Preprint version; final version available at https://ieeexplore.ieee.org/document/8685136
Transactions on Robotics
DOI: 10.1109/TRO.2019.2904461

Probabilistic Real-Time User Posture Tracking for Personalized Robot-Assisted Dressing

Fan Zhang, Antoine Cully, and Yiannis Demiris, *Senior Member, IEEE*

***Abstract*—Robotic solutions to dressing assistance have the potential to provide tremendous support for elderly and disabled people. However, unexpected user movements may lead to dressing failures or even pose a risk to the user. Tracking such user movements with vision sensors is challenging due to severe visual occlusions created by the robot and clothes. We propose a probabilistic tracking method using Bayesian networks in latent spaces, which fuses robot end-effector positions and force information to enable camera-less and real-time estimation of the user postures during dressing. The latent spaces are created before dressing by modeling the user movements with a Gaussian Process Latent Variable Model, taking the user's movement limitations into account. We introduce a robot-assisted dressing system that combines our tracking method with hierarchical multi-task control to minimize the force between the user and the robot. The experimental results demonstrate the robustness and accuracy of our tracking method. The proposed method enables the Baxter robot to provide personalized dressing assistance in putting on a sleeveless jacket for users with (simulated) upper-body impairments.**

***Index Terms*—Personalized dressing assistance, probabilistic real-time tracking, user modeling in latent spaces.**

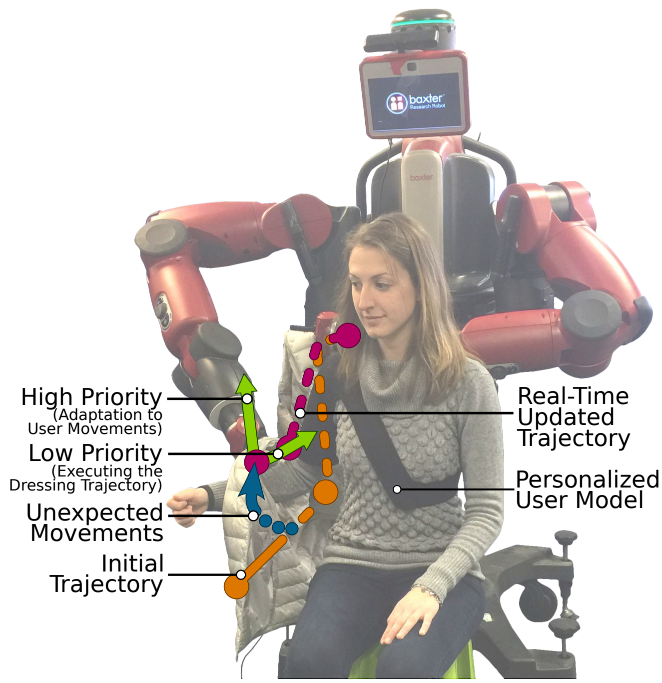


Fig. 1: The Baxter humanoid robot provides personalized dressing assistance in putting on a sleeveless jacket for human users with different simulated upper-body movement limitations. This is achieved with a real-time posture tracking and a user model construction in latent spaces are combined with a multi-task hierarchical control strategy introduced in this paper. Here, healthy users simulate upper-body impairments by wearing braces which limit their movements. The camera on the top of the robot is not used in this paper.

I. Introduction

DRESSING assistance is a basic assistive activity in the daily life of the elderly and disabled people who suffer from impairments. Studies indicate that although over 80% of people in skilled nursing facilities require assistance with dressing [1], there is growing concern regarding the increased costs of daily care and lack of nursing staff [2]. A cross-sectional study [3] of 14,500 Medicare beneficiaries shows that 8.2% of those beneficiaries have reported difficulties with dressing. [3] also reports that of all the activities of daily living, dressing has shown the highest burden on caregiving staff and the lowest use of assistive technologies. In this work, we focus on robot-assisted dressing, which is a growing area of research that has the potential to alleviate this problem.

In recent studies, robot-assisted dressing is usually described as a trajectory planning problem. A common approach to solve this problem is recognizing the user's posture and planning the dressing trajectory accordingly using feedback from vision sensors [4-8]. However, severe visual occlusions are created when the robot, the clothes and the user body are in close contact during the dressing sequence, which often lead to failures in real-time human posture tracking [9, 10]. For instance, depth cameras and motion capture systems, which are often used for human posture tracking, typically fail to extract the user's skeleton due to such occlusions. As a result, user postures may not be observed accurately in real time when the human user moves during the dressing sequence, and thus the dressing trajectory cannot be planned reliably.

Current research usually assumes that users remain still during dressing [10-13]. However, human users may move their arms during the dressing sequence, for instance, performing other activities (*e.g.*, grasping objects) or adapting their arm postures when they feel uncomfortable or tired. Such unexpected movements may result in undesired large forces between the robot and the user, which may lower the comfort of the user, or even put the user at risk and lead to dressing failures. In our previous research [14], we have developed methods to ensure that the robot will not interfere with user movements, while providing dressing assistance. We have represented the dressing task as two hierarchical subtasks: 1)

The authors are with the Personal Robotics Laboratory, Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, U.K. (e-mail:f.zhang16@imperial.ac.uk; a.cully@imperial.ac.uk; y.demiris@imperial.ac.uk). This research is supported by a Royal Academy of Engineering Chair in Emerging Technologies to Yiannis Demiris.

the high-priority subtask adapts robot motions to minimize the force applied between the user and the robot caused by human user movements, and 2) the low-priority subtask completes the dressing trajectory such that the human is properly dressed. Due to unexpected user movements, the dressing trajectory in the low-priority subtask needs to be updated in real time. However, updating the dressing trajectory requires real-time user posture tracking during the dressing sequence which is very challenging through vision-based approach due to severe occlusions.

In order to achieve precise and real-time user posture tracking, in this paper, we present a probabilistic filtering method using two sources of information, *i.e.*, robot end-effector positions, and force information between the end-effector and the user. We formulate the problem in the framework of recursive Bayesian estimation in latent spaces, enabling online fusion of the two sources of information using a particle filter. As vision sensors often cannot directly view task-relevant aspects of the users due to severe occlusions from clothes and the robot's own body during dressing, robot end-effector positions and applied forces provide critical information for posture tracking. This is inspired in part by the notion that a human can use what he/she feels while performing the dressing task for someone else to infer where the arm is or if the dressing is finished. The proposed method follows this principle, and enables camera-less user posture tracking during the dressing process.

As the features representing human movements are high-dimensional and non-linear, they are particularly challenging to study [15, 16]. Especially in our case, it is challenging to implement probabilistic posture tracking using Bayesian networks in the high-dimensional Cartesian space, as the computational requirements of the particle filter scale exponentially with the number of dimensions of the state space. Expensive explicit parameterization (*e.g.*, morphological properties of the user's arm including arm lengths, angular limits) is required when using a particle filter in high-dimensions to avoid unergonomic or degenerated arm postures being sampled. In this work, we propose to model the user postures using latent spaces to make the tracking more tractable. In our previous research [14], we used a Gaussian Process Latent Variable Model (GP-LVM) [17] to project the user's high-dimensional upper-body movements onto a 2-dimensional latent space, which preserves the morphological properties of the user's arm. The model also captures the specificities of different upper-body movement limitations, taking into account a trade-off between the users' movement capabilities and their preferences, which are otherwise challenging to manually describe in high-dimensional space. In the case of impaired users, such a personalized user model is crucial as it captures only the postures that users can reach and prefer, and thus guarantees that only these postures could be estimated in our tracking system.

Experimental validations have been carried out by using the Baxter humanoid robot to provide personalized dressing assistance in putting on a sleeveless jacket for 20 human users with simulated upper-body movement limitations (Fig. 1). The experiments cover a range of challenging conditions, including severe visual occlusions and varying user movements. We quantitatively and qualitatively demonstrate the robustness and accuracy of our tracking method. The evaluation of the whole system shows that it outperforms three baselines from the state-of-the-art, including using a pre-defined dressing trajectory (similarly to the vast majority of the literature [10, 18, 19]) with hierarchical control to minimize the interaction, updating the dressing trajectory with our tracking approach but without the hierarchical control, and the robot-assisted dressing approach that we introduced in [14]. The experiments here assume that the user's hand is already inside the sleeve opening.

In our previous research [14], we presented the hierarchical multi-task control approach (Section IV) and the user model construction method (Section V). For the completeness of the whole proposed system, we will briefly outline these two methods to describe how they are combined with the probabilistic real-time posture tracking method (Section VI) proposed in this paper.

The main contributions of this paper can be summarized as follows:

1) A precise, real-time user posture tracking method is proposed based on a probabilistic filter using multi-modal information. This method enables camera-less user posture tracking during the dressing process. The depth camera is only needed for initial posture recognition and user model construction before dressing.
2) The tracking method is performed in low-dimensional latent spaces provided by user model construction, which would be challenging in the original Cartesian space because of the inherent high-dimensionality feature of human movements. The user model also captures the specificities of upper-body impairments, which enables the Baxter humanoid robot to provide personalized dressing assistance for users with different upper-body movement limitations.
3) A robot-assisted dressing system is proposed, combining the tracking method with a hierarchical multi-task controller that adapts robot motions to minimize the force between the user and the robot.

II. Related Work

A. Robot-Assisted Dressing

There exist several approaches to robot-assisted dressing in home environments. For example, reinforcement learning has been used to enable a robot to wrap a scarf [4] and to dress a T-shirt on mannequins with different head and shoulder inclinations [5, 6]. In [11], learning by demonstration has been proposed to encode and adapt dressing skills. Clothing models [12, 20] and clothes-user topological relationships [7, 8, 21] have been studied to optimize the dressing trajectory. While these different studies lead to successful and promising results, they usually use fixed mannequins for experiments and thus assume that the human user stays perfectly still. However, unexpected user movements from human users can make the robot apply large forces on users, which may lead to failures in dressing assistance or even pose a risk to the user.

