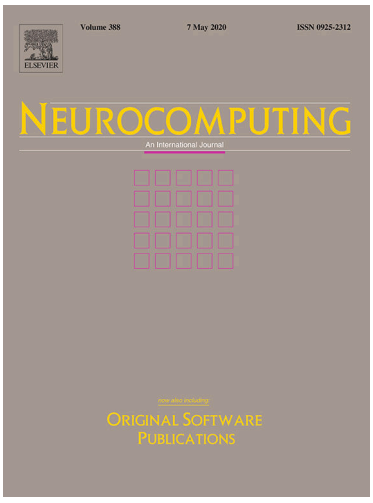
Journal Pre-proofs



Learning Robot Anomaly Recovery Skills from Multiple Time-driven Demonstrations

Hongmin Wu, Wu Yan, Zhihao Xu, Shuai Li, Xuefeng Zhou

PII: S0925-2312(21)01220-0
DOI: https://doi.org/10.1016/j.neucom.2021.08.036
Reference: NEUCOM 24201

To appear in: *Neurocomputing*

Received Date: 27 January 2021
Accepted Date: 8 August 2021

Please cite this article as: H. Wu, W. Yan, Z. Xu, S. Li, X. Zhou, Learning Robot Anomaly Recovery Skills from Multiple Time-driven Demonstrations, *Neurocomputing* (2021), doi: https://doi.org/10.1016/j.neucom.2021.08.036

Learning Robot Anomaly Recovery Skills from Multiple Time-driven Demonstrations

Hongmin Wu[a,b], Wu Yan[a], Zhihao Xu[a,b], Shuai Li[c], Xuefeng Zhou[a,*]

[a] *Institute of Intelligent Manufacturing, Guangdong Academy of Sciences, Guangdong Key Laboratory of Modern Control Technology, China.*
[b] *Pazhou Lab, Guangzhou, 510330, China.*
[c] *School of Engineering, Swansea University, Swansea, United Kingdom.*

Abstract

Robots are prone to making anomalies when performing manipulation tasks in unstructured environments, it is often desirable to rapidly adapt the robotic behavior to avoid environmental changes by learning from experts' demonstrations. We propose a framework for learning robot anomaly recovery skills from time-driven demonstrations based on a Gaussian process regression with prior mean derived by Gaussian mixture regression, named as mean-prior GPR (MP-GPR), which allows an end-user to adjust the anomalous trajectory intuitively by simultaneously considering the variability of the demonstrations and the adaptation of recovery skills. Evaluations are divided into two phases, a benchmarking dataset with robot reaching, pushing, writing, and pressing tasks are first used to verify the path accuracy and variability, and then a real-time robot bin-picking task for evaluating the adaptation of the framework. Our method has a fair comparison with probabilistic-based methods in the field of robot learning from demonstrations, including Gaussian mixture regression, probabilistic movement primitives, and kernelized movement primitives. The results indicate that our proposed method can efficiently encode the variability from multiple demonstrations and rapidly anomaly recovery skills learning by modulating a learned trajectory to safe via-points.

Keywords: Robot skill learning, Gaussian process, Robot recovery

*Corresponding author: Xuefeng Zhou (xf.zhou@giim.ac.cn)

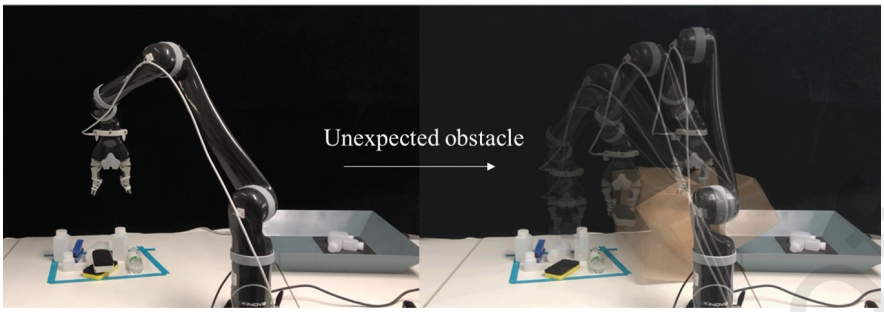


Figure 1: Robots are prone to making anomalies when performing manipulation tasks in unstructured environments, e.g. robot collides with the unexpected obstacle.

1. Introduction

With the rapid development of human-robot collaboration [1], robots may encounter unseen situations such that they are prone to making anomalies when performing manipulation tasks in unstructured environments [2, 3], e.g. robot collides with the unexpected obstacle during object transporting, as shown in Figure 1. Besides the detection and classification of performing anomalies [4, 5], robots that can recover appropriately to common anomalies have the potential to provide more effective and safer collaboration.

In our previous works [6], we proposed a SPAIR (Sense-Plan-Act-Introspect-Recover) framework for learning transition policy for robot anomaly recovery, which mainly investigated the problem of anomaly detection, classification as well as recovery policies during robot manipulation tasks. We investigated the adaptation recovery policy based on dynamical movement primitives (DMP) that can effectively adapt to the target points, but it can't lean the variability from human demonstrations and pass through the via-points such that the safety considerations are insufficient. These experiences make us investigate a method that can modulate the anomalous trajectory by given safe via-points and easily adapt to the environmental changes based on human demonstrations.

For learning recovery skills, the environmental changes or human interven-

tion can be considered as new task constraints, which require the robot to adapt its trajectory to pass through constrained via-points or end-points. To satisfy those new constraints, the conventional time-consuming robot programming methods performed by engineers have become a key obstruction that influences the development of industrial manufacturing towards intelligence and flexibility [7]. Motivated by the continued success of robot learning from demonstrations (LfD) [8, 9] to generate trajectory derived from observations of a human's own performance thereof, rather than requiring users to analytically decompose and manually program the desired behavior. Thus, the generation of adaptive trajectory for anomaly recovery by learning from human demonstrations plays a very important role as the robot works aside with human [10].

In the context of LfD, the typical algorithms such as Dynamical Movement Primitives (DMP) [11], Gaussian mixture regression (GMM) [12, 13], and probabilistic movement primitives (ProMP) [14, 15] have been developed to generate desired trajectories, which can directly learn variability from time-driven trajectories, but they introduce many open parameters (including basis functions and weighting coefficients) and can't address the problem with high-dimensional inputs. To address the problem of learning of high-dimensional trajectories, the GMM/GMR scheme has been proposed [16], where GMM is employed to model the joint distribution of input variables and complemented with GMR to retrieve the desired trajectory. The GMR is employed not only for modeling the movement but also encoding the EMG-based stiffness for robot skill learning [17, 18]. Despite GMM/GMR can effectively learn the variability of high-dimensional demonstrations to generate adaptive trajectories, it is difficult to pass through via-points and end-points. Moreover, it is worth mentioning that DMP can only adapt trajectories towards end-points, but can't satisfy the via-points constraints. ProMP can adapt trajectories towards new via-points and end-points simultaneously.

Thus, there are two properties that should be considered in robot learning recovery skills, including the variability of the demonstrations and the extrapolation of the reproductions. Note that we consider the reproduced trajectories

to pass through desired via-points and end-points as the extrapolation capability. Generally, learning the variability of the demonstrations for redefining parts of the demonstrated trajectories according to the environmental changes, while maintaining the general trajectory shape as in demonstrations. Moreover, improving the extrapolation by modeling the presence of uncertainty in the reproduction for adapting the robot behavior such that compliance at different phases of the tasks [19, 20].

