Imitation Learning for Coordinated Human-robot Collaboration based on Hidden State-space Models

Likun Wang[a], Guoyan Wang[b,*], Shuya Jia[c], Alison Turner[a], Svetan Ratchev[a]

[a]Centre for Aerospace Manufacturing, University of Nottingham, UK
[b]School of Computer Science and System, Bauman Moscow State Technical University, Russia
[c]Renishaw plc, Gloucestershire, UK.

Abstract

This paper proposes a novel coordinated human-robot collaboration framework based on the hidden state-space model, which probabilistically clones the human behaviour and presents dynamic features in a nonparametric form. Derived from the filter prediction techniques and the theory of exact moment matching, this framework could provide an analytical approximation of the posterior distribution, and hence infer the hidden state variables of the collaborative robot given the external observation and its uncertainties. Not akin to the other cutting-edge movement-primitive based algorithms or coordinated human-robot collaboration methods, our collaboration framework not only preserves the adaptation functionalities of imitation learning but also propagates state variables and their uncertainties during real-time coordinated implementation. By leveraging on the binary Gaussian process classification, additional functionality, such as multiple task recognition is proposed to enhance the generalisation capability of our framework. The application feasibility is verified from both theoretical comparison simulation and real-world experiments.

Keywords: Imitation learning, Coordinated human-robot collaboration, Hidden state-space model, Gaussian process, Binary classification

1. Introduction

In modern industrial manufacturing, robots play a significant role in various crucial applications, such as painting components in automotive industry, shot-peen in aerospace manufacturing, and agile packing in food production [1, 2]. These industrial applications are usually implemented in several work cells operated by skilled workers with a given set of instructions. Evidently, the robots as a powerful tool are utilised to fulfill tasks or assist operators to achieve desirable industrial processes, which is usually referred to as human-robot collaboration [3]. However, coordinated physical collaboration between human and robots is still challenging, since robots are required to adapt to environment changing and both motions should be aligned spatially and temporally. Imitation learning is a potential candidate to clone human behaviour and transfer human motion skills to robots. More importantly, the extension of imitation learning model could be used in the applications such as force-driven interaction control [4], manufacturing product handover [5, 6] and obstacle avoidance motion generation [7].

Dynamic Time Warping (DTW) has been applied in human-robot collaboration, for the purpose of temporally aligning human actions and robot predicted trajectories, which are retrieved according to various human demonstrations [8] or directly learned from a probabilistic representation [9]. Nevertheless, as DTW heavily relies on the distance measurements among different full trajectories [10], on-line applications with full or partial observations [11] are not desirable.

Although Hidden Markov Model (HMM) is a typical time series prediction method, it can be used for generating robot trajectories in the human-robot collaboration applications, such as tool delivery [12], low-level primitive control [13] and human action prediction [14]. Additionally, in [15] the temporal issues are solved with transition probabilities during on-line implementation. However, these methods do not model the robot trajectories explicitly. Alternatively, they rely on the mixture of the trajectory demonstrations, which are separated from the methods themselves.

Alternative human-robot collaboration approaches could be referred to as the movement-primitive-based learning algorithms. For instance, Dynamical Movement Primitives (DMP) [16] could provide variation of temporal modulation with an explicit phase indicator, which can be seen as a useful tool to govern trajectories of several agents. Moreover, a new variation of movement primitive approach, so-called Kernerlized Movement Primitives (KMP) [18] which is derived from the nonparametric ridge regression, focuses more on the nonlinear mapping between human motion and robot pose. In [17], both human motion and robot trajectories are correlated together using interaction Probabilistic Movement Primitives (iProMP).

In order to correctly interpret human observation, filtering techniques are applied in [19, 20]. Consequently, the robot trajectories are inferred given the measurement from the inertial measurement unit or force sensor. In addition, the human inten-

---

Table 1: Qualitative comparison.

Approach	Probabilistic	Dynamic system	Uncertainty propagation	Muli-task
DMP [16]	–	√	–	–
iProMP[17]	√	–	–	√
KMP [18]	√	–	–	–
Our approach	√	√	√	√

tion for assembly cell could also be recognised by using Kinect depth camera as detailed in [21]. The extension of their work [22, 23] focuses on the safety issues of the human-robot interaction with wearable sensors and AR techniques. The safety issues in a share workspace are also discussed in [24] by using cyber physical system. A collision-free algorithm for human-robot collision based on context awareness is introduced in [25]. Besides, machine learning algorithms such as Bayesian estimation [26], reinforcement learning [27], deep learning [28] and Gaussian process [29] are applied for human intention as well as interaction impedance in human-robot collaboration. Nevertheless, these nonlinear human intention estimation algorithms do not provide dynamic functionalities.

According to the above analysis, the temporal alignment is one of the main issues of the coordinated human-robot collaboration. Although DTW could provide a straightforward solution, it suffers from computation burden. HMM is usually seen as an alternative. However, HMM only presents discrete state variables and requires additional knowledge to output continuous probabilistic trajectories.

Since the implementation of robot motion is always related to temporal parameters, a dynamic system might be a better option as introduced in DMP. Nevertheless, it must combine with a separate canonical system and does not support probabilistic trajectory prediction. Both recent cutting-edge movement-primitive algorithms, i.e., KMP and iProMP are derived from the basis function regression. The KMP approach seeks to use nonlinear mapping to model human-robot collaboration, which may lose the crucial temporal information. Leveraging on the filtering technique, the iProMP framework is able to retrieve adapted trajectory given the human observations. However, the framework always updates the whole trajectory, even the part that has already been implemented. This could be a disadvantage if the next-step predicted robot state of the new trajectory has significant drift compared with current robot state. Then, the jerk motion of the robot is unavoidable. In addition, the selection of the width of the basis function is also crucial, as it may cause overfitting problems.

Therefore, the challenges of coordinated human-robot collaboration can be summarised into three aspects. Firstly, the human behaviour as well as corresponding robot trajectories should be effectively captured and learned. Secondly, the robot system must present its dynamic functionalities both from spatial and temporal aspects. More specially, human observation should be considered in real-time, along with its uncertainties. Thirdly, the developed framework of the coordinated human-robot collaboration should have the ability to accommodate different collaboration tasks to enhance the generalisation capability.

Despite the large volume of excellent work, there are still some open issues to address as presented in the qualitative comparison in Table 1. Consequently, we propose a novel probabilistic coordinated human-robot collaboration framework that could retrieve novel adapted robot trajectory given human observation and its uncertainties. Therefore, we highlight our contributions as follows:

1. Firstly, we propose a novel imitation learning approach based on differential Gaussian process that could encode the human behaviour probabilistically;

2. Secondly, the dynamic hidden state-space model of human-robot collaboration is introduced, along with the uncertainty propagation based on the exact moment matching. The adapted trajectory for the robot can be obtained from the conditional inference given the Gaussian approximation of two consecutive states;

3. Furthermore, multi-task recognition is discussed derived from the binary Gaussian process classification to further improve the generalisation capability;

4. Finally, we conduct a comparison experiment and two real-world applications to further verify the feasibility of our proposed framework.