Force information during the dressing sequence has been used in recent research to adjust the robot motion to minimize the interaction force [22]. For example, some studies [23, 24] have presented solutions to use the force resistance applied between the robot and the human body to address classification-type problems, such as determining the garment type or the outcome of a robot pulling the sleeve of a hospital gown. Other studies have used detected force [10, 18, 19] or real-time predicted force through a deep recurrent network [25] to adjust the robot motion with a position controller. Although these methods can adjust the robot trajectory based on the force information, this is not achieved in real time. In these studies, the human users usually remain still, and forces usually occur due to either snagging of the garment at some point on the user arm or the friction between the garment and the user. In these cases, the delayed trajectory update still yields some good results. However, when dealing with forces generated by fast moving users' arms, such delay in the trajectory update may put the user in an uncomfortable or even dangerous state.

One common solution in assistive robotics to account for user interactions in real time is using compliant controllers [26]. They have achieved competitive performance in several robot-assisted scenarios (*e.g.*, drawing [27], dual-arm grasping in [28]). In our previous research [14], we presented another alternative, a hierarchical multi-task controller, which can automatically adapt the robot motions in accordance with unexpected user movements by constantly minimizing the force applied between the user and the robot, while executing the dressing task. The architecture presented in this paper will use this approach to minimize the force applied on the user during the dressing.

B. Human Posture Tracking

In order to plan the dressing trajectory in real time, user postures should be obtained during the dressing sequence. Vision information has been shown to be effective in human posture tracking. For example, a RGB camera [11], a depth camera [29] and a high-cost motion capture system [5] have been used to detect the poses of mannequins or human users. In [30, 31], optical flow has been calculated to plan the trousers dressing trajectory. Although these studies yield some successful posture recognition and dressing behaviors, only the initial postures of the fixed mannequins or immobile human users are detected by vision sensors. However, the cameras easily lose track of the human skeletons when users move their arms during the dressing sequence, with the underlying reason being the severe occlusions created by the clothes and the robot, and the morphological similarities between the arms of the robot and the human users.

Deep learning has enjoyed substantial attention in the human pose estimation community [32]. Recent approaches usually train deep neural networks to regress 2D or 3D poses directly from images. OpenPose [33] has efficiently detected the 2D pose of multiple people in an image. DensePose [34] has achieved impressive performance in mapping the pixels of humans in 2D images, containing also background and occlusions, to a 3D representation of the human body. Several other studies [35-37] tackle the 3D human pose estimation task by training deep network architectures. Another approach to predict occluded parts of a posture is to rely on the parts of the posture that can still be observed. This approach has been proposed by [37], in which a Long Short-Term Memory network has been trained for predicting the left elbow position of a user being dressed, based on features observed from a depth-camera placed in front of the robot. However, the proposed model offered predictions with an accuracy of over 5cm on each of the X/Y/Z axis when using a Kinect camera.

Only data from the vision modality has been used in the above presented approaches for posture tracking. There exist studies using multi-modal information to compensate for the disadvantages of using vision information only. [38] has designed a capacitive proximity sensing system to track user movements during dressing, which however requires the user poses to be approximately parallel to the capacitive sensor. In [39], in order to avoid the errors of forward kinematics and frame transformations, a recursive Bayesian estimation algorithm that fuses joint measurements with depth images has been formulated to achieve robot arm tracking. In [40], a Bayesian filtering framework has been presented, which fuses point cloud information from an RGB-D camera with tactile information from a contact sensor, to track objects in-hand manipulated by the robot. Bayesian networks have also been successfully used in [41-44] to track human postures. However, to our knowledge, no paper from the literature uses Bayesian networks to track human postures in the application of robot-assisted dressing. In this work, we propose to use Bayesian networks, which fuse multi-modal information (*i.e.*, robot end-effector positions, and force information), to yield a precise, real-time estimate of the human arm postures. This method enables camera-less user posture tracking during dressing. Real-time updates of the trajectory also ensure a one-shot dressing, which avoids multiple trials and errors in [5, 6, 10, 11] that may put the users at risk. The proposed method also avoids markers attached on users or clothes [11, 12, 45] that may cause inconvenience in daily dressing scenarios.

C. User Modeling

Human movements vary significantly according to the arm configurations users may prefer and the different impairments users may suffer from, which are challenging to manually describe. To provide assistance for different users, one typical approach is to construct personalized user movement models, which are capable of capturing preferred and/or reachable movements for different impaired users.

Previous work [46-50] has learned user preferences by constructing user models. Multi-modal data (*e.g.*, speech communication, user speed and verbal feedback) have been used to infer user preferences in [47, 48]. [49, 50] aim to analyze which robot hand-over configurations users prefer most. In our previous work [14], we projected the upper-body movements onto a latent space and computed its density plot. The denser locations in the density plot represent the upper-body postures which are more frequently adopted by the users, and we

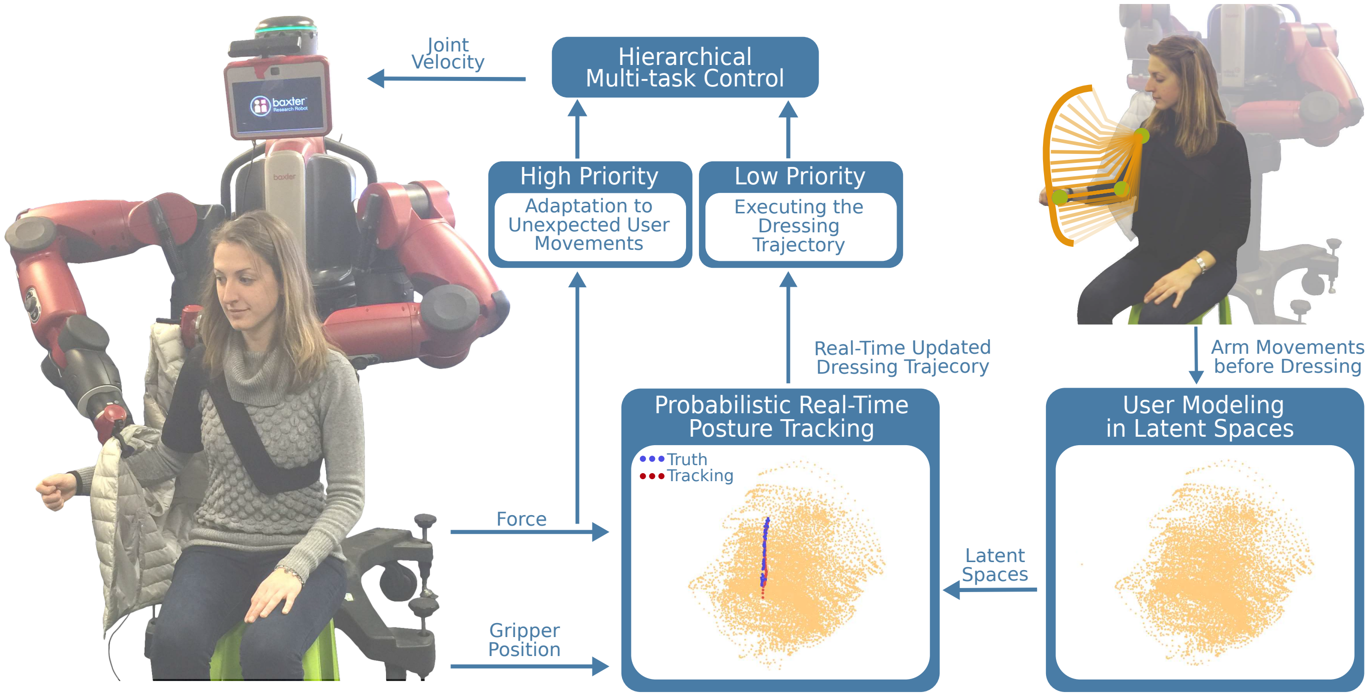


Fig. 2: Proposed framework for robot-assisted dressing. The dressing process is represented as two hierarchically organized subtasks. The high-priority subtask adapts robot motions in accordance with unexpected user movements. The low-priority subtask executes the dressing trajectory, which is updated in real time with the proposed probabilistic posture tracking method using multi-modal data, *i.e.*, force information between the end-effector and the user, and end-effector positions. The user model is constructed before dressing, which provides latent spaces for posture tracking and captures the specificities of the user's impairment.

have used this information to update the robot trajectory accordingly. Other studies [51-54] have described the user movement capabilities to personalize robot assistance rather than user preferences. Multivariate Gaussian distributions have been successfully used to statistically identify the range of hand movement capabilities in a planar space for robotic rehabilitation [51, 52]. [53] has modeled the movement velocity of stroke survivors' arms with a multivariate Gaussian kernel, and accordingly personalized a robot-applied force intervention that encourages under-expressed movements. [54] has used support vector machine, with the fusion of surface electromyography and mechanomyography signals, to describe the elbow spasticity. In [13], Gaussian Mixture Models (GMMs) have been used to model the movement space of upper-body joints (hand, elbow and shoulder) separately in 3-D coordinates. However, GMMs cannot capture the correlations of the upper-body joint movements, *e.g.*, the impact of hand and shoulder movements on elbow movements.

As discussed in Section I, user posture tracking using a particle filter suffers from the curse of dimensionality (*e.g.*, the expensive computation and explicit parameterization in high-dimensions). Many studies have extended the particle filter to alleviate such problems by adopting different searching strategies to reduce the dimensionality. For example, [44] has proposed to track human motion by decomposing high-dimensional space into several lower-dimensional spaces to reduce the computational time of particle filters. In [39], a coordinate particle filter [55] has been used to estimate robot end-effector poses by computing the approximate weights of one dimension at a time to avoid the degeneracy in high-dimensional state spaces. Although the above methods have successfully reduced the dimensionality, the particle filter is still calculated in relatively high-dimensional spaces, and a parameterized user model or robot URDF (Unified Robot Description Format) is required.

Other methods [15, 16] have used dimensionality reduction algorithms to model human movements directly in low-dimensional latent spaces. The non-linearity feature of human behaviors makes linear dimensionality reduction approaches (*e.g.*, Principal Component Analysis [56], Probabilistic Principal Component Analysis [57]) often fail to accurately model human movements. Instead, Nonlinear Latent Variable Models are attractive for describing human movements due to their property of capturing nonlinearities efficiently [58]. It has been shown that Gaussian Process Latent Variable Model (GP-LVM) [17] can efficiently embed human actions into a 2-D latent space while preserving most stylistic information (*e.g.*, the morphological properties of the user's arm) [59], which makes GP-LVM attractive for modeling the upper-body movements.