As those two properties are taken into further consideration, the existing LfD methods can be typically divided into three main categories, including the dynamical system-based methods, e.g. DMP; the probabilistic methods, e.g. ProMP, GMM/GMR, kernelized movement primitives (KMP) [21] or Gaussian process regression (GPR) [22, 23], and the dynamical system-based probabilistic methods, e.g. Stable Estimator of Dynamical Systems (SEDS) [24]. Due to the covariance matrices of the prediction distributions computed by GMR, ProMPs and KMP can precisely encode the variability of the predicted trajectory, making the probabilistic methods obtain increasing attention in recent years. However, most of the aforementioned probabilistic methods without taking both the variability of the demonstrations and the extrapolation of the reproductions into consideration simultaneously.

2. Related Work

Some studies and reviews have been undertaken to explore the robot learning recovery skills from anomalies or environmental changes in the last five years, the fields as diverse as robot assembly [25, 26, 27], robot grasping and obstacle avoidance [28, 29], quadrupedal robot [30] as well as legged robot [31]. Besides the robust robot control algorithms [32], corrective trajectories by learning from human demonstrations has been widely used due to an end-user with the straightforward comprehension of task. Among them, the most popular representation to encode motion from demonstrated trajectories is DMP, as described in [33], because of its well-defined attractor properties and a non-linear

force term that allows modeling of almost any complex trajectories. A reversible DMP [34] is also proposed, which has the significant potential for learning recovery skills that modelling the reversible execution, due to its reversibility of a learned trajectory. A robot learning system based on adaptive neural control and DMPs is proposed in [35, 36] for considering both motion generation and trajectory tracking to compensate for the effect caused by the dynamic environments. Moreover, GMM/GMR, ProMP, KMP are proposed as an alternative recovery skills representation in probabilistic formulation, they learn a trajectory distribution from multiple demonstrations and modulate the movement by conditioning on desired target states. Therefore, a method considering both of the variability of the demonstrations and the extrapolation of the reproductions is proposed for learning robot anomaly recovery skills.

To meet the requirements of variability and extrapolation, [37] propose to use mixture density networks (MDN) in an imitation learning context to predict both variability and uncertainty, but it assumes the outputs are uncorrelated. Moreover, Heteroscedastic Gaussian Processes (HGP) [38] introduce an input-dependent noise model into the regression problem, but require the training of separate HGP models when tasks with multiple outputs, thus output correlations are not straightforward to learn in the standard formulation. Recently, [39] first discuss a fundamental difference between the type of variance encapsulated by the predictions between GMR and GPR, in which GMR relying on a previously trained GMM, computes full covariance matrices encoding the correlation between output variables such that the predictions measure the variability of demonstration, while GPR quantify the uncertainty during reproduction. With this finding of GMR and GPR, [20] introduce an uncertainty-aware probabilistic model using KMP for predicting both variability and uncertainty with a single imitation learning framework, which takes the robot uncertainty about its actions, the variability and correlations in the data into account, to design optimal controllers from demonstrations. Besides the extensions of KMP, [23] propose a novel multi-output Gaussian Process (MOGP) based on GMR that can encapsulate the variability retrieved from the demonstrations in the covariance of

the MOGP and implement the trajectories modulation given new observations, which we have adopted. Moreover, a GP-based model for robot learning from demonstration is proposed in recent work [40], which allows to generalize over multiple demonstrations, and encode variability along the different phases of the task.

Thus, we attempt to learn the robot recovery skills when it collides with an unexpected static obstacle, where the behavior not only tracking the via-points by leaning the variability based on the GMM/GMR scheme but also easily adapt to task uncertainty with GPR. That's, the proposed method can learn the variability of demonstrations for generating adaptive trajectory with constrained points, and generalize for the various tasks with high extrapolation capability.

3. Methods

3.1. Problem Formulation

In this paper, we address the question of whether it is possible to automatically interpret a corrective trajectory by passing through given safe start-points, via-points as well as end-points, made by an end-user, while still maintaining the general trajectory shape (mean and variability) as in the demonstrations. Here, we mainly focus on the time-driven demonstrated trajectories such that time t considered as input and the Cartesian positions of a robot end-effector $\mathbf{y} = \{x, y, z\}$ considered as outputs in the following model explanations.

3.2. Gaussian Mixture Regression

Gaussian Mixture Regression (GMR) constructs a sequence of Gaussian Mixture Models (GMM) for the joint density of the input datapoints, and then derive conditional density and regression functions from each model [41]. That's, GMR estimates the distribution of output data given input data by relying on the learned joint distribution by GMM. The computational complexity of GMR is mainly dependent on the number of GMM components, which is the

only parameter that needs to be specified. Therefore, GMR is well adapted for real-time applications.

In the scheme of learning from demonstrations, we assume a dataset with N datapoints/observations collected when a robot performs a task under an end-user's demonstration. Generally, each observation contains the time-step $t_i, i \in \{1, 2, ..., N\}$ and its corresponding multimodal information $y_i \in \mathbb{R}^D$ of D dimensions, which are $i\text{-}th$ realizations of dataset $Y = \{t_i, y_i\}_{i=1}^N$. Hence, the multimodal datapoints are used to encode the GMM with K components by

$$\begin{aligned} p(\mathbf{y}_i) &= \sum_{k=1}^{K} p(k)p(\mathbf{y}_i|k) \\ p(k) &= \pi_k \\ p(y_i|k) &= \mathcal{N}(\mathbf{y}_i; \mu_k, \mathbf{\Sigma}_k) \\ &= \frac{1}{\sqrt{(2\pi)^D |\mathbf{\Sigma}_k|}} e^{-\frac{1}{2}((\mathbf{y}_i - \mu_k)^T \mathbf{\Sigma}_k^{-1} (\mathbf{y}_i - \mu_k))} \end{aligned} \tag{1}$$

where $p(k)$ is the prior, $p(\mathbf{y}_i|k)$ is the conditional probability density function, $\mathcal{N}(\mathbf{y}_i; \mu_k, \mathbf{\Sigma}_k)$ denotes the normal distribution of an observation $\mathbf{y}_i$. The parameters of each component in Equation 1 are described as prior probability π_k, mean vector μ_k, and covariance matrix $\mathbf{\Sigma}_k$. The GMM parameters can be trained in a batch mode using the Expectation-Maximization (EM) algorithm [42], which iteratively optimizes the model using maximum likelihood estimates.