The remainder of this paper is organised as follows: after the introduction, the coordinated human-robot collaboration framework for a single task is detailed in Section 2; the multi-task recognition is further studied in Section 3; Finally, the simulation comparison and real-world experiments with an ABB YuMi robot and an Oculus VR are implemented in Section 4 and the conclusion is drawn in the last section.

2. Coordinated Human-robot Collaboration Framework

In this section, the novel proposed collaboration framework is introduced into three parts. First, a novel imitation learning method based on differential Gaussian process is detailed in Subsection 2.1. Then, the hidden state-space collaboration model is addressed in Subsection 2.2. Finally, the conditional inference with exact moment matching is given in Section 2.3.

2.1. Learning from Human Demonstration

The human demonstrations could be defined as the collection of the observation pairs $\{x_i, y_i\}$ and differential observation
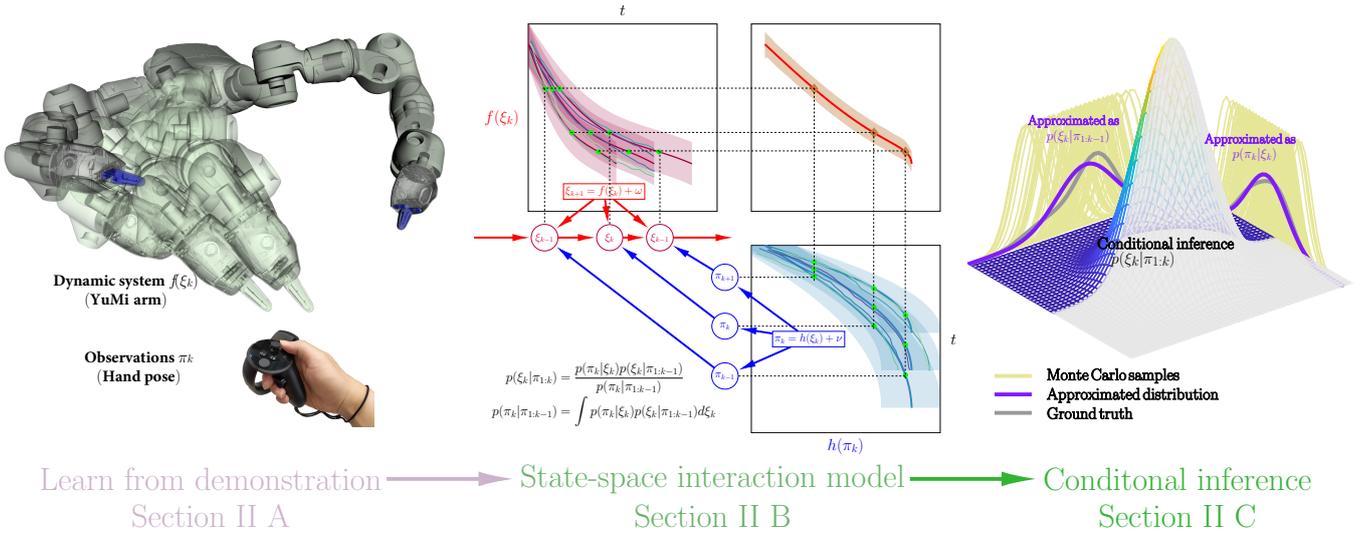
Figure 1: The graphic explanation of the physical human-robot collaboration framework. The collaboration framework can be divided into three parts. Firstly, both human and robot trajectories are learned from the demonstration in Subsection 2.1. Then the robot dynamic system state ξ_{k+1} in the interaction model is only dependant on the previous state ξ_k and the human observation π_k as detailed in Subsection 2.2. Finally, the novel retrieved trajectory is obtained from the conditional inference of the exact moment matching approximation in subsection 2.3.

pairs $\{x_i^d, y_i^d\}$

$$\mathcal{D} = \{\{x_i, y_i\}_{i=1}^M, \{x_i^d, y_i^d\}_{i=1}^N\}, \tag{1}$$

where y_i and x_i are the observation and its timestamp, respectively. Similarly, y_i^d and x_i^d are denoted as the differential observation and its timestamp. M and N are the total number of the observation pairs and differential observation pairs, correspondingly. In this paper, i and j are defined as the index number of the timestamp and observation variables.

Consequently, the human demonstrations are encoded by a nonparametric Gaussian process model and in this paper, we refer to as Gaussian process movement primitives. Consequently, we define a squared exponential covariance function as follows

$$k_x = \sigma_0^2 \exp(\frac{1}{2}\|x_i - x_j\|_\lambda^2) + \sigma_f^2 \delta_{ij}, \tag{2}$$

with σ_0 the vertical scale, λ the horizontal scale, and σ_f the noise variance. For further addressing the differential observations, the squared exponential covariance function should be rewritten to include the differential kernel

$$k_{dx} = \frac{\partial k_x(x_i^d, x_j)}{\partial x_i^d} \tag{3}$$

$$k_{ddx} = \frac{\partial^2 k_x(x_i^d, x_j^d)}{\partial x_i^d \partial x_j^d}. \tag{4}$$

and hence the compact forms [30] could be given as

$$k_{dx} = -k_x(x_i^d, x_j)\lambda(x_i^d - x_j)$$
$$k_{ddx} = k_x(x_i^d, x_j^d)\lambda$$
$$\quad - k_x(x_i^d, x_j^d)\lambda(x_i^d - x_j^d)(x_i^d - x_j^d)^T \lambda.$$

In preparation for retrieving the position and velocity of a novel trajectory, the conditional distribution given an enquiring point x_* is defined as

$$\mu(x_*) = k_*(\boldsymbol{K} + \sigma_f^2 \boldsymbol{I})^{-1} \boldsymbol{y} \tag{5}$$

$$\Sigma(x_*) = k_{**} - k_*^T(\boldsymbol{K} + \sigma_f^2 \boldsymbol{I})^{-1} k_*. \tag{6}$$

Accordingly the term $\boldsymbol{y}$ should include the differential observation defined as $\boldsymbol{y} = [\{y_i\}_{i=1}^M, \{y_i^d\}_{i=1}^N]^T$. Consequently, the covariance k_{**} and cross-covariance k_* are given as $k_{**} = k_x(x_*, x_*)$ and $k_* = [k_x(x_*, x_i)\}_{i=1}^M, \{k_{dx}(x_*, x_i^d)\}_{i=1}^M]^T$, separately. The Gram matrix $\boldsymbol{K}$ should also be rewritten regarding the additional differential observations x_i^d

$$\boldsymbol{K} = \begin{bmatrix} [k_x(x_i, x_j)] & [k_{dx}(x_i^d, x_j)] \\ [k_{dx}(x_i, x_j^d)] & [k_{ddx}(x_i^d, x_j^d)] \end{bmatrix}.$$

Since the above novel covariance functions expressed in Equ. 3 and Equ. 4 are in closed-form, the prediction is still Gaussian.