In this paper, we use a GP-LVM to model the user movement limitations, which provides latent spaces for real-time human posture tracking. It also automatically captures all the user parameters in the latent representation. Such a latent model is also capable of capturing the specificities of different upper-body impairments (as demonstrated in the first experiment of this paper).

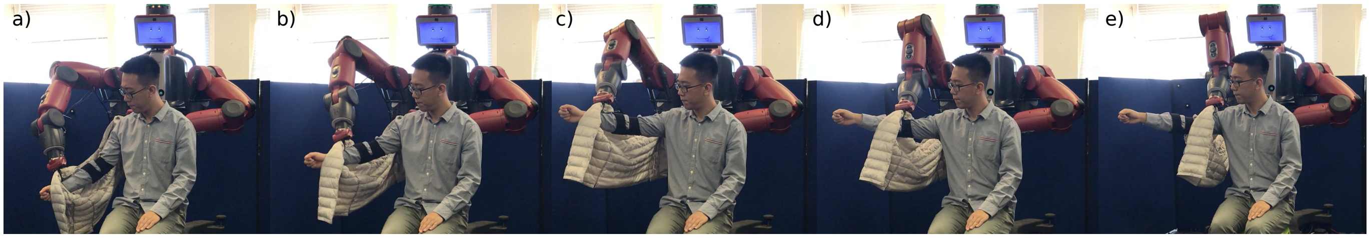


Fig. 3: Snapshots of the robot-assisted dressing sequence from a user with simulated elbow movement limitations. (a) The robot reaches the initial hand position and starts the dressing trajectory. (b) The user moves his arm suddenly. (c) The robot adapts its motion accordingly and updates its trajectory in real time. (d) The robot reaches the updated elbow position. (e) The robot finishes dressing the right arm.

III. Framework Description

In this section, we present the framework of our proposed robot-assisted dressing system, as shown in Fig. 2. As discussed in Section I, we presented a hierarchical multi-task control approach (Section IV) and a user model construction method (Section V) in our previous research [14]. We briefly outline these two components here to describe how they are combined with the probabilistic real-time posture tracking method (Section VI) proposed in this paper.

Similarly to the scenarios in which a human assistant helps another person to get dressed, the humanoid robot dresses a user's forearm first, and then the upper arm. Therefore, the dressing sequence is defined as: hand, forearm, elbow, upper arm, and shoulder. During the initialization, the user's hand, elbow and shoulder positions, detected by the depth camera, are defined as the waypoints of the initial dressing trajectory. Then the robot gripper starts to move the clothes to the initial hand position. During the dressing sequence, the robot motion is designed based on the hierarchical multi-task controller presented in Section IV. We have defined the dressing process as two hierarchically-arranged subtasks. The high-priority subtask adapts robot motions in accordance with unexpected user movements to minimize the interaction force. The low-priority subtask executes the dressing trajectory, which needs to be updated in real time due to the unexpected user movements. In Section VI, we describe a probabilistic real-time posture tracking method to update the dressing trajectory using multi-modal data.

A user model provides low-dimensional latent spaces for the proposed user posture tracking method, which is challenging to perform in the Cartesian space because of the inherent high-dimensionality feature of human movements. The user model is constructed before dressing and only once, and takes different upper-body movement limitations into account. This is achieved by asking the user to move his/her arm randomly prior to the dressing scenario. We record the arm posture data and project such data onto a 2-dimensional latent space with GP-LVM. This method is detailed in Section V.

Meanwhile, the motion of the other gripper is designed to cooperate with the dressing motion. For instance, if the robot is dressing the right impaired arm, the left robot gripper will grasp the other opening of the clothes and hold the clothes at the user's right shoulder. After the robot finishes dressing the right arm, the left robot gripper will move the clothes to the user's left area so that the user can wear the left part of the clothes without reaching difficulty. In this paper, restricted by the workspace of the robot manipulating the clothes, we assume that only users' right arms suffer from impairments and thus require dressing assistance.

Based on the preliminary experiments, we have observed that shoulder movements are usually small enough to have no impact on the dressing trajectory. Therefore, we assume for this study that the shoulder position remains constant during the dressing process. In our future work, we plan to lift this assumption and take these small changes into account when assisting users with real disabilities. Fig. 3 shows an example of the dressing sequence when the robot provides dressing assistance for a user with simulated elbow movement limitations. This dressing sequence shows the robot motion adaptation when the human user moves unexpectedly during the dressing sequence and the real-time update of the dressing trajectory.

IV. Hierarchical Multi-task Control for Robot-Assisted Dressing

Following the method proposed in our previous research [14], we model the dressing process as two hierarchically organized subtasks to cope with unexpected user movements during the dressing sequence. Although the current hierarchical system only has two levels, it has been designed to accommodate additional tasks of different priority if needed in the future.

A. Task Priorities for Robot-Assisted Dressing

A strict hierarchy of subtasks is adopted in this work, where a hierarchical ordering of the subtasks is predefined. The low-priority subtasks are solved in the null-space of the high-priority subtasks [60, 61]. The two subtasks are combined in a velocity-input hybrid force/position control scheme according to the equation

$$\dot{\boldsymbol{q}} = \boldsymbol{J}_c^{\dagger}\boldsymbol{v} + (\boldsymbol{I} - \boldsymbol{J}_c^{\dagger}\boldsymbol{J}_c)\dot{\boldsymbol{q}}_0 \tag{1}$$

where $\dot{\boldsymbol{q}}_0$ denotes the desired joint velocity of the low-priority subtask, $\boldsymbol{v}$ denotes the desired end-effector velocity of the high priority task and $\dot{\boldsymbol{q}}$ denotes the joint velocity. The matrix $\boldsymbol{J}_c$ is obtained by projecting the geometric Jacobian Matrix $\boldsymbol{J}$ along the end-effector force direction, $\boldsymbol{J}_c^{\dagger}$ is the pseudoinverse of $\boldsymbol{J}_c$, and $(\boldsymbol{I} - \boldsymbol{J}_c^{\dagger}\boldsymbol{J}_c)$ represents the orthogonal projection matrix in the null-space of $\boldsymbol{J}_c$.

The projection of the geometric Jacobian Matrix $\boldsymbol{J}$ along the end-effector force direction implies that the robot is allowed to move in directions that are orthogonal to the direction of the force to achieve the low-priority subtask, while moving according to the direction of the force to achieve the high-priority subtask.

B. High-Priority Subtask

The high-priority subtask ensures that the robot compensates for human movements and allows the user to perform other activities during the dressing sequence without lowering user's comfort and security, using a force controller. As there is no force sensor at the wrist of the Baxter, we estimate the external force and torque applied at the end-effector of the Baxter robot from the measured joint torques, as proposed in [60]:

$$\boldsymbol{h} = (\boldsymbol{J}^T)^\dagger f_{dz}(\boldsymbol{\tau} - \boldsymbol{k}(\boldsymbol{q})) \tag{2}$$

where $\boldsymbol{h}$ is the end-effector force, $\boldsymbol{\tau}$ is the joint torque, $\boldsymbol{k}(\boldsymbol{q})$ is the estimated gravity compensation torques for the joint configuration $\boldsymbol{q}$, and the function $f_{dz}(\cdot)$ is a deadband filter added to counter the effect of measurement noise. The force $\boldsymbol{h}$ is then mapped into desired velocity of the end-effector $\boldsymbol{v}$ using the standard generalized damper approach [62].

C. Low-Priority Subtask

A position controller is used to complete the low-priority subtask which executes the dressing trajectory such that the user gets properly dressed. The joint velocities $\dot{\boldsymbol{q}}_0$ of the low-priority subtask in equation (1) used to drive the robot at a particular end-effector velocity are calculated using the Jacobian damped-least-squares controller [63]:

$$\dot{\boldsymbol{q}}_0 = \boldsymbol{J}^\dagger \frac{([\boldsymbol{g}_f, \boldsymbol{r}]^T - [\boldsymbol{g}_c, \boldsymbol{r}]^T)}{t_d} \tag{3}$$

where $\boldsymbol{g}_f, \boldsymbol{g}_c$ are the target and current positions of the robot end-effector in 3-D coordinates respectively which are waypoints of the dressing sequence, $\boldsymbol{J}^\dagger$ is the pseudoinverse of $\boldsymbol{J}$, and $\boldsymbol{r}$ is the orientation of the end-effector. The time interval t_d is calculated to maintain a constant robot end-effector speed. In this work, the orientation of the robot gripper is fixed during the whole dressing sequence, and $\boldsymbol{g}_f$ is defined based on the real-time updated dressing trajectory (Section VI).

V. User Modeling in Latent Spaces

Following the method introduced in our previous work [14], we model the high-dimensional and non-linear movements of the users in latent spaces using Gaussian Process Latent Variable Model (GP-LVM) [17]. GP-LVM creates a Gaussian Process (GP) that maps a low-dimensional latent space $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N]^T$ to a high-dimensional dataset represented by $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N]^T$, such that

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

where $\boldsymbol{y}_i \in \mathbb{R}^D$, $\boldsymbol{x}_i \in \mathbb{R}^Q$, D and Q are the dimensionality of the corresponding dataset respectively, N is the number of samples, and f denotes the GP mapping.

In this work, we map the high-dimensional upper-body movements onto a 2-D latent space. The observed dataset for our model is therefore

$$\boldsymbol{Y} \equiv \{\boldsymbol{y}_i = [\boldsymbol{P}_i^h,\ \boldsymbol{P}_i^e,\ \boldsymbol{P}_i^s]\}_{i=1,2,\cdots,N} \tag{4}$$

where $\boldsymbol{P}^h, \boldsymbol{P}^e, \boldsymbol{P}^s$ are the positions of human hand, elbow and shoulder joint in 3-D robot coordinates respectively, which correspond to the 9-dimensional arm posture. We use the standard GP prediction [64] to reconstruct the latent data to the high-dimensional observations. The resulting projections of the 9-D arm posture data onto 2-D latent spaces will be detailed in Section VII-B and visually presented in Fig. 7.

VI. Probabilistic Real-Time Posture Tracking

In this section, we present a probabilistic filter in latent spaces to obtain an accurate estimation of user postures used to update the dressing trajectory in real time.