GMR computes the conditional distribution of the GMM joint distribution to infer the output vector (multimodal observations) corresponding to a given input vector (time steps). A GMM with parameters $\mathbf{\Theta} = \{\pi_k, \mu_k, \mathbf{\Sigma}_k\}$ encoding the demonstrated trajectories $Y = \{t_i, y_i\}_{i=1}^N$, the temporal and spatial values of the Gaussian component k are described by

$$\begin{aligned} \mu_k &= \{\mu_k^t, \mu_k^y\}; \\ \mathbf{\Sigma}_k &= \begin{bmatrix} \Sigma_k^{tt} & \mathbf{\Sigma}_k^{ty} \\ \mathbf{\Sigma}_k^{yt} & \mathbf{\Sigma}_k^{yy} \end{bmatrix} \end{aligned} \tag{2}$$

the expected $\mathbf{y}_k$ distribution of given t_k is formulated as

$$p(\mathbf{y}_k|t_k) = \mathcal{N}(\mathbf{y}_k; \hat{\mathbf{y}}_k, \hat{\mathbf{\Sigma}}_k^{yy}) \tag{3}$$

with the componentwise conditional means and covariances

$$\begin{aligned}\hat{\mathbf{y}}_k &= \mu_k^y + \mathbf{\Sigma}_k^{yt}(\mathbf{\Sigma}_k^{tt})^{-1}(t - \mu_k^t) \\ \hat{\mathbf{\Sigma}}_k^{yy} &= \mathbf{\Sigma}_k^{yy} - \mathbf{\Sigma}_k^{yt}(\Sigma_k^{tt})^{-1}\mathbf{\Sigma}_k^{ty}\end{aligned} \tag{4}$$

Therefore, the distribution of output data $\mathbf{y}$ given input data t can be estimated with the conditional means and covariances of the multivariate Gaussian distributions associated with individual components using the laws of total mean and covariance, so that

$$\begin{aligned}p(\mathbf{y}|t) &= \sum_{k=1}^{K} \beta_k(t)\mathcal{N}(\mathbf{y}; \hat{\mathbf{y}}_k, \hat{\mathbf{\Sigma}}_k^{yy}) \\ \beta_k(t) &= \frac{p(k)p(t|k)}{\sum_{i=1}^{K} p(i)p(t|i)}\end{aligned} \tag{5}$$

where $p(t|k) \sim \mathcal{N}(\mu, \Sigma)$ stands for the probability density function with mean μ and covariance matrix Σ, $\beta_k(t)$ denotes the probability of the component k to be responsible for time t. The observation can be estimated at different time step t with the mean $\hat{\mathbf{y}} = \sum_{k=1}^{K} \beta_k(t)\hat{\mathbf{y}}_k$, covariance matrix $\hat{\mathbf{\Sigma}}^{yy} = \sum_{k=1}^{K}(\beta_k(t))^2\hat{\mathbf{\Sigma}}_k^{yy}$.

3.3. Gaussian Process Regression

Gaussian Processes (GPs) are non-parametric technique with explicit uncertainty models for learning deterministic input-output relationship based on a Gaussian prior over potential objective functions [43]. GPs rather try to infer how all the measured data is correlated, instead of fitting the parameters of a selected basis functions. An individual GP is a Gaussian random function, and is fully specified by a mean function $\mu(t)$ and covariance function $\gamma(t, t')$

$$y(t) \sim \mathcal{GP}(\mu(t), \gamma(t, t')) \tag{6}$$

To define a GP, the form for the mean function and covariance function needed to be chosen but there is no prior knowledge about the mean function in most applications. Since GP is a linear combination of random variables with Normal Distribution, this is commonly assumed the mean function $\mu(t)$ to be zero. However, we are modeling the variability of human demonstration, which should

be explicitly different than zero. The covariance function $\gamma(t, t')$ generate a non-negative definitive covariance matrix M, which can be formulated by a great set of possible functions with two arguments. Here, we take two frequently used covariance functions into consideration [44], including the squared exponential kernel γ_{se} and the Matern kernel γ_{ma}.

$$\begin{aligned} \gamma_{se}(t, t') &= \sigma^2 e^{-\frac{1}{2\ell^2} d(t,t')^2} \\ \gamma_{ma}(t, t') &= \sigma^2 \frac{1}{\Gamma(\nu) 2^{\nu-1}} (\frac{\sqrt{2\nu}}{\ell} d(t, t'))^{\nu} \mathcal{K}_{\nu}(\frac{\sqrt{2\nu}}{\ell} d(t, t')) \end{aligned} \tag{7}$$

where ℓ is the lengthscale parameter of the kernel for determining the trajectory smoothness (the larger, the smoother), $d(\cdot, \cdot)$ is the Euclidean distance, and the output variance σ^2 is a scale factor that determines the average magnitude away from the mean. Meanwhile, σ and ℓ are hyper-parameters, which can significantly influence the performance of the GP. For the Matern kernel, ν controls the smoothness of the resulting function, $\mathcal{K}_{\nu}(\cdot)$ is a modified Bessel function, and $\Gamma(\cdot)$ is the gamma function. Specifically, when $\nu = 1/2$, $\gamma_{ma}(t, t') = \gamma_{se}(t, t')$.

The output of the Gaussian process model is a normal distribution, expressed in terms of mean and variance. The mean value represents the most likely output and the variance can be interpreted as the measure of its confidence. In contrast with the aforementioned methods in learning from demonstrations, both the variability information of the demonstrations and the presence of uncertainty in reproduction are encoded in a single GP.

3.4. Mean-prior Gaussian Processes Regression

In this section, we define the prior mean of the GP as equal to the mean provided by GMR, and formed as a Mean-prior Gaussian Processes Regression (MP-GPR), which can effectively learn the variability of human demonstrations for adapting to the uncertainty of reproduced trajectory. As formulated in Equation 6, the prior mean of MP-GPR can be obtained by Equation 5 and a kernel in the form of a sum of K separable kernels associated with the K

components of the considered GMM.

$$\begin{aligned} \mu(t) &= p(\mathbf{y}|t) \\ \gamma(t,t') &= \sum_{k=1}^{K} \beta_k(t)\beta_k(t')\hat{\mathbf{\Sigma}}_k^{yy}\gamma_k(t,t') \end{aligned} \tag{8}$$

where the prior mean of Equation 5 allows the MP-GPR to generate trajectory without considering the training trajectories. Moreover, the MP-GPR with non-stationary covariance functions due to its spatially-varying coregionalization matrices $\beta_k(t)\beta_k(t')\hat{\mathbf{\Sigma}}_k^{yy}$, which allows the model to adapt to functions whose smoothness varies with the inputs [45, 46]. The covariance matrices $\hat{\mathbf{\Sigma}}_k^{yy}$ allows the dependencies between the output data to be described for each GMM component and both of the coregionalization matrices and the number of separable kernels are determined by GMR. Therefore, the hyperparameters of the kernels γ_k can be estimated by maximizing the likelihood of the GP. Moreover, the variance parameters σ of the kernels γ_k are commonly fixed to 1.

As formulated in Equation 8, MP-GPR belongs to the generative models such that the reproduced trajectories can be easily generated through sampling and conditioning. This allows to ignore the human demonstrations in the generation of corrective trajectories and to consider only new observations constraints as observed in reproduction, which can be performed well in learning robot anomaly recovery when given safe via-points. Moreover, the tracking precision of new observations can be formulated as a function of the retrieved covariance of MP-GPR, which allows us to demand the robot to track the new observations while lowering the required tracking precision in regions of high variability.

Particularly, in case of learning robot recovery skills, we aim to adapt generated trajectories towards new time-driven via-points $(\mathbf{t}, \mathbf{v}) = \{(t_1, v_1), ..., (t_n, v_n)\}$, where $v_i \in \mathbb{R}^D$ and those new observations are used to infer the posterior distribution of MP-GPR as formulated in Equation 8. Moreover, MP-GPR is that each kernel γ_k can be chosen individually and their parameters are determined separately. With this property, we can modify the local behavior of a robot in

different phases. However, the standard GMM/GMR scheme can't generate a conditioning trajectory towards new observations because the covariance terms between to datapoints are equal to zero, and the predicted mean of standard Gaussian processes regression strongly depend on the kernel parameters such that lack of extrapolation capability.