2.2. State-space Interaction Model

In the following, we assume two retrieved trajectories are derived from two different agents, respectively, such as two YuMi arms or a human operator and a robot manipulator. Consequently, the two trajectories are encoded with Gaussian process movement primitives denoted by $\{\xi_k\}$ and $\{\pi_k\}$, separately.

Without loss of generality, for an autonomous robotic system, the system dynamic equation can be expressed as

$$\xi_k = f(\xi_{k-1}) + \omega, \tag{7}$$

where $\{\xi_k \sim \mathcal{N}(\mu_\xi^k, \Sigma_\xi^k)\}$ is the state variable and ω is the system noise. Given the observation $\{\pi_k\}$, which is simultaneously obtained from another object, the task defined in this subsection

is the inference of the hidden state variable ξ_k according to the measurement system defined as

$$\pi_k = h(\xi_k) + \nu, \qquad (8)$$

with $\pi_k \sim \mathcal{N}(\mu_\pi^k, \Sigma_\pi^k)$ and ν the noise of the measurement system. We would like to point out that this inference is not exactly the same as filtering techniques, which the measurement system observes the same object and usually has direct connection with state variables. In our scenario, however, the observation is obtained from a second agent and the mapping between the two agents is established in Equ .8.

Given the prior information of the previous step estimation $p(\xi_{k-1}|\pi_{k-1})$, the posterior distribution of the state parameter ξ_k could be obtained according to the Bayesian theory

$$p(\xi_k|\pi_{1:k}) = \frac{p(\pi_k|\xi_k)p(\xi_k|\pi_{1:k-1})}{p(\pi_k|\pi_{1:k-1})}, \qquad (9)$$

where the likelihood $p(\pi_k|\xi_k)$ is derived from the measurement system in Equ. 8. The prior knowledge can be obtained from the dynamic system as presented in Equ. 7

$$p(\xi_k|\pi_{1:k-1}) = \int p(\xi_k|\xi_{k-1})p(\xi_{k-1}|\pi_{k-1})d\xi_{k-1}. \qquad (10)$$

However, the marginal likelihood $p(\pi_k|\pi_{1:k-1})$, which has the following expression

$$p(\pi_k|\pi_{1:k-1}) = \int p(\pi_k|\xi_k)p(\xi_k|\pi_{1:k-1})d\xi_k, \qquad (11)$$

is in most interesting cases analytically intractable. This is because the propagation of a Gaussian distribution learned from the human demonstrations is not a Gaussian after going through a nonlinear function.

Although typical solution such as Monte Carlo sampling can provide an exact result, it suffers from computational expense. In this paper, we apply an alternative analytical solution, so-called exact moment matching [31], which matches the mean and the covariance of the true distribution and approximates this distribution as a normal distribution. Consequently, the propagated mean of function $f(\xi_{k-1})$ with uncertainty input $\xi_{k-1} \sim \mathcal{N}(\mu_\xi^{k-1}, \Sigma_\xi^{k-1})$ can be expressed as

$$m_f(\xi_{k-1}) = \int \mu(\xi_{k-1})p(\xi_{k-1})d\xi_{k-1}$$
$$= \int \mu(\xi_k)\mathcal{N}(\xi_k|\mu_\xi^{k-1}, \Sigma_\xi^{k-1})d\xi_{k-1} = \alpha^T q, \qquad (12)$$

where $\alpha = (K + \sigma_f I)^{-1}y$ and q is defined as

$$q = \sigma_f^2|\Sigma_\xi^{k-1}\lambda^{-1} + I|^{-\frac{1}{2}}$$
$$\exp(\frac{1}{2}(\xi_{k-1} - \mu_\xi^{k-1})^T(\Sigma_\xi^{k-1} + \lambda)^{-1}(\xi_{k-1} - \mu_\xi^{k-1})).$$

The propagated variance given the uncertainty input ξ_{k-1} could

be given as

$$\sigma_f(\xi_{k-1}) = \int (\Sigma_\xi^{k-1})^2 p(\xi^{k-1})d\xi^{k-1} - m_f^2(\xi_k)$$
$$+ \int (\mu_\xi^{k-1})^2 p(\xi_{k-1})d\xi_{k-1} \qquad (13)$$
$$= \alpha^T Q\alpha + \sigma_0^2 - tr((K + \sigma_f^2)^{-1}Q) - m_f^2(\xi_k),$$

where Q_{ij} is given by

$$Q_{ij} = \frac{k_x(x_i, \mu_\xi^{k-1})k_x(x_j, \mu_\xi^{k-1})}{|2\Sigma_\xi^{k-1}\lambda^{-1} + I|^{\frac{1}{2}}}$$
$$\exp(z_{ij} - \mu_\xi^{k-1})^T(\Sigma_\xi^{k-1} + \frac{1}{2}\lambda)^{-1}\Sigma_\xi^{k-1}\lambda)^{-1}(z_{ij} - \mu_\xi^{k-1}), \quad (14)$$

with $z_{ij} = (x_i + x_j)/2$. Note that the Equ. 11 can be solved analytically by setting $\xi_{k-1} = \xi_k$ and f as h. Therefore, the analytical expressions of $p(\xi_k|\pi_{1:k-1}) \sim \mathcal{N}(m_f(\xi_{k-1}), \sigma_f(\xi_{k-1}))$ and $p(\pi_k|\xi_k) \sim \mathcal{N}(m_h(\xi_k), \sigma_h(\xi_k))$ are obtained.

2.3. Conditional Inference

According to the above analysis, the coordinated human-robot collaboration model including dynamic system and measurement system, is nonparametric and relies on the data itself. Not akin to the inference in [16] or [17], the collaboration model is explicit and also the uncertainties of the data (trajectories and observations) are considered at every step propagation. These issues could be addressed by assuming known or tractable distributions propagated in predict-update-project cycles.