We formulate the problem of posture tracking in the framework of recursive Bayesian estimation, enabling real-time fusion of two sources of information without vision sensors during the dressing sequence. We use a particle filter to recursively estimate the arm posture given the robot gripper positions (*i.e.*, $\boldsymbol{g}_c$ in Equation (3)) and force information between the gripper and the user (*i.e.*, the estimated end-effector force $\boldsymbol{h}$ in Equation (2)). A preview of the framework is shown in Fig. 4. We define that the current time step t starts when a group of gripper position and force is received, which are published by the Baxter robot at a frequency of 100Hz.

A. Initialization

The initial arm posture $\boldsymbol{y}_s$ is detected by the depth camera. As GP-LVM only provides the link from the latent data to the observed data, in order to locate the initial arm posture in the latent space, we project $\boldsymbol{y}_s$ onto the latent space by finding its nearest neighbor $\boldsymbol{y}^*$ ($\boldsymbol{y}^* \in \boldsymbol{Y}$) from all observed data $\boldsymbol{y}_i$ in equation (4). This is achieved by calculating the Euclidean distance between $\boldsymbol{y}_s$ and all observed data $\boldsymbol{y}_i$ and find $\boldsymbol{y}^*$ that has the smallest distance. Then the corresponding latent point $\boldsymbol{x}^*$ is regarded as the initial arm posture in the latent space.

In this work, we model the estimation of the arm posture at each time step as a Gaussian distribution to capture the uncertainty of the prediction

$$\mathcal{P} = \mathcal{N}(\boldsymbol{\mu}_x^t, \boldsymbol{\Sigma}_x^t)$$

where $\boldsymbol{\mu}_x^t$ and $\boldsymbol{\Sigma}_x^t$ are the mean value and the covariance matrix of the Gaussian distribution respectively. $\boldsymbol{\mu}_x^t$ denotes the arm posture in the latent space in the previous time step. For initialization, $\boldsymbol{\mu}_x^0$ is the initial arm posture $\boldsymbol{x}^*$ in the latent space, and $\boldsymbol{\Sigma}_x^0$ is the noise of depth-camera. After initialization, the definitions of $\boldsymbol{\mu}_x^t$ and $\boldsymbol{\Sigma}_x^t$ are detailed in Section VI-E.

B. Particles Sampling with Random Walk

In order to capture the arm movement dynamics between two time steps, a random walk is used by adding a constant

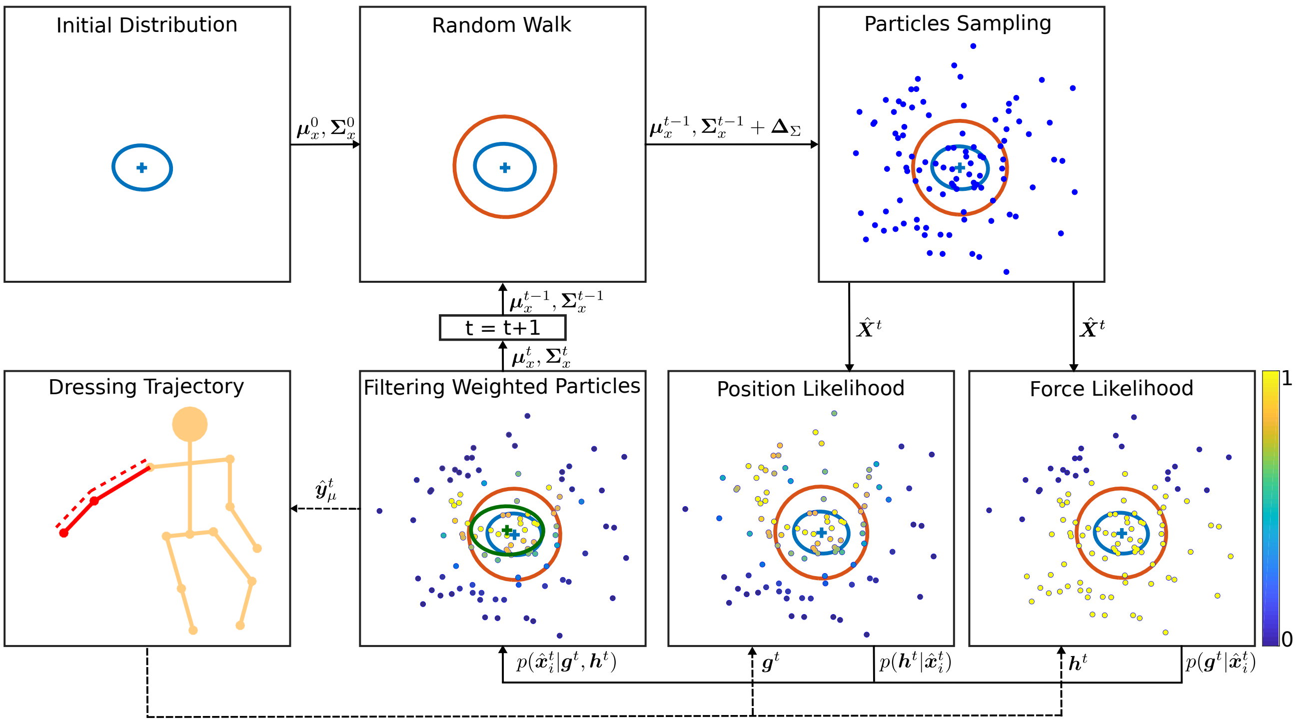


Fig. 4: Proposed framework for probabilistic real-time posture tracking. A particle filter is used for the recursive Bayesian estimation of arm postures, with the real-time fusion of gripper position information and force information. Colors are presented to detail the weights of particles. Warmer colors denote higher weights. For clarity, the latent space is not presented in this figure. Best viewed in color.

diagonal matrix to the covariance matrix of the Gaussian distribution. We then sample particles based on the new Gaussian distribution after the random walk. The sample particles can be presented as

$$\hat{\boldsymbol{X}}^t = [\hat{\boldsymbol{x}}_1^t, \dots, \hat{\boldsymbol{x}}_M^t]^T$$

where M is the number of particles, and t denotes the current time step.

The sampling follows the random walk

$$p(\hat{\boldsymbol{x}}_i^t) = \mathcal{N}(\hat{\boldsymbol{x}}_i^t | \boldsymbol{\mu}_x^{t-1}, \boldsymbol{\Sigma}_x^{t-1} + \boldsymbol{\Delta}_{\boldsymbol{\Sigma}})$$

where $\boldsymbol{\Delta}_{\boldsymbol{\Sigma}}$ is a constant diagonal matrix (presented in Appendix A) that captures the arm movement dynamics between two time steps.

Given all the sampled particles $\hat{\boldsymbol{X}}^t$, we reconstruct their corresponding high-dimensional arm postures $\hat{\boldsymbol{Y}}^t = [\hat{\boldsymbol{y}}_1^t, \dots, \hat{\boldsymbol{y}}_M^t]^T$. In the following three sections, we calculate the likelihood of each reconstructed posture based on the robot end-effector position and force information, to recursively design the Gaussian distribution for the estimation of the arm posture at each time step.

C. Posture Likelihood based on Force Information

In this section, the external force $\boldsymbol{h}$ is used to calculate the likelihood of each reconstructed posture.

We classify the user movements during the dressing sequence as two classes: pushing the gripper and pulling the gripper through the clothes (Fig. 5). In some other dressing scenarios, the user's arm may move without physical interaction with the robot, for instance, moving forward and back in the sleeve. However, based on the preliminary experiments, these movements are usually small enough to have no impact on the dressing trajectory. In our future work, we plan to investigate tracking these small movements without physical interaction between the robot and the user.

We calculate the perpendicular $\hat{\boldsymbol{n}}_i^t$ from the reconstructed posture $\hat{\boldsymbol{y}}_i^t$ to the current gripper position $\boldsymbol{g}^t$. A friction cone $\hat{\boldsymbol{c}}_i^t$ is defined with the central axis $\hat{\boldsymbol{n}}_i^t$, as shown in Fig. 5. The perpendicular intersects either the forearm or the upper-arm of the reconstructed posture depending on which part of the arm is being dressed. In the force model proposed here, the direction of current force $\boldsymbol{h}^t$ should lie in the friction cone, which indicates that the angle δ_i^t between the perpendicular $\hat{\boldsymbol{n}}_i^t$ and the force $\boldsymbol{h}^t$ should be smaller than the predefined cone angle δ_h. Therefore, we manually define the likelihood of the particle $\hat{\boldsymbol{x}}_i^t$ as 1 when the force direction $\boldsymbol{h}^t$ is in the friction cone $\hat{\boldsymbol{c}}_i^t$; the likelihood decreases following a Gaussian function when the force direction is out of the friction cone. Therefore, the likelihood of observing the force $\boldsymbol{h}^t$ given each posture is defined as

$$p(\boldsymbol{h}^t | \hat{\boldsymbol{x}}_i^t) = \begin{cases} 1, & \text{if } \delta_i^t < \delta_h \\ \mathcal{N}(\delta_i^t | \delta_h, \sigma_h^2), & \text{if } \delta_i^t > \delta_h \end{cases}$$

where σ_h^2 is an artificial parameter defined as the variance of the Gaussian function.

D. Posture Likelihood based on Position Information

In this section, the gripper positions are used to calculate the likelihood of each reconstructed posture.

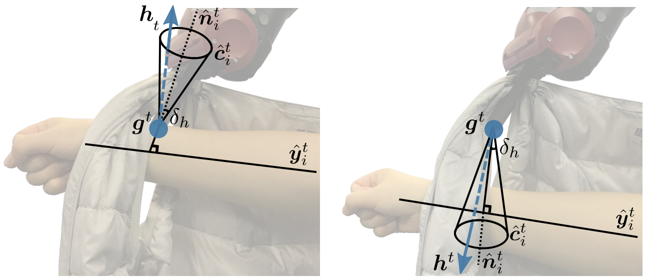


Fig. 5: (Left) The scenario of the user pushing the gripper. (Right) The scenario of the user pulling the gripper through the sleeve opening. The posture likelihood is calculated based on the relationship between the force direction and the friction cone.