4. Verification and Performance Comparisons

4.1. Baselines and Common Parameters Settings

In this paper, we compare the performance between MP-GPR with the other probabilistic-based methods that widely used for robot skill learning from human demonstrations, i.e. ProMP (few of the utility functions are used from *promplib*[1]) and KMP (original code implementation[2] in *Matlab* provide by *Yanlong Huang*). The main advantage of those probabilistic methods is that they not only encode the variability by means of a covariance matrix from multiple demonstrations, but also allow for trajectory adaptations with via-points. Note that MP-GPR and KMP can be implemented with the requirements of high-dimensional inputs and outputs because their kernel treatment. However, there are several coupled parameters that should be carefully given in training the ProMP and KMP models.

In our following experiments, the ProMP is empirically formulated by 35 Gaussian basis functions with $\sigma = 0.05$ and the Gaussian noise distribution with $\epsilon \sim \mathcal{N}(0, 0.01)$, in which the detail formulations as described in [14]. KMP is evaluated as well, the time-driven demonstrations are first encoded by a 4 components GMM with input t and output x, y, z, and then a probabilistic reference trajectory is retrieved through GMR. Note that KMP will adapt the distribution according to the covariance given by GMR when concerning the via-points. Subsequently, a Gaussian kernel $\gamma(t, t') = \sigma^2 \exp\left(-\frac{||t-t'||^2}{\ell}\right)$ with variance $\sigma = 0.05$, lengthscale $\ell = 0.1$, and the prior regularization term

[1] https://github.com/baxter-flowers/promplib

[2] https://github.com/yanlongtu/robInfLib-matlab

$\lambda_{mean} = 0.1, \lambda_{cov} = 0.1$ for the mean and covariance that is multiplied by the covariance, respectively, the detail formulations as described in [21]. On the contrary, the MP-GPR prior and posterior models based on the mean of GMR can be conveniently obtained by determining the two kinds of hyperparameters, including the number of GMM is 4 as well as the variance $\sigma = 0.05$ and lengthscale $\ell = 0.1$ of kernels. The detailed implementations are described in the following experimental sections.

4.2. Experiment A: quantitative evaluations on an open-source benchmarking dataset

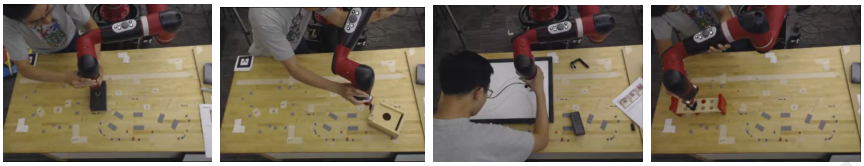
To have a fair comparison between our method and the baselines, we firstly consider an open-source bench-marking dataset for robot skill learning from demonstrations. The dataset was released in literature [47] by *M. Asif Rana*, which including four tasks by human kinesthetic teaching, i.e. reaching, pushing, writing, and pressing. Figure 2 shows the experimental setup and five demonstrated trajectories of each task.

In this paper, two quantitative metrics are considered to evaluate the considered methods. First, a point-wise root mean squared error (*RMSE*) is introduced, which measures the deviation of reproduced motions from demonstrated trajectories. This widely-used metric called path accuracy that has been promising described in two benchmarking studies [47, 48] of robot skill learning, which formulated by

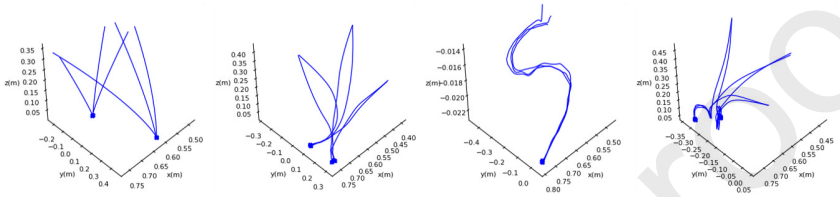
$$RMSE = \sqrt{\frac{1}{M}\sum_{t=1}^{N} ||Y_{demo}(t) - Y_{repo}(t)||} \tag{9}$$

where $||\cdot||$ denotes the L^2-norm, N is the number of datapoints, and M being the number of non-anomalous demonstrations. Y_{demo} and Y_{repo} represent the trajectories of demonstration and reproduction, respectively.

Except for the *RMSE* metric, the variability $\mathcal{V}_{repo}$ of reproduced trajectory is also proposed for quantitative evaluation in this paper, which has been treated as the confidence or uncertainty in performing an action [49, 50]. According



(a) reaching (b) pushing (c) writing (d) pressing



(e) reaching (f) pushing (g) writing (h) pressing

Figure 2: Illustrations of the experimental setups of reaching, pushing, writing, and pressing. (Top) A non-technical end-user to demonstrate those tasks by kinesthetic teaching. (Bottom) Five trajectories are demonstrated for each task, respectively. Further information can be found in the detailed descriptions[4]

to the definitions of MP-GPR, ProMP, and KMP methods, we assume the variability is related to the variance σ_t^2 of reproduced trajectory at each time step. Therefore, the variability is computed as the average magnitude over each Cartesian dimension of the reproduced trajectory by

$$\mathcal{V}_{repo} = \frac{1}{N}\sum_{t=1}^{N}\sqrt{\sigma_{x,t}^2 + \sigma_{y,t}^2 + \sigma_{z,t}^2} \tag{10}$$

where the subscripts x, y, z is used to represent the Cartesian dimension independently.

Due to the demonstrated trajectories in the raw dataset is redundant for our applications, we only select the first five demonstrations of a participant ($M = 5$ in Equation 9), as shown in the bottom row of Figure 2. And then, we evaluate the performance of encoding the variability from multiple demonstrations based on the parameters settings in Section 4.1 for reaching, pushing, writing, and pressing task, respectively, the derived results as shown in Figure 3. Table 1 presents the results of *RMSE* and $\mathcal{V}_{repo}$ using ProMP, KMP, and MP-GPR

Table 1: The quantitative comparisons of MP-GPR, ProMP, and KMP, including *RMSE* denotes the deviation of reproduced trajectory from demonstrations, and $\mathcal{V}_{repo}$ represents the confidence of reproduced trajectory.

Task	**RMSE***			**Variability***		
	ProMP	KMP	MP-GPR	ProMP	KMP	MP-GPR
reaching	1.717	1.735	1.723	0.365	1.070	**0.173**
pushing	0.986	0.992	0.995	0.193	0.645	**0.099**
writing	0.195	0.216	0.214	0.041	0.196	**0.020**
pressing	0.583	0.587	0.591	0.110	0.409	**0.059**
average	0.871	0.881	0.881	0.177	0.580	**0.088**

* The lower the value, the better the reproduced performance.

across four tasks accordingly, where the *RMSE* values are not too different but the variability doesn't. The reasons for these results including 1) the ProMP with the smallest *RMSE* by introducing weight vector ω to compactly represent the demonstrated trajectory with 35 basis functions, which outperforms the others. Whereas the KMP and MP-GPR are kernel-based methods and benefit from GMM/GMR such that they get the similar *RMSE* values. 2) Gaussian processes have been promisingly used to extrapolate with an indication of uncertainty, therefore the MP-GPR can simultaneously encapsulate the variability and uncertainty of multiple demonstrations by using the prior information of GMM/GMR and kernel functions, resulting in the smallest $\mathcal{V}_{repo}$ value of each task and significantly outperform the others. The average variability of MP-GPR across four tasks concerning the baselines is 50.28% better than ProMP[5], and 84.83% better than KMP.