Consequently, the posterior inference of the hidden state variable ξ_k can be deduced by the conditional inference $p(\xi_k|\pi_{1:k})$ given a joint distribution

$$p(\xi_k, \pi_k|\pi_{1:k-1}) = p(\xi_k|\pi_k)p(\xi_k|\pi_{1:k-1})$$
$$\sim \mathcal{N}(\begin{bmatrix} m_f(\xi_{k-1}) \\ m_h(\xi_k) \end{bmatrix}, \begin{bmatrix} \sigma_f(\xi_{k-1}) & \sigma_{fh}(\xi_{k-1}, \xi_k) \\ \sigma_{hf}(\xi_k, \xi_{k-1}) & \sigma_h(\xi_k) \end{bmatrix}). \quad (15)$$

The mean term $m_f(\xi_{k-1})$ and $m_h(\xi_k)$ could be derived from exact matching as detailed in the previous subsection respectively, along with the covariance $\sigma_f(\xi_{k-1})$ and $\sigma_h(\xi_k)$. In terms of exact expression of cross-covariance $\sigma_{fh}(\xi_k, \xi_{k-1})$, we omit it but refer to the excellent work in [31].

Since we have all the expressions of the joint distribution $p(\xi_k, \pi_k|\pi_{1:k-1})$, the posterior distribution of the hidden-state ξ_k can be finally obtained by

$$p(\xi_k|\pi_{1:k}) \sim \mathcal{N}(\xi_k|m_k^p, \sigma_k^p), \qquad (16)$$
$$m_k^p = m_f + \sigma_{fh}\sigma_f^{-1}(\pi_k - m_h), \qquad (17)$$
$$\sigma_k^p = \sigma_f - \sigma_{fh}\sigma_f^{-1}\sigma_{fh}^T. \qquad (18)$$

The graphic explanation of our proposed framework is concluded in Fig. 1. We would like to point out that the procedure in Fig. 1 is an example of one coordinated human-robot collaboration task.
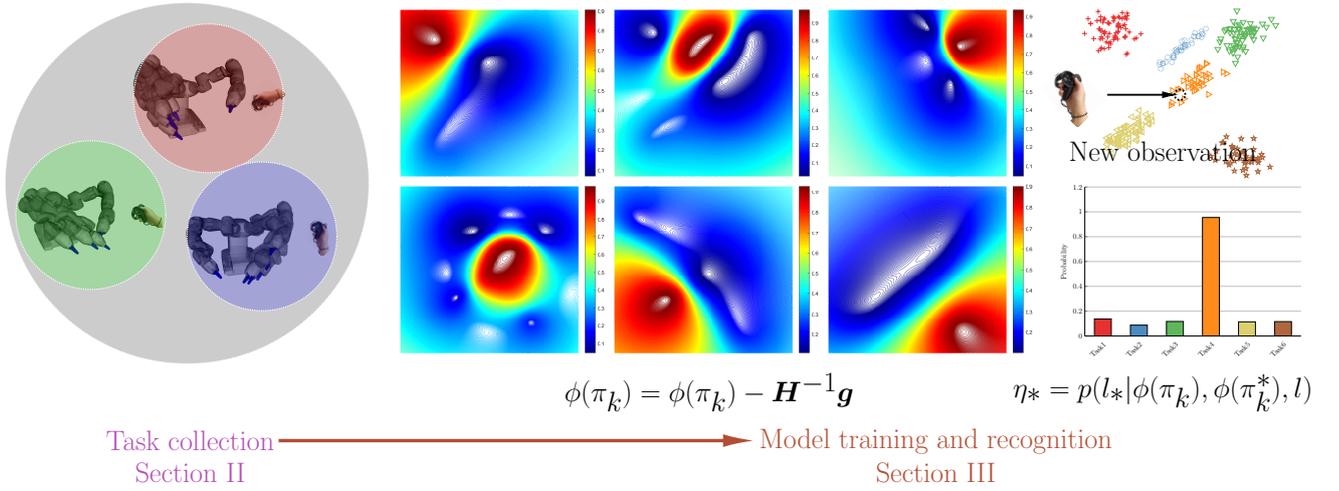
4

Figure 2: The graphic explanation of the multi-task physical human-robot collaboration framework. The framework detailed in Section 2 is a single task template. Therefore, the recognition of multiple tasks is further developed in Section 3. Each task recognition model is trained with iteratively reweighted least squares. Subsequently, given a new observation π_k^*, every binary classification model outputs the probability η_* of the prediction. The task label l_* is corresponding to the maximum probability among these predictions.

3. Multiple Tasks Recognition

The template of the coordinated human-robot collaboration task has been detailed heretofore. However, a single collaboration model cannot always interpret all the tasks exactly. Subsequently, it is not realistic to expect various tasks to own the same temporal and spatial features. In this section, additional feature such as multiple tasks recognition with Gaussian process classification has been developed to enhance the generalisation capability of our framework.

3.1. Multi-task Recognition

The multiple tasks recognition is developed based on Gaussian process binary classification [32]. Thus, the task collection model is defined as $p(l|\pi_k) = \sigma(l\phi(\pi_k))$, where $\sigma(\cdot) = sigm(\cdot)$ is the logistic regression and l_i is the task label.

Basically, the tasks of each human-robot collaboration should be collected regarding the coordinated framework introduced in Section 2. The dataset for recognition, such as time-series human observation π_k, must be collected initially. Then, the log function of the unnormalised posterior distribution is given as follows

$$\mathcal{L}(\phi(\pi_k)) = \log p(l|\phi(\pi_k)) + \log p(\phi(\pi_k)|\pi_k). \quad (19)$$

The negative log function $-\mathcal{L}(\phi)$ is hence optimised using iteratively reweighted least squares algorithm. More specifically, the update of the task collection has the following form

$$\phi(\pi_k) = \phi(\pi_k) - \boldsymbol{H}^{-1}\boldsymbol{g}, \quad (20)$$

with $\boldsymbol{H} = -\nabla\nabla\log p(l|\phi) + \boldsymbol{K}^{-1}$ the Hessian expression of the log likelihood, along with the gradient $\boldsymbol{g} = \nabla p(l|\phi) + \boldsymbol{K}^{-1}\phi$. After convergence, the posterior is thus defined as

$$p(\phi(\pi_k)|\pi_k, l) \sim \mathcal{N}(\hat{\phi}(\pi_k), \boldsymbol{H}^{-1}). \quad (21)$$

The multi-task recognition could be seen as the prediction of the task collection model given the new human observation π_k^*. The mean of the prediction according to the Equ. 21 can be defined as

$$\mathcal{E}[\phi(\pi_k^*)|\pi_k^*, \pi_k, l] = \int \mathcal{E}[\phi(\pi_k^*)|\phi, \pi_k^*, \pi_k, l]p(\phi|\pi_k, l)d\phi$$
$$= k_*^T \boldsymbol{K}^{-1}\mathcal{E}[\phi|\pi_k, l] \approx k_*^T \boldsymbol{K}^{-1}\hat{\phi}. \quad (22)$$