We compute the distance between reconstructed arm postures $\hat{\boldsymbol{y}}_i^t$ and the current gripper position $\boldsymbol{g}^t$. Here, each distance d_i^t is computed between $\boldsymbol{g}^t$ and either the forearm or the upper-arm of the reconstructed posture, depending on which part of the arm is being dressed. We assume that the posture likelihood follows a Gaussian distribution of distance. If the user is pushing the gripper, the mean value of the Gaussian distribution is the radius of the human arm d_a. If the user is pulling the gripper through the sleeve opening, the mean value of the Gaussian distribution is the length of the opening d_o. Here we manually define d_o as a constant parameter of the sleeve opening, considering the deformation when the clothes are being pulled. Different d_o are manually designed according to different sleeve openings. Therefore, the likelihood of observing a contact at the gripper position $\boldsymbol{g}^t$ given each posture is defined as

$$p(\boldsymbol{g}^t|\hat{\boldsymbol{x}}_i^t) = \mathcal{N}(d_i^t|\mu_d, \sigma_d^2)$$

where σ_d^2 is an artificial parameter defined as the variance of the Gaussian function, and μ_d can be written as

$$\mu_d = \begin{cases} d_a, & \text{if pushing the gripper} \\ d_o, & \text{if pulling the gripper} \end{cases}$$

E. Weighted Particles Filtering

Having calculated the posture likelihood given position and force information, we propose to fuse these two sources of information. Given the gripper position $\boldsymbol{g}^t$ and the force $\boldsymbol{h}^t$ at the current time step, the particle weights w_i^t are calculated as

$$p(\hat{\boldsymbol{x}}_i^t|\boldsymbol{g}^t, \boldsymbol{h}^t) \propto p(\boldsymbol{g}^t|\hat{\boldsymbol{x}}_i^t) \cdot p(\boldsymbol{h}^t|\hat{\boldsymbol{x}}_i^t) \cdot p(\hat{\boldsymbol{x}}_i^t) = w_i^t \cdot p(\hat{\boldsymbol{x}}_i^t)$$

Then the distribution of particles in the latent space is modeled as a new weighted Gaussian distribution

$$\mathcal{P} = \mathcal{N}(\boldsymbol{\mu}_x^t, \boldsymbol{\Sigma}_x^t)$$

where $\boldsymbol{\mu}_x^t$ is the mean vector, and $\boldsymbol{\Sigma}_x^t$ is the covariance function

$$\boldsymbol{\mu}_x^t = \sum_{i=1}^{M} w_i^t \hat{\boldsymbol{x}}_i^t$$

$$\boldsymbol{\Sigma}_x^t = \boldsymbol{A}\boldsymbol{A}^T$$

$$\boldsymbol{A} = [(w_1^t)^{\frac{1}{2}} \hat{\boldsymbol{x}}_1^t - \boldsymbol{\mu}_x^t, \ldots, (w_M^t)^{\frac{1}{2}} \hat{\boldsymbol{x}}_M^t - \boldsymbol{\mu}_x^t]$$

We reconstruct $\boldsymbol{\mu}_x^t$ to its high-dimensional posture $\hat{\boldsymbol{y}}_\mu^t$, which is regarded as the arm posture at the current time step. $\hat{\boldsymbol{y}}_\mu^t$ is sent to the robot and the dressing trajectory is updated accordingly. The next time step of estimation is carried out with the same method. The values of the parameters used in our method are presented in Appendix A.

VII. Experiments and Results

We conduct three experiments to investigate: 1) how the user modeling captures the specificities of users with different movement limitations (User Modeling Evaluation), 2) the robustness and the accuracy of our proposed real-time posture tracking method (Posture Tracking Evaluation), and 3) the quality of the whole proposed system compared against three experimental baselines (Robot-Assisted Dressing Evaluation). A supplementary video illustrates the performance of the proposed method and how it works with different users.

A. Experimental Setup

We implement the proposed method on the Baxter humanoid robot. An ASUS Xtion Pro RGB-D camera is placed in front of the robot and users. The camera is used for collecting data for user modeling and initial posture recognition. The skeleton of the user is extracted with the OpenNI skeleton tracker. The coordinate transformation between the camera and the robot has been determined prior to the experiments.

We define scenarios in which our users have limitations with their right arm movements and require assistance from the Baxter robot. Fig. 6a-c respectively presents three scenarios of right arm movement limitations considered in this paper: 1) limitations of both elbow and shoulder movements; 2) limitations of elbow movements; 3) limitations of shoulder movements. In order to simulate the movement limitations, our healthy participants wear elbow and shoulder braces which are used to limit their movements.

A total of 20 able-bodied users (ages 23-32, mean: 27, std: 2.29) have participated in these experiments, wearing braces to simulate elbow, shoulder, or elbow-shoulder movement limitations respectively (presented in Appendix B).

For clarity, note that in the Posture Tracking Evaluation (Section VII-C), in order to collect ground-truth data from a motion capture system, a rope is used instead of the clothes to simulate the sleeve opening for all 20 participants. This helps avoid occluding the markers of the motion capture system, as shown in Fig. 6d. In the Robot-Assisted Dressing Evaluation (Section VII-D), the robot provides dressing assistance for all 20 users in putting on an actual sleeveless jacket.

B. User Modeling Evaluation

In the first experiment, we investigate how the user models capture the specificities of users with different movement limitations. Each user wears braces to simulate the three scenarios of right arm movement limitations. In each scenario, the users are asked to move their restricted arm randomly for around 5 minutes within their full reachable spaces while the corresponding arm posture data is recorded as the dataset

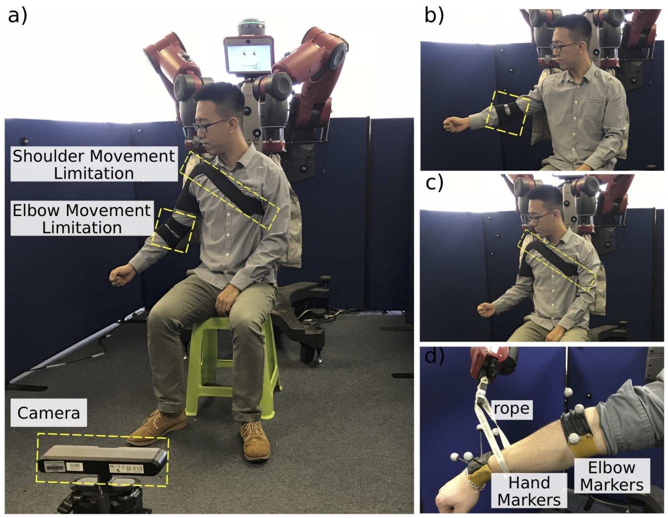


Fig. 6: Experiment environment setup. Healthy users wear braces to simulate elbow, shoulder, or elbow-shoulder movement limitations respectively. The camera is only used for initial posture recognition and user model construction before dressing.

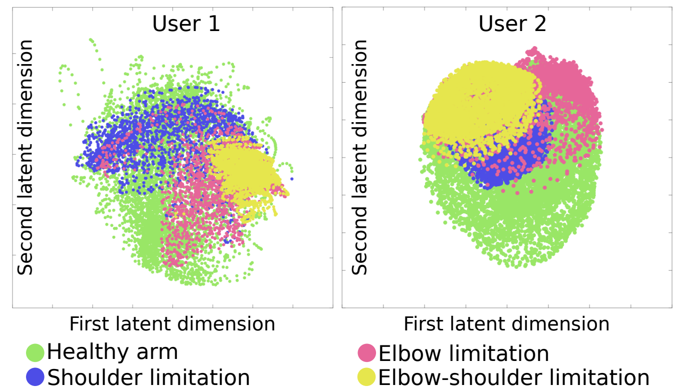


Fig. 7: The 2-D latent representations of the 9-D arm posture data produced with GP-LVM. The different colored subspaces denote the arm movements with different impairments. Best viewed in color.

of GP-LVM shown in equation (4). The movements of the user without limitations (*i.e.*, without the braces) have also been recorded to provide reference data. For the training of GP-LVM models, we set 500 inducing points used for the variational approximation and 200 optimization iterations. We conduct our experiments on a desktop machine with an Intel i7 3.4 GHz processor.

Fig. 7 represents the examples of resulting latent spaces from two users. For the first user, 29565 healthy arm posture data, 29001 elbow-limited arm posture data, 28434 shoulder-limited arm posture data, and 28403 arm posture data when both elbow and shoulder are limited have been recorded. For the second user, 29374, 29136, 29314, and 29358 arm posture data have been recorded respectively.

For each user, we project all the unrestricted-motion data and the restricted-motion data of the listed three scenarios together onto a 2-dimensional latent space with GP-LVM. The different colored subspaces represent the range of arm movements with different limitations. Three immediate observations can be made from this graph: 1) Different arm movement limitations lead to different subspaces; 2) The elbow-limited arm movement subspace (pink), shoulder-limited arm movement subspace (blue), and elbow-shoulder-limited arm movement subspace (yellow) are parts of the healthy arm movement subspace (green) respectively; 3) The elbow-shoulder-limited arm movement subspace (yellow) overlaps with the elbow-limited arm movement subspace (pink) and the shoulder-limited arm movement subspace (blue).

The results demonstrate that the specificities of different users and different movement limitations have been captured by user modeling in latent spaces. Such a personalized model is crucial for impaired users as it captures only the postures that the users can reach, and thus guarantees that only these reachable postures could be estimated in our tracking system. The model intrinsically takes into account a trade-off between the users' full reachability and their comfort zones. Although the users are asked to move their restricted arms within their full reachable spaces, they are likely to remain mainly in their comfort zones where the unexpected movements are also likely to happen during dressing. This is a beneficial effect as it will guide the tracking system to predict postures that are also part of the users' comfort zones.

C. Posture Tracking Evaluation

The second experiment investigates the robustness and the accuracy of our proposed real-time posture tracking method. Each user simulates only one of the three scenarios of movement limitations. The personalized user model for each user is built using the arm posture data of the corresponding impairment collected in Section VII-B.