4.3. Experiment B: evaluations on environmental changes

After analyzing the path accuracy and variability of the reproduced trajectory, we evaluate the trajectory adaptation of MP-GPR in an autonomous

[5] $(\mathrm{V}^p_{repo} - \mathcal{V}^m_{repo})/\mathcal{V}^p_{repo}$

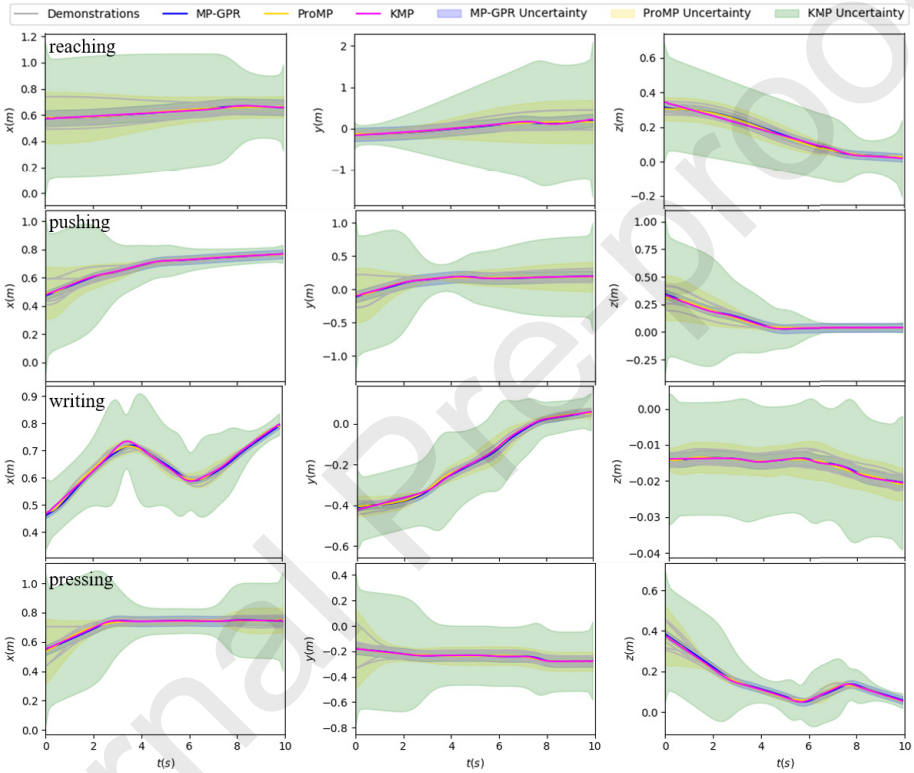
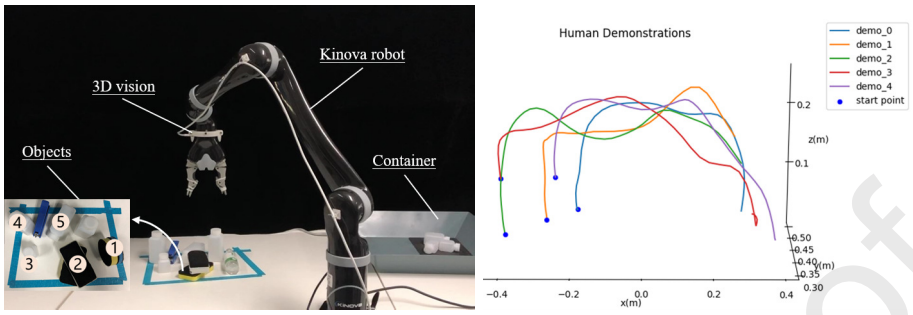


Figure 3: Illustrates the performance of encoding the variability from multiple demonstrations using MP-GPR, ProMP, and KMP, respectively, where the solid lines represent the reproduced trajectory and demonstrated trajectories and the shaded areas denote the variance from the mean.



(a) Experimental setups (b) Five demonstrations

Figure 4: Experimental setup for robot bin-picking the demonstrated trajectories by kinesthetically guiding the robot to grasp the objects and then place them into the container in one-at-a-time fashion, where the blue dot represents the grasping position of each object.

bin-picking task achieved by the 6-DoF Kinova robot and mounted with an Intel RealSeanse 3D camera (for object recognition and picking pose estimation), which demand the robot collects various objects on the desktop into the container one by one, as shown in Figure 4(a). Note that the implementation of bin-picking with 3D vision isn't the considering scope of this paper, which can accurately pick-up objects and don't affect the verification of the proposed method in this paper. In practical applications, the environmental changes should be seriously considered when you want to quickly configure a robot bin-picking systems and the robot is expected to grasp and place in a way that is preconceived by an end-user.

4.3.1. Human Demonstrations

During data acquisition, we collect the demonstrations by human kinesthetically guiding the robot to grasp the objects and then place them into the container in a one-at-a-time fashion without obstacle. Meanwhile, we use the rosbag [6] command to record all the published topics of the Kinova robot synchronously when demonstrating. Particularly, the training dataset consists of 5 human demonstrations by picking the 5 objects marked in the left bottom of

[6] http://wiki.ros.org/rosbag

Figure 4(a) and each demonstration is a three-dimensional time-driven trajectory that consists of Cartesian position x, y, z of robot end-effector, as shown in Figure 4(b). The recorded demonstrations with different durations due to the uncertainty of end-user and object's picking pose, we first pre-treated all the demonstrations using DTW [7] for time alignment within 10 seconds, and then a GMM/GMR model is first trained.

4.3.2. Implement Processes and Results

To effectively evaluate the trajectory adaptation, we first train a GMM/GMR for each dimension based on those five time-driven trajectories according to Equations 1 and 5, with the same parameters settings in Section 4.1, respectively. The resulting means and variances of x, y, z are shown in Figure 5, where a probabilistic prior trajectory $\{\hat{\mathbf{Y}}_i\}_{i=1}^{N}$ in red can be retrieved via GMR and each point $\hat{y}_i$ is associated with t_i is described by a conditional probability distribution with mean $\hat{\mu}_i$ and covariance $\hat{\Sigma}_n$, i.e. $\hat{y}_i|t_i \sim \mathcal{N}(\hat{\mu}_i, \hat{\Sigma}_n)$. As formulated in Equation 5, the prior trajectory encapsulates the variability in the time-driven demonstrations. Subsequently, we take advantage of the mean (prior trajectory) of GMR to consider the uncertainty (via-points) during reproduction using GPR. This permits to ignore the training data in the generation of new trajectories and to consider only via-points constraints as observed data.