Similarly, the variance of the prediction can be computed using the rule of iterated variance in Equ. 21

$$var[\phi(\pi_k^*)] \approx k_{**} - k_*^T \boldsymbol{H}^{-1}k_*. \quad (23)$$

Therefore, the distribution of the prediction given the observation π_k^* is expressed as $p(\phi(\pi_k^*)|\pi_k^*, \pi_k, l) \sim \mathcal{N}(\mathcal{E}[\phi(\pi_k^*)], var[\phi(\pi_k^*)])$. Consequently, to identify the task label, this distribution η_* for binary response is finally defined as

$$\eta_* = p(l_*|\phi_k, \phi_k^*, l) \approx \int \sigma(\phi_k^*)p(\phi(\pi_k^*)|\pi_k^*, \pi_k, l)d\phi_k^*. \quad (24)$$

A graphic explanation of the multi-task framework is presented in Fig. 2. The training dataset is derived from the initial data stream of the human observations of several different tasks. This setting enables the robot to quickly recognise the task label and respond to the human requirement.

4. Evaluation

Three human-robot collaboration experiments are conducted in this section. Firstly, the comparison between iProMP and our proposed framework is analysed. Then two more real-world applications, i.e., product brochure handover, and multiple tasks recognition are further implemented in order to verify our proposed framework.
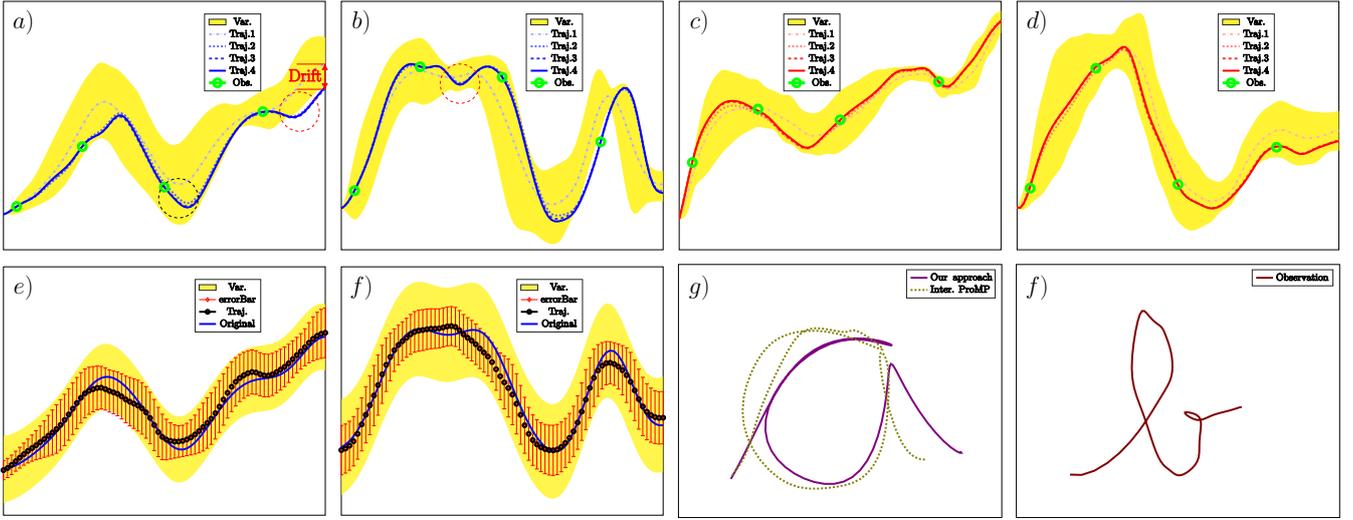
5

Figure 3: The comparison between iProMP and our proposed framework. (a), (b) the axis components of X and Y of letter "a" of the iProMP; (e), (f) the axis components of X and Y of letter "a" of our framework; (c), (d) the axis components of X and Y of letter "b". (g) the comparison result letter "a" of two frameworks; (h) the observation letter "b". In the comparison, handwritten letter "a" and "b" are referred as robot training trajectories and human movement primitives.

4.1. Comparison

In this experiment, two datasets of the handwritten letter "a" and "b" are used separately as robot trajectories and human demonstrations. The capacity of each dataset is 40 trajectories. The task is that when the letter "b" is being written, a new trajectory which is the letter "a" should be written at the same time according to the observation of letter "b". This could be utilised in surface inspection process for Aerospace manufacturing environment.

Before implementing the experiment, two different demonstration letters are trained with the Gaussian process movement primitives presented in Subsection 2.1. Then, the interaction model between two trajectories can be further established according to Subsection 2.2. After the preparation, the implementation closely follows the conditional inference detailed in Subsection 2.3. The compared framework, iProMP [17] is a state-of-the-art coordinated human-robot collaboration approach, which seeks to provide a probabilistic collaboration solution based on Kalman filtering techniques.

The comparison result is given in Fig. 3. Since iProMP responds to sparse observations (four observations in green), the whole trajectories update four times as shown in Fig. 3 (a). This leads to unpredictable trajectory drift, such as the third observation in Fig.3 (a) as presented in the black dashed circle. In the real-world application, the robot cannot respond to this observation at once due to the inertial properties. Thus, a jerk of motion is sometimes unavoidable. Nevertheless, our framework performs a real-time updating manner given the same real-time observation (not four observations, but the whole observation dataset). More specifically, it only updates the next step state variables as given in Fig. 3 (a) and (b). Besides, if further compared, the drift in Fig. 3 (a) is much larger than ours in Fig. 3 (e).

Fundamentally, the iProMP is derived from the linear basis function, which is sometimes easily overfitting, as illustrated in the red dashed circle in Fig. 3 (a) and Fig. 3 (b). Moreover, the iProMP only deals with trajectories which are not dynamic systems. Thus, whenever human observation comes, the whole trajectory must update, even the previous part. Compared with iProMP, the main advantage of our framework is that it is not only an imitation learning trajectory but also a nonparametric hidden state-space dynamic system, which takes full account of the uncertainties propagation. Moreover, our framework is based on Bayesian inference and essentially there is no data fitting in it. Thus, it doesn't suffer from overfitting issues.

4.2. Product Brochure Handover

The product brochure handover experiment aims to provide further explanation of our framework. The platform shown in Fig. 6 consists of a YuMi collaboration robot, a set of Oculus Rift VR devices and a monitor. The VR devices include a headset, two hand touches, and a switch and a controller. In our experiment, the touches are used as a measurement device to collect the human observations. The real-time robot motion can be observed from the screen.

