Kinematic representation given by joint angles $D_K = 14$ and 2) Task space representation given by the end-effector pose of both arms with Cartesian position $P_X, P_Y, P_Z \in \mathbb{R}^3$ and orientation $O_X, O_Y, O_Z, O_\omega \in \mathbb{R}^4$ forming a 14-dimensional space $D_T = 14$. We set the dimension of the latent space as $q = 5$. However, the dimensionality is eventually inferred through the training of the ARD kernel weights as explained in Section 2.1.

There can be several types of failure scenarios when the robot performs clothing tasks. To recover from these failures, not only is the trajectory of the robot important, but also the speed of execution. Imparting these skills through kinesthetic movement of the arms can be difficult for a bulky robot and could lead to noisy demonstrations. To address this problem, we have implemented a real-time controller that gets an input signal from the BGPLVM latent space. The BGPLVM model learns a mapping from the low dimensional latent space to the robot kinematic space such that a trajectory of latent points generates a trajectory on the dual-arm robot. This interface can be used as a tool for Learning from Demonstration (LfD) where the necessary clothing skills are imparted to the robot by using cursor control over the latent space. The controller was able to reproduce the clothing demonstration even when the latent trajectory was different from the training latent points.

2.3 Latent Space Reinforcement Learning

In this section, we formulate a policy search framework in the BGPLVM latent space. The objective is to learn a high-dimensional robot trajectory for performing the clothing task on an unseen posture of the mannequin by searching within this latent space. Firstly, a dataset of successful clothing assistance trajectories is used to train a latent space that encodes the motor skills. Each of the trajectory is now transformed into a sequence of points in the latent space forming latent space trajectories. The policy search is performed using the cross-entropy method (CEM) [6] which is a general algorithm for approximately solving global optimization problems. It has been widely used in robotics to learn controllers for performing various tasks.

We consider Dynamic Movement Primitives (DMP) [7] as policy representation. It is the combination of a point attractor dynamical system and a non-linear forcing term ($f(s)$) learned using Locally Weighted Regression (LWR) [8]:

$$\begin{aligned} \tau\ddot{x} &= K(g-x) - D\dot{x} + (g-x_0)f \\ f(s) &= \frac{\sum_i w_i \psi_i(s) s}{\sum_i \psi_i(s)}, \text{ where } \tau\dot{s} = -\alpha s \end{aligned} \tag{5}$$

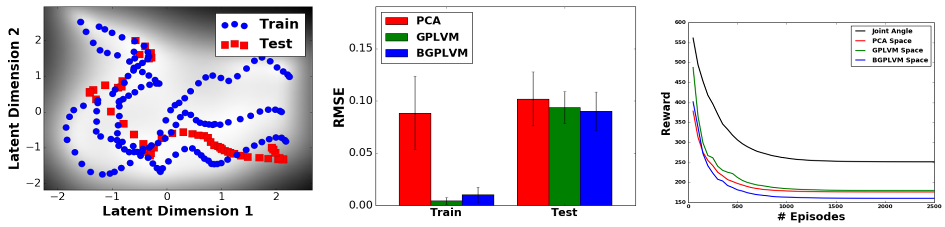
We train a DMP on the latent points corresponding to one of the training trajectory and the LWR weight coefficients are used as policy parameters which are modified to generalize to unseen environmental settings. The cost function for policy improvement is designed in the high-dimensional action space. In the current setting, we obtain a demonstration for the unseen posture and consider this as the optimized trajectory to be achieved. The optimized trajectory is efficiently encoded by via-points extracted using the minimum jerk criterion [9] that the robot needs to pass through. The cost function is given by the sum of all errors between the reconstructed robot trajectory and the desired via-points:

$$R(\tau) = \sum_{i=1}^{n_{\text{dims}}} \sum_{j=1}^{n_{\text{via}}} \|V_{i,j} - x_i^{recons}(t_{i,j})\|^2 \tag{6}$$

where $R(\tau)$ is the total reward for trajectory τ, $V_{i,j}$ is the j^{th} via point of i^{th} dimension and $x_i^{recons}(t)$ is the value at time t for i^{th} dimension of reconstructed trajectory.

3 Results

In this section, we present the applicability of BGPLVM for encoding clothing skills. The evaluation dataset contains 4 clothing trials that were obtained for 4 different postures of the mannequin wherein the elevation of the arms were varied forming different angles $\{65^o, 70^o, 75^o, 80^o\}$ with the



(a) Latent Space (b) Reconstruction Error (c) Learning Curves for RL

Figure 2: Experimental results of applying BGPLVM to clothing assistance demonstrations for the joint angle scenario.

mannequin's body. These arm inclinations cover the entire range for which the robot can successfully perform clothing tasks thereby spanning all feasible postures. The clothing trial for 75^o was left out as the test data and latent variable models were trained for the remaining 3 clothing trials. The ARD weights on training resulted in 3 active latent dimensions for the joint angle scenario and 2 latent dimensions for the end effector scenario. This implies that the encoded motor skills varies depending on the input action space. The latent space learned for the joint angle scenario is shown in Figure 2a. The computational complexity of the proposed framework depends on the forward mapping from the BGPLVM latent space to joint angle space which scales with the size of the training dataset and the number of inducing points used to approximate the dataset i.e. $O(nm^2)$ [2]. In the current framework, the training dataset remains fixed during runtime and so precomputed values of the kernel matrix can be used to reduce time-complexity to $O(n)$.