The MP-GPR prior (without new observations) is trained based on Matern kernel defined in Equation 7 with smoothness parameter $\nu = 2$. As shown in Figure 5, MP-GPR is that the variability of demonstrations is included in the prior mean derived by GMR and the covariance after determining the hyper-parameters, therefore the variability obtained by MP-GPR is similar to GMR but not equal and the training demonstrations can be ignored during trajectory reproduction. Additionally, although the GMM/GMR computes full covariance matrices encoding the correlation between output variables, it does not mea-

[7] https://github.com/DynamicTimeWarping/dtw-python

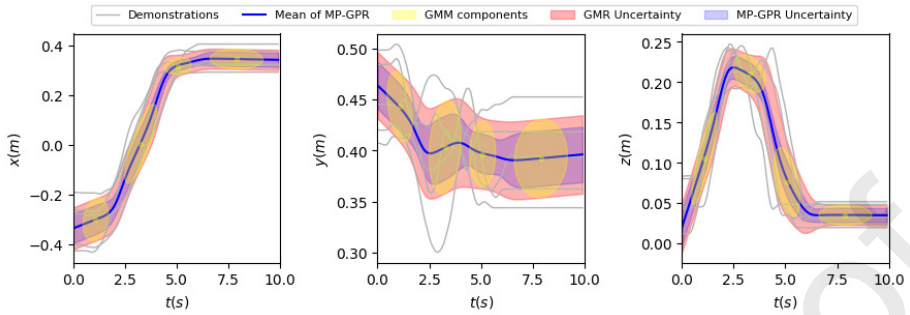


Figure 5: Illustration of the trained MP-GPR through GMM/GMR prior model based on those five demonstrated time-driven trajectories without additional via-points. where, 1) the gray lines denote the human demonstrations; 2) the blue solid curve is the mean that exactly equal to the prior trajectory derived by GMR; 3) the shaded area display the uncertainty of the GMR and MP-GPR distribution with ± 1 standard deviation of the mean trajectory; 4) the light yellow blocks depict the estimated GMMs with the ellipses representing Gaussian components. Note that the variability obtained by MP-GPR is similar to GMR, but not equal and we can generate new trajectories from MP-GP through sampling.

sure the uncertainty in reproduction, and the covariance of the closest Gaussian component is selected by default when a query point is far from the model such that the generated trajectory with new observations is discontinuous. Thus, the GMM/GMR can't be modulated to the new observations.

And then, an obstacle with length (0.35m), width (0.17m), and height (0.27m), added in between the picking position and the container, as shown in Figure 6. Currently, to avoid the obstacle, we need to program the robot transporting trajectories by adding via-points manually and debugging repeatedly, but the robot is hard to maintain the general trajectory shape (mean and variability) and lack of the exploration capability for rapidly adapting to experimental changes. With the proposed MP-GPR, we only need to ask an end-user to kinesthetically guiding the robot to grasp an object and then place it into the container while avoiding the obstacle in an intentional manner, as shown in 6. This demonstration was also recorded and the safe via-points are straightforwardly determined by the end-user.

Due to the theoretical formulations of ProMP and KMP are designed with

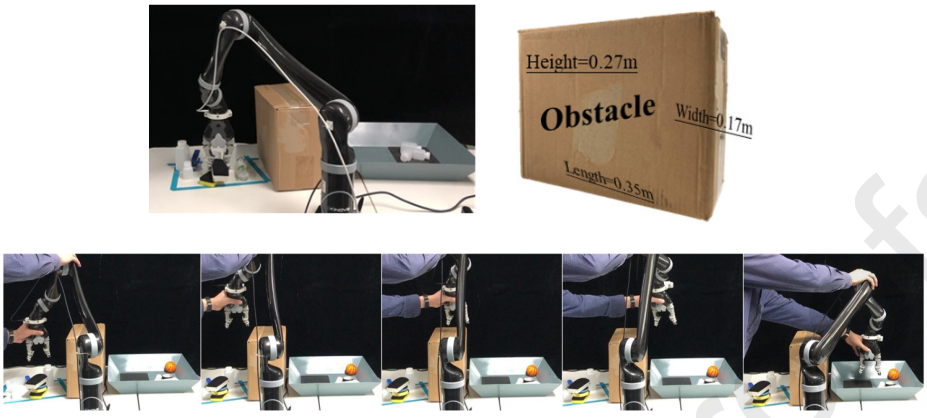


Figure 6: Top: unexpected obstacle and (Bottom) human demonstration: The experiment is initially configured for robot bin-picking using 3D vision, in which human kinesthetically demonstrated the robot how to pick and place. There would have been a collision if an unexpected obstacle was placed in the middle of picking and placing because lots of perturbations were introduced without human intentions. Hence, an end-user should be asked for providing a corrective demonstration above the obstacle, where all the observations are safe and can be added to the prior movement models for trajectory modulation.

the capability for passing through via-points. For a fair comparison, the training data also consist of 5 demonstrations that have been used for MP-GPR evaluation. Based on the aforementioned parameters setting, Figure 7 illustrates the different trajectory modulation using MP-GPR, ProMP, and KMP under the same via-points configuration, and the corresponding snapshots of real-robot applications are shown in Figure 8.

Results indicate that both the MP-GPR and KMP can be effectively modulated to the safe via-points, but the ProMP doesn't, therefore we only further compared the computational efficiency between our method and KMP. KMP modulates the whole trajectory distribution according to the via-points' covariance because it uses the covariance matrix of the observation noise to represent the variability and needs to find the nearest datapoint by a manually predefined threshold. Whereas, MP-GPR can turn the hyper-parameters of the specific kernel at robot different phase for precisely avoiding the anomalies and keep the normal execution for those normal phases according to the Equation 8.

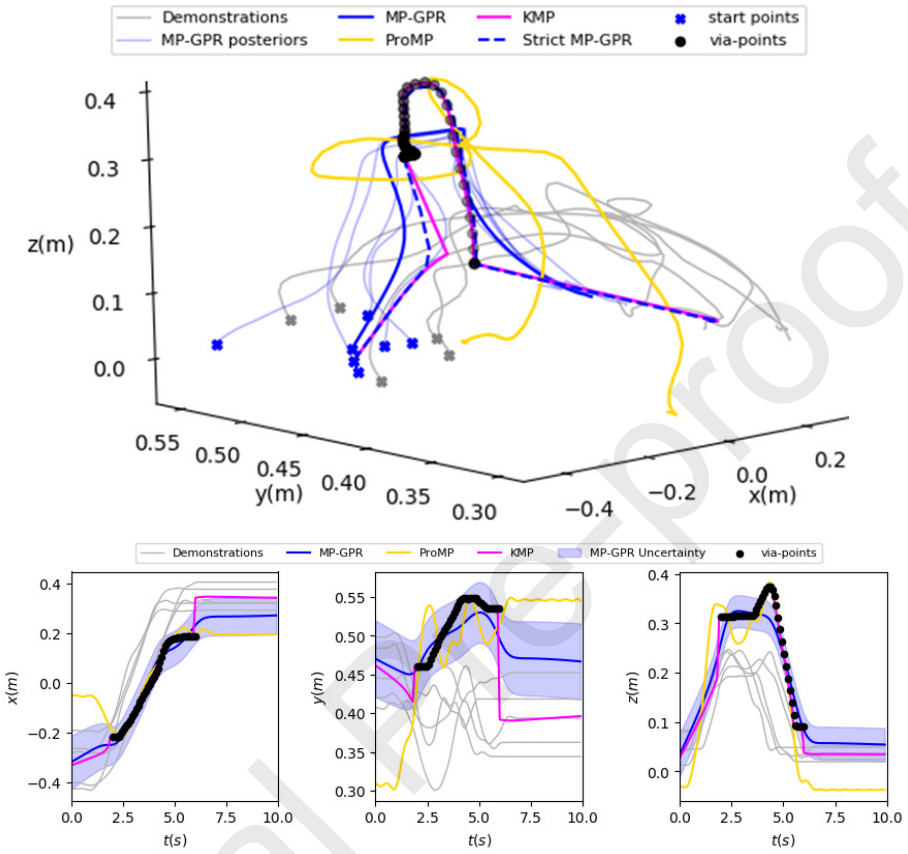


Figure 7: Illustrates the comparison between MP-GPR model with ProMP, KMP with via-points, where the number of GMM components for training the prior mean of MP-GPR and the reference trajectory of KMP are 4. Specifically, 1) the demonstrated trajectories without obstacle are shown in gray, in which the time t as input and the Cartesian position x, y, z of robot end-effector as outputs; 2) The mean or reference trajectory generated by adding the constraint of safe via-points (in black) by MP-GPR, ProMP and KMP are represented in blue, gold and magenta, respectively. As shown, the proposed MP-GPR can effectively learn the variability of the demonstrations and the uncertainty of reproductions simultaneously, which the modulated trajectory can avoid obstacles while maintaining the underlying shape of demonstrations. However, the modulated trajectory of ProMP can't encode the variability and uncertainty simultaneously and KMP strictly considers the via-points such that ignore the variability of demonstrations.