Before implementing the experiment, both human and robot trajectories should be collected at first. The human trajectories are collected by using VR touch, which is basically an IMU device. The robot trajectories are obtained with manually guiding (lead through [33]). More specifically, the human demonstrations collected from the VR touch are presented in light grey lines in Fig. 5 (d), (e) and (f). The YuMi robot trajectories are obtained by manual guidance are given in light grey in Fig. 5 (a), (b) and (c). Then, the whole dynamic system and collaboration model are learned given the details in Section 2.

After the experiment implementation, a sequence records of the product brochure handover is presented in Fig. 4. The VR touch which is utilised to collect the human motions is in the left
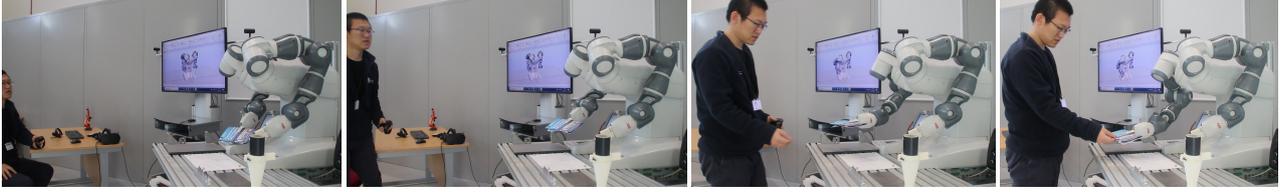
Figure 4: The sequence records of the product brochure handover. The human motion is record during the experiment, along with the movement of YuMi. The human motion is obtained with VR touch in his left hand.
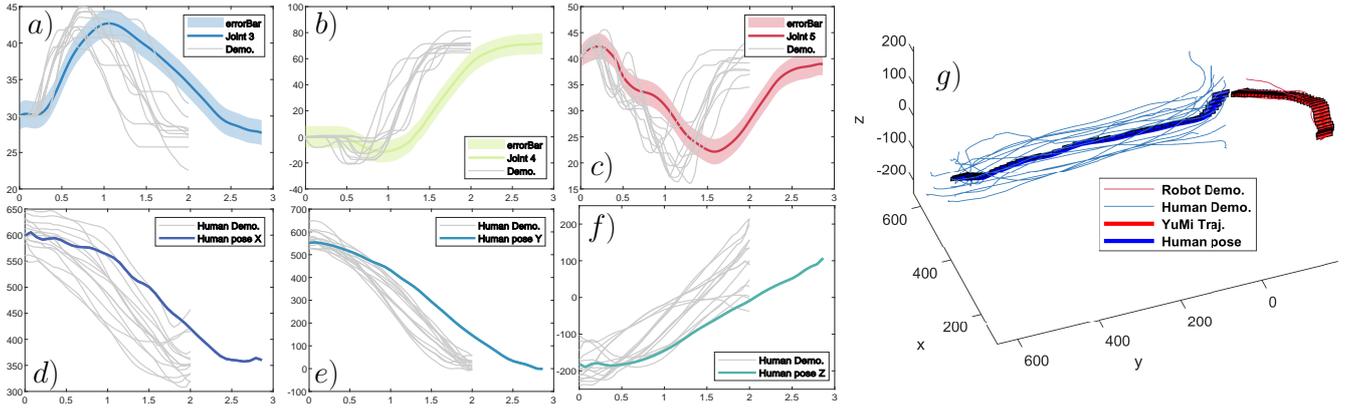

Figure 5: The representation of the collaboration trajectories. (a), (b) and (c), the joint trajectories of the joint 3, joint 4 and joint 5 of the YuMi robot. (d), (e) and (f), three axes trajectories of the human motion. (h), the whole trajectories of the human operator and YuMi robot presented in three-dimension coordinates.
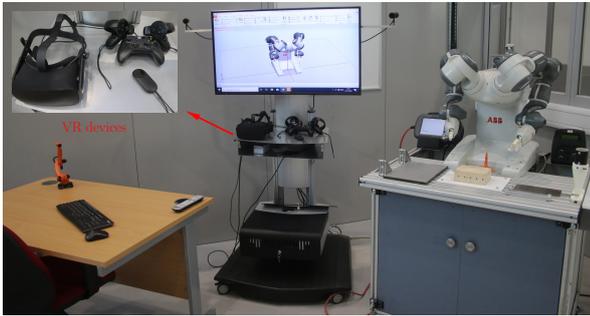

Figure 6: Experiment setting. The experimental platform includes a collaborative robot YuMi, an Oculus Rift VR system and a monitor to check the robot's motion in real-time.

hand of the operator. Subsequently, when the operator walks towards the YuMi, the robot hands over the product brochure to the operator. The human as well as the YuMi trajectories are shown in Fig. 5 in dark blue and dark red, respectively. More specifically, the default learning time period is 2 seconds. However, in the real-time implementation, the duration is nearly 3 seconds, as the human operator could not follow very precise procedure. Therefore, the left six figures in Fig. 5 present the trajectories of the joints and axes during the experiments. In addition, this demonstrates the automatic time alignment introduced in our framework.

Compared with the simulation result presented in Fig. 3, the trajectories in color given in Fig. 5 (a), (b), and (c) do not quite match with the human demonstrations in gray. However, they

are still in the area that covered by the human demonstrations (if ignore the time scale). This is because even the human operators try their best to repeat the same tasks with same motions again and again, there is still difference among each demonstration as shown in gray lines in Fig. 5. In addition, during the real-time experiment, the VR touch does input certain amount of noise to the control system as illustrated in Fig. 5 (d), (e), and (f). This is the reason that the human-robot collaboration in this paper is established based on the dynamic hidden state-space models. Also this is the reason that the uncertainty is considered and propagated through the dynamic and measurement systems. With the uncertainty propagation, the proposed framework is aware of the human intention and outputs the predicted robot trajectories and their confidential intervals, which are the error bars as shown in Fig. 5.

4.3. Multi-task Experiment

The multi-task functionality aims to improve the generalisation capability and allow the robot to quickly recognise the operator's intention from observations. More specifically, the collaborative robot should automatically choose the correct task from the task container and physically collaborate with a human operator according to the multi-task recognition model. In this subsection, the task container consists of three tasks here, i.e., rivet handover, nutplate handover and sealing assistance. According to the graphic explanation in Fig. 2, each task is constructed under the coordinated human-robot collaboration framework as detailed in Section 2.