During the dressing, each user is asked to move his/her arm to perturb the dressing sequence. This action is used to simulate the user executing another action during the dressing sequence. For each user, the experiment is replicated five times, in which the user is asked to execute five different perturbing actions. The five actions are defined as follows: 1) pushing, 2) pulling the sleeve opening, 3) pushing followed by another push, 4) pushing followed by pulling, and 5) pushing twice and pulling in sequence. These actions are artificially simplified versions of everyday motions that can happen during dressing and require combinations of pushing and pulling (*e.g.*, hand-over, answering the phone, adapting arm postures when feeling uncomfortable or tired). The supplementary video illustrates these scenarios of representative motions that a user might take in everyday life.

A motion capture system is used to observe the ground truth of the arm postures. The motion capture system uses eight cameras mounted on the ceiling around the robot. Markers have been attached to the user's hand and elbow. As discussed in Section III, we observe that shoulder movements are usually small enough to have no impact on the dressing trajectory, and thus assume that the shoulder position remains constant during the dressing process. We have investigated human posture tracking in real time with the motion capture system in our previous research, which still yielded failures due to the

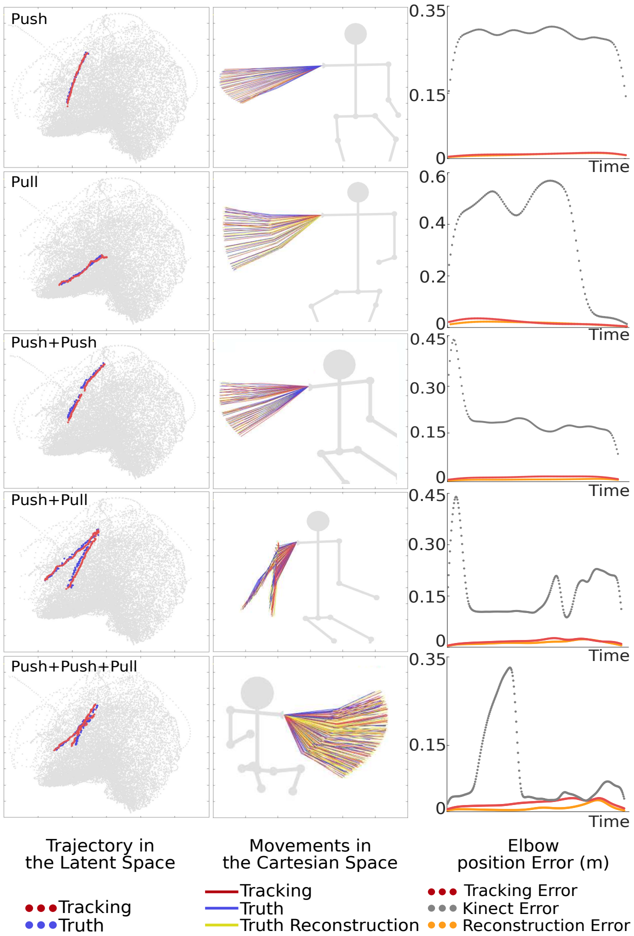


Fig. 8: Results of Posture Tracking Evaluation from one user. The user takes five different movements to perturb the dressing process. (Left) The ground truth and tracking trajectories in the latent space. (Center) The arm movements and the tracking results in the Cartesian space. (Right) The elbow position errors. For clarity, in this figure the errors have been smoothed by Gaussian Convolution Filter [65]. Best viewed in color.

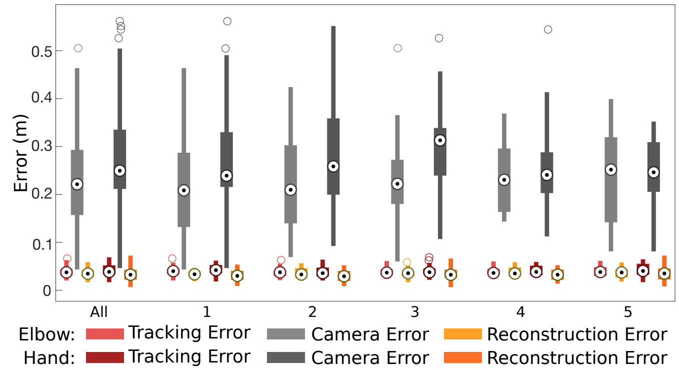


Fig. 9: The boxplot pictures the hand and elbow position errors from all 20 users. The central dot corresponds to the median value of the errors, while the sides of the box refer to the first and third quartiles of the data. The whiskers extend to the most extreme data points that are not considered as outliers, which are displayed as individual circles. The x-axis corresponds to the errors across all trials and the errors of five different actions described in Section VII-C (*i.e.*, the errors of: all actions together (column all); pushing action (column 1); pulling action (column 2); pushing followed by another pushing action (column 3); pushing followed by pulling action (column 4); and pushing twice and pulling in sequence action (column 5)). Best viewed in color.

severe visual occlusions created by the clothes and the robot. Therefore, in this experiment, we use a rope, instead of the real clothes, to simulate the sleeve opening, as shown in Fig. 6d. We also use a depth camera to capture arm postures through the OpenNI skeleton tracker during dressing, and compare the depth sensor observation with our method. We have designed a framework using Robot Operating System (ROS) to integrate all devices for synchronous data recording. The coordinate transformation between the camera and the motion capture system has been determined before the experiments. To compare our method against the truth data, we have also calibrated the arm skeleton lengths obtained from the two vision sensors. Note that here the motion capture system and the markers are only used for collecting truth data, but are not needed for our proposed method. Similarly, the depth camera is only used to collect data for comparison.

We first present the results from one user with five trials, as shown in Fig. 8. The left part of the figure shows the truth trajectory and the tracking trajectory achieved with our method in the latent space. With the same method described in Section VI-A, we project the ground truth arm postures onto the latent space by finding its nearest neighbor. We then reconstruct the ground truth trajectory from the latent space to the high-dimensional arm movements, as shown in Fig. 8-center, together with the ground truth arm movements and the tracking arm movements in the Cartesian space. The elbow position errors at each time step are calculated respectively between the truth data and: 1) the tracking data, 2) the reconstruction data (*i.e.*, the reconstructed movements in Fig. 8-center), and 3) the depth camera data.

Three observations can be made from Fig. 8: 1) The trajectories in the latent space and the arm movements in the Cartesian Space show that our method consistently yields precise arm posture estimates; 2) In Fig. 8-right, the proposed method (red) shows similar errors level to the reconstruction (orange), and much smaller errors than the camera observation (gray). This result illustrates that the error of our method is mostly caused by the GP-LVM projection and the dimensionality reduction. In [12], GP-LVM has already presented a better performance in reconstruction than some other dimensionality reduction methods, like PCA. We hope that through a better Latent Variable Model in the future, the accuracy of our tracking method could be increased; 3) The depth camera observation shows errors that can be as high as 45cm. The experiments show that the camera easily loses track of the human skeleton because of the morphological similarities between the arms of the robot and the human users. The robot arm motions can interfere with the camera tracking, which makes the camera track the moving robot joints instead of the human user arm. It is important here to recall that this experiment does not use real clothes, but a rope to simulate the sleeve opening due to the visual occlusions. The camera shows even worse results when dressing with clothes. Conversely, our proposed approach consistently yields precise estimations in real time.

Fig. 9 illustrates the results of the hand and elbow position errors from all 20 users. We can see from this figure that the same conclusion as the previous figure can be drawn independently from the user: our proposed method consistently yields precise estimations of the user postures in real time. More precisely, the median accuracy of the proposed tracking method is 3.3cm for the hand position and 3.1cm for the elbow position. In the worst case, the tracking error is only of 6.5cm, while the dressing is still successfully completed in this case as it falls within the sleeve opening lengths. The median error of our proposed method is around 6.0 times smaller than the median error of the camera observation (hand: p-value=6.1e-34; elbow: p-value=4.8e-34 when all the trials are considered), and is around 1.2 times larger than the median error of reconstruction (hand: p-value=4.9e-2; elbow: p-value=5.4e-2). This result shows that: 1) the error of our method is mostly caused by the GP-LVM dimensionality reduction; 2) The depth camera easily loses track of human skeleton.

Three observations can be obtained when analyzing the error of our method: 1) The stochastic nature of the particle filter contributes to the error; 2) The error is mostly caused by the GP-LVM dimensionality reduction; 3) In some robot arm configurations, the noise of the joint torque $\boldsymbol{\tau}$ in Equation (2) obtained from the Baxter ROS topic is large, which decreases the accuracy of force likelihood calculation. In order to evaluate the accuracy of the external force estimation, we have compared it with a FTSens 6-axis semiconductor-based force and torque sensor[1] from IIT, installed at the Baxter end-effector, which has a resolution of 0.25N. The Baxter is programmed to execute the same trajectory multiple times carrying a box (1.6kg), while the forces estimated and obtained from the sensor are recorded respectively. The results show that the average value error of the estimated force is approximately 2.8N, while the average angle error is approximately 8.1 degrees. However, in some cases, especially when the robot arm is in vertical configurations, the value error can reach 8.9N and the angle error can reach 40.7 degrees. An external force sensor mounted at the end of the Baxter arm would significantly improve the tracking accuracy in these configurations.

D. Robot-Assisted Dressing Evaluation

In the last experiment, we investigate how the proposed system reduces the force felt by the user during the dressing process. In this experiment, the robot provides dressing assistance in putting on an actual sleeveless jacket for all 20 users. During the dressing, each user is asked to move his/her arm to perturb the dressing sequence. We compare our proposed method against three baselines: 1) Pre-defined dressing trajectory with hierarchical control; 2) Dressing trajectory updated in real-time without hierarchical control; and 3) The approach proposed in our previous research [14].

1) Pre-defined dressing trajectory with hierarchical control (Hierarchical control only): This baseline corresponds to the classic approach of using a pre-defined trajectory, in which the hierarchical multi-task controller is still used to minimize the interaction force. The dressing trajectory is not updated and the gripper constantly executes the initial dressing trajectory (as customary in the literature [10, 18, 19]).
2) Dressing trajectory updated in real-time without hierarchical control (Trajectory update only): The method is equivalent to using only a position controller to execute the dressing trajectory. In this context, the gripper cannot respond to the external force caused by human movements. The mechanical compliance of the Baxter robot allows the physical interaction between the user and the robot, but does not minimize the discomfort of the user. The update of the dressing trajectory is still executed with our proposed tracking method.
3) The approach proposed in our previous research [14]: Hierarchical multi-task control is used to adapt robot motions. The trajectory is updated online based on the density of the latent space, which shows how frequently each upper-body posture is adopted by the user. Instead of accurately tracking user movements, we proposed to find a posture in the latent space that maximizes the density, while being within a certain distance to the gripper. This updated dressing trajectory ensures that the user can reach the clothes without difficulty by taking movement limitations into account. Please refer to [14] for more details.