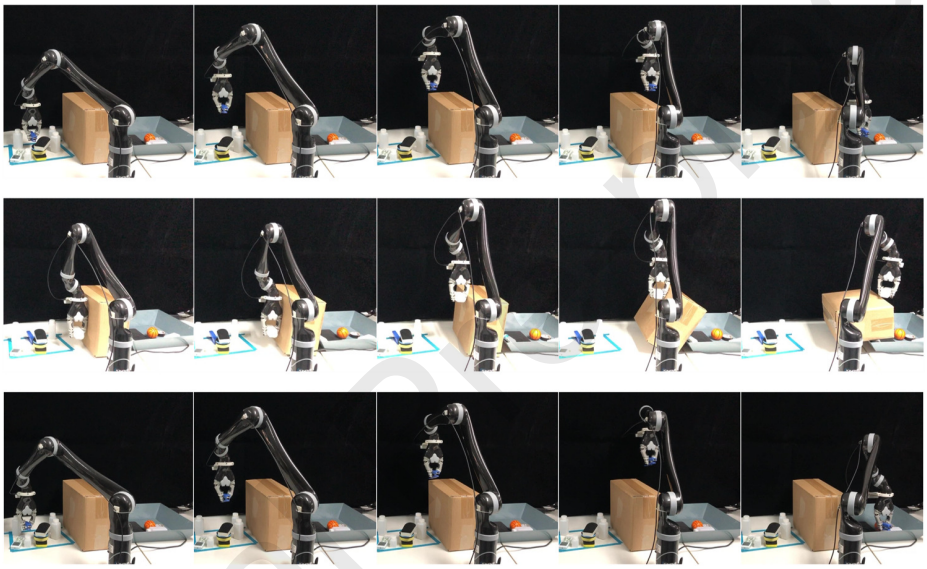


Figure 8: Illustrates the effective performance of MP-GPR (top) and KMP (bottom) for learning robot anomaly recovery skills based on five time-driven trajectories, but the modulated trajectory of ProMP (middle) can't avoid the obstacle as shown in Figure 7. Results indicate that the proposed MP-GPR method allows the robot to avoid the obstacle by modulating trajectory to the safe via-points and easily adapt to environmental changes, e.g. different target position and the shape of obstacles.

This characteristic is different from the KMP implementation for strictly pass through all the given via-points and ignore the variability of demonstrations. As a consequence, the computational efficiency of KMP is lower than our method given the same via-points. In our case, assuming the derived trajectory of 100 datapoints and 42 via-points, the average time cost of adding a via-point of KMP is 0.373s, but MP-GPR only takes 0.097s. The relative ratio is 73.99%. This property specifically important in robot anomaly recovery skills learning.

5. Conclusion and Future work

To allow robots rapidly adapt to environmental changes, we present a framework for learning robot recovery skills from human demonstrations by passing through given safe via-points. The proposed MP-GPR method takes advantage both of GMM/GMR and GPR, which results in effectively learning the variability from human demonstrations and the uncertainty in reproduction simultaneously. It adapts the prior trajectory generated from demonstrations to avoid obstacles in the desired manner while maintaining the underlying shape and variability. An open-source benchmarking dataset and a robot bin-picking experiment are presented to evaluate the performances and compare with commonly used probabilistic methods. Results indicate that MP-GPR can be conditioned uniquely on via-points by ignoring the variability derived in demonstrations by defining a prior mean as GMM/GMR. Besides the applications of learning recovery skills, the MP-GPR would be widely used in other applications such as imitation learning and human-robot collaboration scenarios. The online implementation of robot learning recovery skills using MP-GPR based on visual perception will be further investigated in the future.

Acknowledgment

This work is supported by Guangdong Province Key Areas R&D Program (Grant No. 2019B090919002), Basic and Applied Basic Research Project of

Guangzhou (Grant No. 202002030237), GDAS' Project of Thousand doctors(post-doctors) Introduction (2020GDASYL-20200103128), Foshan Key Technology Research Project(Grant No. 1920001001148), Guangdong Province International Cooperation Project of Science and Technology (Grant No. 2019A050510040), National Science Foundation of China (Grant No. 61950410758).

References

[1] A. Cherubini, R. Passama, A. Crosnier, A. Lasnier, P. Fraisse, Collaborative manufacturing with physical human–robot interaction, Robotics and Computer-Integrated Manufacturing 40 (2016) 1–13.

[2] O. Ogorodnikova, Methodology of safety for a human robot interaction designing stage, in: Conference on Human System Interactions, 2008.

[3] M. Vasic, A. Billard, Safety issues in human-robot interactions, in: 2013 IEEE International Conference on Robotics and Automation, 2013.

[4] H. Wu, Y. Guan, J. Rojas, Analysis of multimodal bayesian nonparametric autoregressive hidden markov models for process monitoring in robotic contact tasks, International Journal of Advanced Robotic Systems 16 (2) (2019) 1729881419834840.

[5] D. Park, H. Kim, C. C. Kemp, Multimodal anomaly detection for assistive robots, Autonomous Robots.

[6] H. Wu, H. Lin, Y. Guan, K. Harada, J. Rojas, Robot introspection with bayesian nonparametric vector autoregressive hidden markov models, in: 2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids), IEEE, 2017, pp. 882–888.

[7] C. Zeng, C. Yang, Z. Chen, S.-L. Dai, Robot learning human stiffness regulation for hybrid manufacture, Assembly Automation.

[8] H. Ravichandar, A. S. Polydoros, S. Chernova, A. Billard, Recent advances in robot learning from demonstration, Annual Review of Control, Robotics, and Autonomous Systems 3.

[9] C. Yang, C. Zeng, Y. Cong, N. Wang, M. Wang, A learning framework of adaptive manipulative skills from human to robot, IEEE Transactions on Industrial Informatics 15 (2) (2018) 1153–1161.

[10] B. D. Argall, S. Chernova, M. Veloso, B. Browning, A survey of robot learning from demonstration, Robotics Autonomous Systems 57 (5) (2009) 469–483.

[11] P. Pastor, H. Hoffmann, T. Asfour, S. Schaal, Learning and generalization of motor skills by learning from demonstration, in: Robotics and Automation, 2009. ICRA'09. IEEE International Conference on, IEEE, 2009, pp. 763–768.

[12] S. Calinon, A. Billard, Statistical learning by imitation of competing constraints in joint space and task space, Advanced Robotics 23 (15) (2009) 2059–2076.

[13] C. Ye, J. Yang, H. Ding, Bagging for gaussian mixture regression in robot learning from demonstration, Journal of Intelligent Manufacturing (2020) 1–13.