Besides the same experiment setting as the previous experiment as presented in Fig. 6, the initial data stream of the human
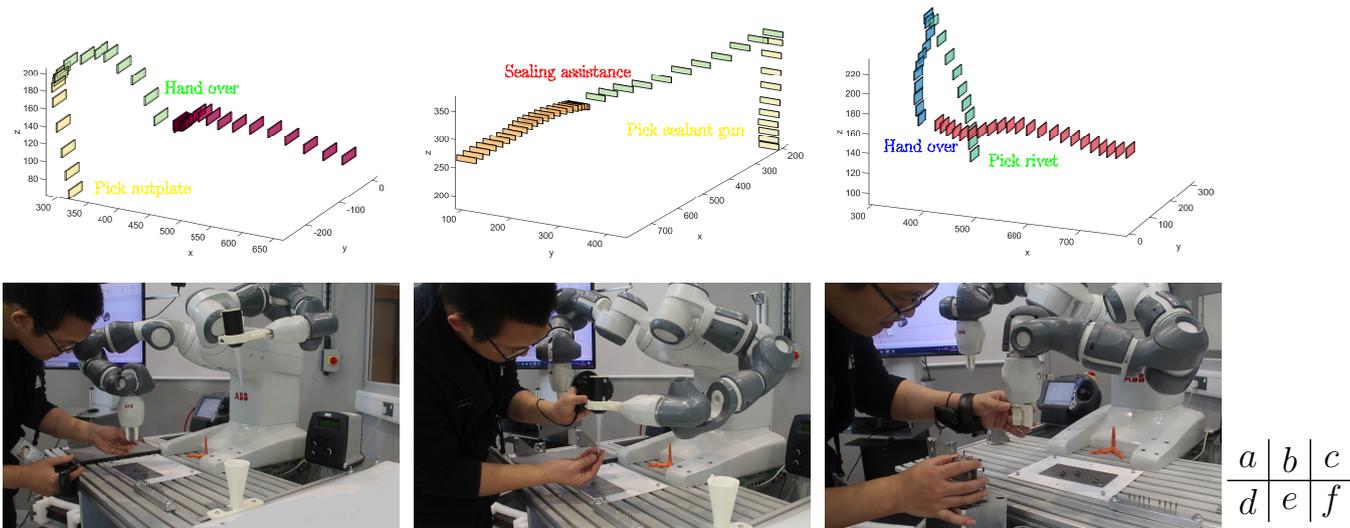
Figure 7: The experiments of the multi-task recognition. (a) and (d), nutplate handover; (b) and (e), sealing assistance; (c) and (f), rivet handover. Three experiments, i.e., nutplate handover, sealing assistance and rivet handover are presented in the above figures. All the key actions during the collaboration, such as nutplate picking and hand over are also listed in the figures.
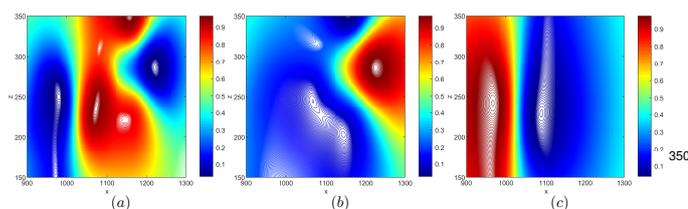

Figure 8: Training result of the task collection model with binary Gaussian process classification. (a) Training result of the nutplate handover. (b) Training result of the sealing assistance. (c) Training result of the rivet handover.

observations for each task should be captured for training multi-task recognition model at first. After the data collection, since the binary labels are used in Gaussian process classification, three recognition models corresponding to each collaboration task should be trained separately, as represented in Fig. 8.

Consequently, there are two regions of interest in each sub-figure or recognition model. The region of interest in red shows a high probability, which means that for example, the data fallen into the red region in Fig. 8 (a) is more likely to be the nutplate handover task. Thus, the YuMi robot would pick the nutplate and hand it over to the operator. On the contrary, the blue region shows a lower probability result, which indicates that the operator chooses a different task but not this task. We would like to point out that the recognition models given in Fig. 8 (a), (b) and (c), are two dimension instance for explanation purpose.

During the experiment, the operators select the task randomly. However, in order to obtain a more statistical analysis, we conduct ten times testing for each task. The input is still the human initial observations, i.e., three-dimension information. The snapshot of three tasks and their trajectories are given in Fig. 7, respectively. Moreover, the probabilities for the three tasks inferred by the multitask recognition model are also recorded in the bar chart as shown Fig. 9.

In Fig. 9, the task labels at the bottom are the ground truth of the human operator selections. For each task, the above three columns indicate the probability information inferred by the learning multitask recognition model. For instance, for the ten times testing of the first task, the recognition model outputs the probability of at around about 0.7 of the first task, which is approximately twice larger than the probabilities of second and third tasks. The results indicate that the human operator is more likely to choose the first task. Therefore, the YuMi robot should response accordingly to the operator and implement the first collaboration task. Similarly, the other two task testings also show the correct probabilities at roughly 0.8. Evidently, the learned results of maximum probabilities match with the ground truth. Thus, the application of the feasibility of our multitask recognition model is verified.
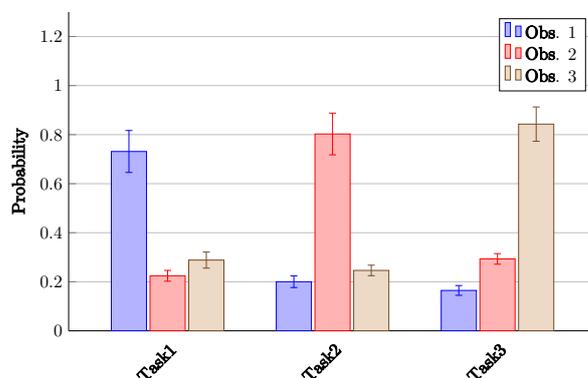

Figure 9: Experimental result of the multi-task recognition. The mean and standard deviation derived from ten test for each task are presented above.

5. Conclusions

This paper proposes a novel imitation learning framework for coordinated human-robot collaboration derived from the hidden state-space models. After learning from the human demonstrations, given only the spatial information of the human observation, this framework can effectively provide a temporal and spatial alignment solution for coordinated human-robot collaboration. Additional crucial function, i.e. multiple task recognition is also addressed in this paper to provide more flexibility.

The proposed nonparametric framework responds to the human observation and its uncertainties based on the conditional inference of the exact moment matching approximation. Based on the hidden dynamic functionalities, both state variables and their uncertainties are propagated through the framework. The simulation experiments not only compare the iProMP approach and our proposed framework but also present the main advantages and functionalities of our work. In addition, the product brochure handover is conducted to further verify the real-world application based on our framework. The multi-task experiment is implemented to demonstrate the recognition of the human intention corresponding to various task labels. Therefore, the future work will be focused on the compliance physical interaction and human-robot collaboration in Cartesian space.

Funding

References

[1] W. Ji, L. Wang, Industrial robotic machining: a review, The International Journal of Advanced manufacturing Technology 103 (1-4) (2019) 1239–1255.