[1]http://wiki.icub.org/wiki/FT_sensor

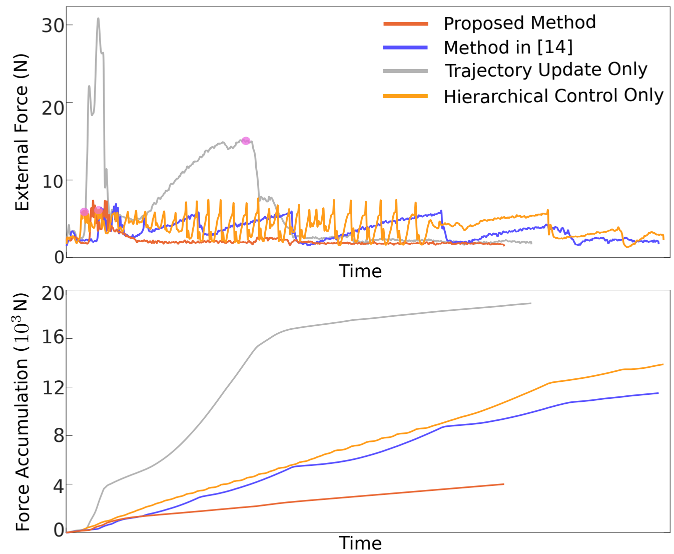


Fig. 10: Results of Robot-Assisted Dressing Evaluation from one user with one trial of each method. (Top) This panel presents the external force recorded at every time step during the dressing sequence. (Bottom) This panel shows the accumulation of the external force over time. Best viewed in color.

For each method and for each user, five replications of trials have been carried out. Overall each user has participated in 20 trials of this experiment. To compare different methods, each user is asked to perform the same action to perturb the dressing sequence during the 20 trials. Then the different actions described in Section VII-B are uniformly distributed among the users.

As shown in Fig. 10, we first present the comparison results from one user with one trial of each method. Fig. 10-top presents the external force $\boldsymbol{h}$ in Equation (2), but without applying the deadband function $f_{dz}(\cdot)$, at every time step. Fig. 10-bottom shows the accumulation of the external force during the dressing sequence. We start to record the force when the gripper starts to move from the hand position. We stop recording when the right arm is dressed. Four conclusions can be drawn from this figure:

1) The approach using the trajectory updated in real-time without hierarchical control shows the largest force (gray). This result is expected and illustrates that hybrid force/position control is required to accomplish the dressing task, while minimizing the force applied between the robot and the user. We can also see that although the dressing trajectory is updated according to the user posture tracking, there is still constant large external force between the robot and the user (*i.e.*, the third peak, marked with a dot). This is because without the hierarchical control, the physical interactions between the user and the robot are allowed by the spring compliance of the Baxter robot. As a consequence, the robot continues its trajectory while maintaining a certain force on the user. This demonstrate that both the hierarchical multi-task controller and the real-time user posture tracking are necessary for dressing assistance;
2) The approach using a pre-defined dressing trajectory with hierarchical control (orange) presents multiple peaks, which means that the robot has collided with the user multiple times. The first two collisions (marked with dots) occur when the user suddenly moves the arm. The other collisions occur when the robot gripper moves to the initial elbow position after the perturbation, while the arm posture has changed. Each time after a collision, the robot adapts its motion to minimize the force based on our hierarchical control method before trying again to move to the initial elbow position (*i.e.*, multiple peaks in Fig. 10-top). Eventually, the gripper will reach the initial elbow position after slipping from the user arm;
3) The approach in [14] (blue) attempts to find an updated elbow position that the user can reach without difficulty, but not to accurately track arm movements. Thus, compared to the initial arm posture, the updated trajectory is closer to the current arm posture, which still leads to multiple but fewer collisions and smaller force accumulation.
4) Conversely, the approach proposed in this paper (red) shows the smallest external force compared to the three other approaches and no multiple collisions. This demonstrates that our approach accurately updates the dressing trajectory in real time.

Fig. 11 compares the results from all 20 users. Fig. 11-top shows the external force accumulation over the dressing sequence. To counter the effect of measurement noise, we consider the data after using the deadband filter of the function $f_{dz}(\cdot)$ in Equation (2), filtering all the noise below 4N. In

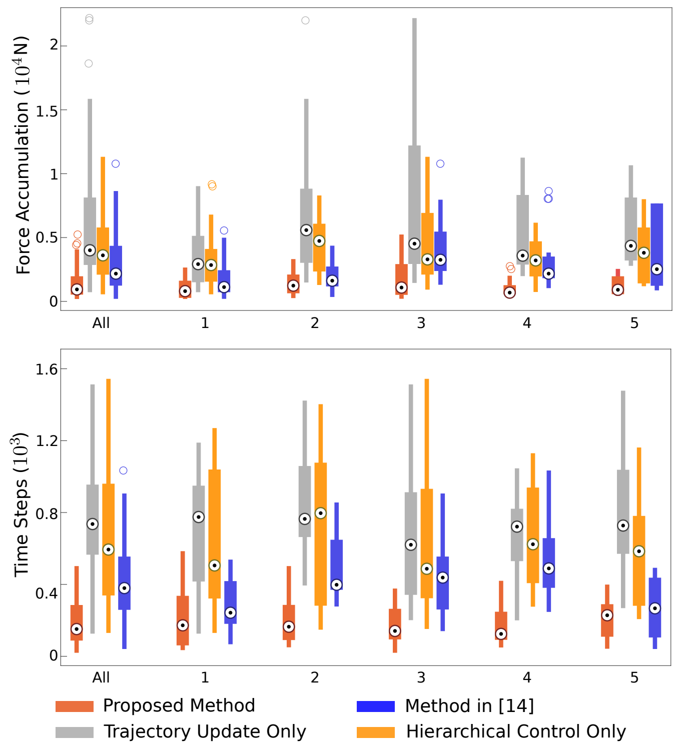


Fig. 11: The boxplots picture the data from all twenty users. (Top) This boxplot represents the external force accumulation over the dressing sequence (when the force is larger than 4N). (Bottom) This boxplot represents the number of time steps during which the perceived force is larger than 4N. The x-axis corresponds to the errors across all trials and the errors of five different actions described in Section VII-C. Best viewed in color.

our preliminary experiments, we found that the parameter 4N is robust to noise in almost all time steps, while users still feel comfortable. We calculate the accumulation of the external force, when it is larger than 4N in each trial and report the results in the boxplot. Fig. 11-bottom reports the number of time steps in which the force is larger than 4N in each trial. We can see from this figure that the same conclusion as the previous figure can be drawn: our proposed approach minimizes the force between the robot and the user, and avoids multiple collisions. This shows that the tracking method has successfully inferred the postures of the user in real time, even when the perturbing actions are particularly complicated. More precisely, the accumulation external force of our proposed method is around 4.3 times smaller than the results of the method without hierarchical control (p-value=3.7e-28 when all the trials are considered). Similar results have been presented when comparing our methods with the method without trajectory update (3.8 times smaller, p-value=1.8e-21), and the method in [14] (2.3 times smaller, p-value=3.7e-13). With respect to the number of the time steps in which a force larger than 4N is perceived, the proposed method also significantly outperforms the other approaches (7.6 times smaller, p-value=1.1e-23 when compared to the variant without hierarchical control; 4.7 times smaller, p-value=2.3e-18 when compared to the variant without trajectory update; and 3.0 times smaller, p-value=1.2e-9 when compared to the method in [14]).

We also note some variations between different users and methods. For instance, Fig. 11-column 1 (*i.e.*, the pushing action) shows the smallest difference between our proposed method and the method in our previous work [14]. This method has yielded some good results when the user performs simple movements like pushing, which is the same movement we have asked users to do in our previous experiments in [14]. However, our current method outperforms the previous method when users perform the other four movements which are more complicated. This is because the density information cannot effectively reflect the complexity of the arm movements. The method in [14] attempts to find an elbow position that the user can reach without difficulty, but not to track the arm movements in real time. Thus when the user has moved his/her arm to the area where the density is low and the posture is not frequently adopted, our previous method fails to update the trajectory based on the density information.

Fig. 11-Column 5 (*i.e.*, the pushing, pushing and pulling action) shows similar number of time steps in which the force is larger than 4N between our proposed method and the method in [14], but smaller force accumulation with the method in this work. This is because the method in [14] leads to several collisions between the robot and the user. Conversely, Fig. 11-column 2 (*i.e.*, the pulling action) shows a similar force accumulation but fewer time steps with the method in this work compared to [14]. Although all users are told to hold their arm postures after movements, some users slightly move their arms when collisions occur, which leads to a small force accumulation with the method in [14]. This could be explained by different confidence levels on the robotic setup from the users. However, we can observe that even in this case, our method still achieves a smaller force accumulation against other three baselines as presented in Fig. 11-column 2.

Overall, the above statistics and observations quantitatively and qualitatively demonstrate the robustness and accuracy of our tracking method under challenging conditions, like severe visual occlusions and varying user movements. The evaluation of the whole system shows that it is capable of generalizing to users with different movement limitations while maintaining better performance against other methods.

VIII. Conclusions and Future Work

In this paper, we focus on the problem of real-time user posture tracking for personalized robot-assisted dressing. The severe visual occlusions created by the robot and the clothes prevent vision sensors, like cameras and motion capture systems, from accurately extracting human skeletons during the dressing sequence.

In order to address this problem, in this study, we have proposed a probabilistic filtering method using Bayesian networks in latent spaces, which fuses multi-modal information, to yield precise and real-time estimate of user postures. Based on our proposed method, the depth camera is only needed for initial posture recognition and user model construction before dressing. In other words, this method enables camera-less user posture tracking during the dressing process. The resulting robot-assisted dressing system is able to adapt its motions in accordance with unexpected user movements during dressing and improve the user's comfort by minimize the force he/she may feel.