[14] A. Paraschos, C. Daniel, J. R. Peters, G. Neumann, Probabilistic movement primitives, in: Advances in neural information processing systems, 2013, pp. 2616–2624.

[15] A. Paraschos, E. Rueckert, J. Peters, G. Neumann, Model-free probabilistic movement primitives for physical interaction, in: Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems (IROS), 2015.

[16] S. Calinon, F. Guenter, A. Billard, On learning, representing, and generalizing a task in a humanoid robot, IEEE Transactions on Systems Man Cybernetics Part B 37 (2) (2007) 286–298.

[17] C. Yang, C. Chen, N. Wang, Z. Ju, J. Fu, M. Wang, Biologically inspired motion modeling and neural control for robot learning from demonstrations, IEEE Transactions on Cognitive and Developmental Systems 11 (2) (2018) 281–291.

[18] C. Zeng, C. Yang, H. Cheng, Y. Li, S.-L. Dai, Simultaneously encoding movement and semg-based stiffness for robotic skill learning, IEEE Transactions on Industrial Informatics.

[19] S. Calinon, T. Alizadeh, D. G. Caldwell, On improving the extrapolation capability of task-parameterized movement models, in: IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS), 2013.

[20] J. Silvério, Y. Huang, F. J. Abu-Dakka, L. Rozo, D. G. Caldwell, Uncertainty-aware imitation learning using kernelized movement primitives, in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.

[21] Y. Huang, L. Rozo, J. Silvério, D. G. Caldwell, Kernelized movement primitives, International Journal of Robotics Research 38 (7) (2019) 833–852.

[22] M. Schneider, W. Ertel, Robot learning by demonstration with local gaussian process regression, in: IEEE/RSJ International Conference on Intelligent Robots Systems, 2010.

[23] N. Jaquier, D. Ginsbourger, S. Calinon, Learning from demonstration with model-based gaussian process, in: Conference on Robot Learning, PMLR, 2020, pp. 247–257.

[24] S. M. Khansari-Zadeh, A. Billard, Learning stable nonlinear dynamical systems with gaussian mixture models, IEEE Transactions on Robotics 27 (5) (2011) 943–957.

[25] J. S. Laursen, U. P. Schultz, L. P. Ellekilde, Automatic error recovery in robot assembly operations using reverse execution, in: IEEE/RSJ International Conference on Intelligent Robots Systems, 2015.

[26] M. Karlsson, A. Robertsson, R. Johansson, Autonomous interpretation of demonstrations for modification of dynamical movement primitives, in: 2017 IEEE International Conference on Robotics and Automation (ICRA), 2017.

[27] A. Muxfeldt, J. Steil, Fusion of human demonstrations for automatic recovery during industrial assembly, in: IEEE CASE, 2018.

[28] A. Rodriguez, M. T. Mason, S. S. Srinivasa, M. Bernstein, A. Zirbel, Abort and retry in grasping, in: 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2011, San Francisco, CA, USA, September 25-30, 2011, 2011.

[29] M. Ren, J. S. Baras, Y. Yang, C. Fermuller, Co-active learning to adapt humanoid movement for manipulation, in: IEEE-RAS International Conference on Humanoid Robots, 2017.

[30] K. Chatzilygeroudis, V. Vassiliades, J. B. Mouret, Reset-free trial-and-error learning for robot damage recovery, Robotics Autonomous Systems (2017) S0921889017302440.

[31] C. Yang, K. Yuan, Q. Zhu, W. Yu, Z. Li, Multi-expert learning of adaptive legged locomotion.

[32] C. Yang, G. Peng, Y. Li, R. Cui, L. Cheng, Z. Li, Neural networks enhanced adaptive admittance control of optimized robot–environment interaction, IEEE transactions on cybernetics 49 (7) (2018) 2568–2579.

[33] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, S. Schaal, Dynamical movement primitives: learning attractor models for motor behaviors, Neural computation 25 (2) (2013) 328–373.

[34] I. Iturrate, C. Sloth, A. Kramberger, H. G. Petersen, T. R. Savarimuthu, Towards reversible dynamic movement primitives, in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019.

[35] C. Yang, C. Chen, W. He, R. Cui, Z. Li, Robot learning system based on adaptive neural control and dynamic movement primitives, IEEE transactions on neural networks and learning systems 30 (3) (2018) 777–787.

[36] C. Yang, C. Zeng, C. Fang, W. He, Z. Li, A dmps-based framework for robot learning and generalization of humanlike variable impedance skills, IEEE/ASME Transactions on Mechatronics 23 (3) (2018) 1193–1203.

[37] S. Choi, K. Lee, S. Lim, S. Oh, Uncertainty-aware learning from demonstration using mixture density networks with sampling-free variance modeling.

[38] K. Kersting, C. Plagemann, P. Pfaff, W. Burgard, Most likely heteroscedastic gaussian process regression, in: Machine Learning, Proceedings of the Twenty-Fourth International Conference (ICML 2007), Corvallis, Oregon, USA, June 20-24, 2007, 2007.

[39] J. Silve´rio, Y. Huang, L. Rozo, S. Calinon, D. G. Caldwell, Probabilistic learning of torque controllers from kinematic and force constraints, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017.

[40] M. Arduengo, A. Colomé, J. Borràs, L. Sentis, C. Torras, Task-adaptive robot learning from demonstration with gaussian process models under replication, IEEE Robotics and Automation Letters 6 (2) (2021) 966–973.

[41] Calinon, Sylvain, A tutorial on task-parameterized movement learning and retrieval, Intelligent Service Robotics 9 (1) (2016) 1–29.

[42] M. I. M. Yusoff, I. Mohamed, M. R. A. Bakar, Improved expectation maximization algorithm for gaussian mixed model using the kernel method, Mathematical Problems in Engineering,2013,(2013-8-21) 2013 (pt.9) (2013) 377–384.

[43] E. S. A, M. S. B, A. K. C, A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions, Journal of Mathematical Psychology 85 (2018) 1–16.

[44] M. A. Alvarez, L. Rosasco, N. D. Lawrence, Kernels for vector-valued functions: a review, Foundations and Trends in Machine Learning 4 (3) (2011) 195–266.

[45] C. Paciorek, M. Schervish, Nonstationary covariance functions for gaussian process regression, Advances in neural information processing systems 16 (2003) 273–280.

[46] A. Gelfand, A. Schmidt, S. Banerjee, C. Sirmans, Nonstationary multivariate process modeling through spatially varying coregionalization, Test 13 (2) (2004) 263–312.

[47] M. A. Rana, D. Chen, J. Williams, V. Chu, S. R. Ahmadzadeh, S. Chernova, Benchmark for skill learning from demonstration: Impact of user experience, task complexity, and start configuration on performance, in: 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2020, pp. 7561–7567.

[48] A. Lemme, Y. Meirovitch, M. Khansari-Zadeh, T. Flash, A. Billard, J. J. Steil, Open-source benchmarking for learned reaching motion generation in robotics, Paladyn, Journal of Behavioral Robotics 6 (1).

[49] S. Chernova, M. Veloso, Interactive policy learning through confidence-based autonomy, Journal of Artificial Intelligence Research 34 (2009) 1–25.

[50] S. Thakur, H. van Hoof, J. C. G. Higuera, D. Precup, D. Meger, Uncertainty aware learning from demonstrations in multiple contexts using bayesian neural networks, in: 2019 International Conference on Robotics and Automation (ICRA), IEEE, 2019, pp. 768–774.