[2] V. Villani, F. Pini, F. Leali, C. Secchi, Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications, Mechatronics 55 (2018) 248–266.

[3] D. Mukherjee, K. Gupta, L. H. Chang, H. Najjaran, A survey of robot learning strategies for human-robot collaboration in industrial settings, Robotics and Computer-Integrated Manufacturing 73 (2022) 102231.

[4] L. Wang, Z. Du, W. Dong, Y. Shen, G. Zhao, Hierarchical human machine interaction learning for a lower extremity augmentation device, International Journal of Social Robotics 11 (1) (2019) 123–139.

[5] A. Sidiropoulos, E. Psomopoulou, Z. Doulgeri, A human inspired handover policy using gaussian mixture models and haptic cues, Autonomous Robots 43 (6) (2019) 1327–1342.

[6] L. Wang, S. Jia, G. Wang, A. Turner, S. Ratchev, Enhancing learning capabilities of movement primitives under distributed probabilistic framework for flexible assembly tasks, Neural Computing and Applications (Oct 2021).

[7] F. Stulp, E. A. Theodorou, S. Schaal, Reinforcement learning with sequences of motion primitives for robust manipulation, IEEE Transactions on robotics 28 (6) (2012) 1360–1370.

[8] A. Coates, P. Abbeel, A. Y. Ng, Learning for control from multiple demonstrations, in: Proceedings of the 25th international conference on Machine learning, 2008, pp. 144–151.

[9] G. Ye, R. Alterovitz, Guided motion planning, in: Robotics research, Springer, 2017, pp. 291–307.

[10] T. Callens, A. van der Have, S. Van Rossom, J. De Schutter, E. Aertbelien, A framework for recognition and prediction of human motions in human-robot collaboration using probabilistic motion models, IEEE Robotics and Automation Letters (2020).

[11] H. B. Amor, G. Neumann, S. Kamthe, O. Kroemer, J. Peters, Interaction primitives for human-robot cooperation tasks, in: 2014 IEEE international conference on robotics and automation (ICRA), IEEE, 2014, pp. 2831–2837.

[12] Y. Tanaka, J. Kinugawa, Y. Sugahara, K. Kosuge, Motion planning with worker's trajectory prediction for assembly task partner robot, in: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2012, pp. 1525–1532.

[13] C. Pérez-D'Arpino, J. A. Shah, Fast target prediction of human reaching motion for cooperative human-robot manipulation tasks using time series classification, in: 2015 IEEE international conference on robotics and automation (ICRA), IEEE, 2015, pp. 6175–6182.

[14] H. S. Koppula, A. Saxena, Anticipating human activities using object affordances for reactive robotic response, IEEE transactions on pattern analysis and machine intelligence 38 (1) (2015) 14–29.

[15] S. Calinon, A tutorial on task-parameterized movement learning and retrieval, Intelligent service robotics 9 (1) (2016) 1–29.

[16] Y. Zhou, M. Do, T. Asfour, Coordinate change dynamic movement primitives—a leader-follower approach, in: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2016, pp. 5481–5488.

[17] G. J. Maeda, G. Neumann, M. Ewerton, R. Lioutikov, O. Kroemer, J. Peters, Probabilistic movement primitives for coordination of multiple human–robot collaborative tasks, Autonomous Robots 41 (3) (2017) 593–612.

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

[19] A. Al-Yacoub, Y. Zhao, W. Eaton, Y. Goh, N. Lohse, Improving human robot collaboration through force/torque based learning for object manipulation, Robotics and Computer-Integrated Manufacturing 69 (2021) 102111.

[20] D. A. Duque, F. A. Prieto, J. G. Hoyos, Trajectory generation for robotic assembly operations using learning by demonstration, Robotics and Computer-Integrated Manufacturing 57 (2019) 292–302.

[21] P. Tsarouchi, A.-S. Matthaiakis, S. Makris, G. Chryssolouris, On a human-robot collaboration in an assembly cell, International Journal of Computer Integrated Manufacturing 30 (6) (2017) 580–589.

[22] S. Papanastasiou, N. Kousi, P. Karagiannis, C. Gkournelos, A. Papavasileiou, K. Dimoulas, K. Baris, S. Koukas, G. Michalos, S. Makris, Towards seamless human robot collaboration: integrating multimodal interaction, The International Journal of Advanced Manufacturing Technology 105 (9) (2019) 3881–3897.

[23] S. Makris, Synthesis of data from multiple sensors and wearables for human–robot collaboration, in: Cooperating Robots for Flexible Manufacturing, Springer, 2021, pp. 321–338.

[24] N. Nikolakis, V. Maratos, S. Makris, A cyber physical system (cps) approach for safe human-robot collaboration in a shared workplace, Robotics and Computer-Integrated Manufacturing 56 (2019) 233–243.

[25] H. Liu, L. Wang, Collision-free human-robot collaboration based on context awareness, Robotics and Computer-Integrated Manufacturing 67 (2021) 101997.

[26] X. Yu, W. He, Y. Li, C. Xue, J. Li, J. Zou, C. Yang, Bayesian estimation of human impedance and motion intention for human-robot collaboration, IEEE transactions on cybernetics (2019).

[27] R. Zhang, Q. Lv, J. Li, J. Bao, T. Liu, S. Liu, A reinforcement learning method for human-robot collaboration in assembly tasks, Robotics and Computer-Integrated Manufacturing 73 (2022) 102227.

[28] S. H. Choi, K.-B. Park, D. H. Roh, J. Y. Lee, M. Mohammed, Y. Ghasemi, H. Jeong, An integrated mixed reality system for safety-aware human-robot collaboration using deep learning and digital twin generation, Robotics and Computer-Integrated Manufacturing 73 (2022) 102258.

[29] T. Hatanaka, K. Noda, J. Yamauchi, K. Sokabe, K. Shimamoto, M. Fujita, Human-robot collaboration with variable autonomy via gaussian process, IFAC-PapersOnLine 53 (5) (2020) 126–133.

[30] E. Solak, R. Murray-Smith, W. Leithead, D. Leith, C. Rasmussen, Derivative observations in gaussian process models of dynamic systems, Ad-

vances in neural information processing systems 15 (2002) 1057–1064.

[31] M. P. Deisenroth, Efficient reinforcement learning using Gaussian processes, Vol. 9, KIT Scientific Publishing, 2010.

[32] C. E. Rasmussen, Gaussian processes in machine learning, in: Summer School on Machine Learning, Springer, 2003, pp. 63–71.

[33] F.-P. Kirgis, P. Katsos, M. Kohlmaier, Collaborative robotics, in: Robotic Fabrication in Architecture, Art and Design 2016, Springer, 2016, pp. 448–453.