The latent dimensions captured by BGPLVM a specific aspect of the clothing motor-skills. For the joint angle scenario, the most significant dimension captured the horizontal motion of the arms along the mannequin while maintaining the constraints for clothing. The second dimension captured various vertical motions of pulling up the T-shirt in the beginning and pulling it down along the torso at the end. The third dimension captured variations in joint configurations across the demonstrations and could explain the constraint of safe human-robot interaction. A video demonstration with the exploration of the latent dimensions is available at `https://youtu.be/2TCZnt_qBHU`. The The predictive performance of BGPLVM was evaluated by comparing the reconstruction error with two other latent variable models (LVM) i.e. Principal Component Analysis (PCA) [10] and Gaussian Process Latent Variable Model (GPLVM) [11]. Figure 2b shows that BGPLVM has the best performance.

We further demonstrate the effectivity of performing policy search in the BGPLVM latent space by comparison with search in the high-dimensional action space and in latent spaces learned using PCA and GPLVM. CEM was run for 50 iterations where 50 rollouts were performed in each iteration. The rollouts were generated from DMP with random sampling of LWR weight parameters. The top 5 rollouts were retained for updating the mean and variance of policy parameters at each iteration. Figure 2c shows the learning curves for all formulations and the BGPLVM scenario performs better than all other search spaces. PCA relies on a linear model thereby constraining its capability. GPLVM has the worst performance among the latent spaces as relies on a MAP estimate of the latent space and it is overfitting to the training data. This validates our hypothesis that encoding motor-skills using Bayesian nonparametric latent variable models is an effective parametrization for learning.

4 Conclusion

In this study, we have presented the use of Bayesian Gaussian Process Latent Variable Model (BGPLVM) as a representation for encoding motor-skills to perform clothing assistance task. The experimental results indicate our method as a promising approach for learning in combination with reinforcement learning. We further demonstrated its applicability as an intuitive and user-friendly tool for Learning from Demonstration (LfD). There can be several extensions to this work. A latent space representation can be learned from multiple observation spaces using Manifold Relevance Determination (MRD) [12]. Our future work will be to learn models that incorporates visual information of the relationship between the human and cloth. The long term goal is to develop a data-efficient policy search reinforcement learning framework that unifies learning the latent space simultaneously with policy learning.

Acknowledgement

This work was supported in part by ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan) 2015-PM07-36-01. This work was also supported in part by the Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (No. 16H01749).

References

[1] T. Tamei, T. Matsubara, A. Rai, and T. Shibata, "Reinforcement Learning of Clothing Assistance with a Dual-arm Robot" in Proc. of the IEEE-RAS Intl. Conf. on Humanoid Robots (Humanoids), 2011.

[2] Titsias, M. K., and Lawrence, N. D, "Bayesian Gaussian process latent variable model" in Proc. of the Intl. Conf. on Artificial Intelligence and Statistics, 2010.

[3] Bitzer, S., Howard, M., and Vijayakumar, S, "Using dimensionality reduction to exploit constraints in reinforcement learning" in Proc. of the IEEE Intl. Conf. on Intelligent Robots and Systems (IROS), 2010.

[4] Luck, K. S., Neumann, G., Berger, E., Peters, J., and Ben Amor, H, "Latent space policy search for robotics" in Proc. of the IEEE Intl. Conf. on Intelligent Robots and Systems (IROS), 2014.

[5] Williams, C., and Rasmussen, C. E., "Gaussian processes for machine learning" by the MIT Press 2.3 (2006): 4.

[6] De Boer, P., Kroese, D. P., Mannor, S., and Rubinstein, R. Y., "A tutorial on the cross-entropy method" in Annals of operations research 134, no. 1 (2005): 19-67.

[7] Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., and Schaal, S., "Dynamical movement primitives: learning attractor models for motor behaviors", in Neural computation 25, no. 2 (2013): 328-373.

[8] Schaal, S., and Atkeson, C. G., "Constructive incremental learning from only local information", in Neural computation 10.8 (1998): 2047-2084.

[9] Wada, Y., and Kawato, M., "A theory for cursive handwriting based on the minimization principle", in Biological Cybernetics 73.1 (1995): 3-13.

[10] Jolliffe, I., "Principal component analysis" by John Wiley and Sons, Ltd, 2002.

[11] Lawrence, N., "Probabilistic non-linear principal component analysis with Gaussian process latent variable models" in Journal of Machine Learning Research 6.Nov (2005): 1783-1816.

[12] Damianou, A., Ek, C., Titsias, M., and Lawrence, N, "Manifold relevance determination" in Proc. of the Intl. Conf. on Machine Learning, 2012.