The formulation of the proposed user posture tracking method is general and allows to accommodate different modalities. We have demonstrated our tracking method with the Baxter robot, but since no *a priori* knowledge has been assumed on the kinematic and dynamic models of the robot for our proposed tracking method, the method can be applied to different robotic platforms that have torque sensing capabilities. Another strength of our user posture tracking approach lies in modeling the force likelihood and the position likelihood in an intuitive way that is realistic enough to achieve high precision, while keeping tractability in mind. Thus the approach has the potential to be generalized to other human-robot physical interaction applications where close contact between the robot and human users occur, when vision sensors are limited in tracking human postures.

In addition to this contribution, we also build user models, which provide latent spaces for tracking. Data collection for user model construction before the dressing is crucial for tracking, as a low amount and diversity of the collected posture data would decrease the tracking accuracy. Instead of projecting the data from all 20 users onto the same latent space, we build personalized user models, as such models are capable of capturing the specificities of different upper-body impairments, taking into account a trade-off between the users' movement capabilities and their preferences. Thus such personalized models guide the tracking system to predict postures that are also part of the users' reachable or comfort zones. The personalized models also automatically preserve all the user morphological properties in the latent representation. Typically, a generic model does not have these types of information, and thus requires personalization to take into account the differences of each user.

Future extensions of the presented robot-assisted dressing system include assisting users to dress with more complicated clothes, such as long-sleeved shirts or polos. While this paper focuses on sleeveless jackets, our preliminary experiments have already presented promising results when assisting users to dress with a long-sleeved jacket. A more complex force model could be implemented to address some challenging problems, like sleeves getting caught on users' arms. Also, the paper assumes that the user's hand is inside the sleeve opening, while the problems of hand insertion will be considered in future works. In this case, force measurements may provide no meaningful information if the user's arm is in the middle of the sleeve opening, while no forces being applied to the user's arm. Adding additional inputs to the system, such as point cloud information with target segmentation (*e.g.*, Mask R-CNN [66]) or capacitive proximity sensing [38] that can directly capture the human body, could be investigated.

The personalized user modeling in latent spaces works well in our experiments. An interesting question for future work is whether other approaches (*e.g.*, knowledge transfer) could rapidly personalize a generic model to capture the upper-body impairments, the user preferences and the morphological properties of a new user. This would significantly decrease the duration of the data collection, which lasts 5 minutes per user in our current experiments and might be uncomfortable for some users.

We also plan to validate if a more accurate force sensor, mounted at the end of the Baxter robot arm, would improve the posture tracking accuracy. A more complex model, taking into account the force information and human arm dynamics, to automatically adjust the random walk instead of a constant diagonal matrix, could also be an interesting future extension.

Appendix A
Parameter Values

This appendix presents all parameter values used in our proposed method.

TABLE I. Parameter values.

Parameter	Definition	Unit	Value
D	Damping constant	N	0.2
$\Delta_{\mathbf{\Sigma}}$	Constant diagonal matrix	–	$[0.01, 0; 0, 0.01]$
δ_h	Cone Angle	Degree	20
σ_h^2	Variance	–	0.01
d_a	Arm radius*	cm	$[2, 4]$
d_o	Sleeve opening length	cm	12
σ_d^2	Variance	–	10

* The arm radius is set approximately depending on different users. In our experiments, the radius ranges from 2cm to 4cm.

Appendix B
Experiment Participants

This appendix presents all the experiment participants, as shown in Fig. 12. A total of 20 able-bodied users (ages 23-32, mean: 27, std: 2.29) have participated in these experiments, wearing braces to simulate elbow, shoulder, or elbow-shoulder movement limitations respectively.

Fig. 12: A total of 20 able-bodied users (ages 23-32, mean: 27, std: 2.29) have participated in our experiments, simulating elbow, shoulder, or elbow-shoulder movement limitations respectively.

References

[1] T. L. Mitzner, T. L. Chen, C. C. Kemp, and W. A. Rogers, "Identifying the potential for robotics to assist older adults in different living environments," *International journal of social robotics*, vol. 6, no. 2, pp. 213–227, 2014.

[2] H. Robinson, B. MacDonald, and E. Broadbent, "The role of healthcare robots for older people at home: A review," *International Journal of Social Robotics*, vol. 6, no. 4, pp. 575–591, 2014.

[3] B. J. Dudgeon, J. M. Hoffman, M. A. Ciol, A. Shumway-Cook, K. M. Yorkston, and L. Chan, "Managing activity difficulties at home: a survey of medicare beneficiaries," *Archives of physical medicine and rehabilitation*, vol. 89, no. 7, pp. 1256–1261, 2008.

[4] A. Colomé, A. Planells, and C. Torras, "A friction-model-based framework for reinforcement learning of robotic tasks in non-rigid environments," in *IEEE International Conference on Robotics and Automation*. IEEE, 2015, pp. 5649–5654.

[5] T. Tamei, T. Matsubara, A. Rai, and T. Shibata, "Reinforcement learning of clothing assistance with a dual-arm robot," in *Humanoid Robots, 11th IEEE-RAS International Conference on*. IEEE, 2011, pp. 733–738.

[6] T. Matsubara, D. Shinohara, and M. Kidode, "Reinforcement learning of a motor skill for wearing a t-shirt using topology coordinates," *Advanced Robotics*, vol. 27, no. 7, pp. 513–524, 2013.

[7] N. Koganti, T. Tamei, T. Matsubara, and T. Shibata, "Estimation of human cloth topological relationship using depth sensor for robotic clothing assistance," in *Proceedings of Conference on Advances In Robotics*. ACM, 2013, pp. 1–6.

[8] N. Koganti, J. G. Ngeo, T. Tomoya, K. Ikeda, and T. Shibata, "Cloth dynamics modeling in latent spaces and its application to robotic clothing assistance," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*. IEEE, 2015, pp. 3464–3469.

[9] W. Yu, A. Kapusta, J. Tan, C. C. Kemp, G. Turk, and C. K. Liu, "Haptic simulation for robot-assisted dressing," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 6044–6051.

[10] Y. Gao, H. J. Chang, and Y. Demiris, "Iterative path optimisation for personalised dressing assistance using vision and force information," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*. IEEE, 2016, pp. 4398–4403.

[11] E. Pignat and S. Calinon, "Learning adaptive dressing assistance from human demonstration," *Robotics and Autonomous Systems*, vol. 93, pp. 61–75, 2017.

[12] N. Koganti, T. Tamei, K. Ikeda, and T. Shibata, "Bayesian nonparametric learning of cloth models for real-time state estimation," *IEEE Transactions on Robotics*, 2017.

[13] Y. Gao, H. J. Chang, and Y. Demiris, "User modelling for personalised dressing assistance by humanoid robots," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*. IEEE, 2015, pp. 1840–1845.

[14] F. Zhang, A. Cully, and Y. Demiris, "Personalized robot-assisted dressing using user modeling in latent spaces," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*. IEEE, 2017, pp. 3610–3603.

[15] J. Martinez-del Rincon, M. Lewandowski, J.-C. Nebel, and D. Makris, "Generalized laplacian eigenmaps for modeling and tracking human motions," *IEEE transactions on cybernetics*, vol. 44, no. 9, pp. 1646–1660, 2014.

[16] J. M. Wang, D. J. Fleet, and A. Hertzmann, "Gaussian process dynamical models for human motion," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, no. 2, pp. 283–298, 2008.

[17] N. D. Lawrence, "Gaussian process latent variable models for visualisation of high dimensional data," *Advances in neural information processing systems*, vol. 16, no. 3, pp. 329–336, 2004.

[18] G. Chance, A. Camilleri, B. Winstone, P. Caleb-Solly, and S. Dogramadzi, "An assistive robot to support dressing-strategies for planning and error handling," in *Biomedical Robotics and Biomechatronics, 6th IEEE International Conference on*. IEEE, 2016, pp. 774–780.

[19] A. Clegg, W. Yu, Z. Erickson, J. Tan, C. K. Liu, and G. Turk, "Learning to navigate cloth using haptics," in *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*. IEEE, 2017, pp. 2799–2805.

[20] L. Twardon and H. Ritter, "Active boundary component models for robotic dressing assistance," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2016, pp. 2811–2818.

[21] N. Koganti, T. Tamei, T. Matsubara, and T. Shibata, "Real-time estimation of human-cloth topological relationship using depth sensor for robotic clothing assistance," in *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*. IEEE, 2014, pp. 124–129.

[22] Z. Erickson, A. Clegg, W. Yu, G. Turk, C. K. Liu, and C. C. Kemp, "What does the person feel? learning to infer applied forces during robot-assisted dressing," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 6058–6065.

[23] G. Chance, A. Jevtić, P. Caleb-Solly, and S. Dogramadzi, "a quantitative analysis of dressing dynamics for robotic dressing assistance," *Frontiers in Robotics and AI*, vol. 4, 2017.

[24] A. Kapusta, W. Yu, T. Bhattacharjee, C. K. Liu, G. Turk, and C. C. Kemp, "Data-driven haptic perception for robot-assisted dressing," in *Robot and Human Interactive Communication, 25th IEEE International Symposium on*. IEEE, 2016, pp. 451–458.

[25] Z. Erickson, H. M. Clever, G. Turk, C. K. Liu, and C. C. Kemp, "Deep haptic model predictive control for robot-assisted dressing," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–8.

[26] A. Calanca, R. Muradore, and P. Fiorini, "A review of algorithms for compliant control of stiff and fixed-compliance robots," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 2, pp. 613–624, 2016.

[27] F. Ficuciello, L. Villani, and B. Siciliano, "Variable impedance control of redundant manipulators for intuitive human–robot physical interaction," *IEEE Transactions on Robotics*, vol. 31, no. 4, pp. 850–863, 2015.

[28] H.-C. Lin, J. Smith, K. K. Babarahmati, N. Dehio, and M. Mistry, "A projected inverse dynamics approach for multi-arm cartesian impedance control," in *IEEE International Conference on Robotics and Automation*, 2018.

[29] S. D. Klee, B. Q. Ferreira, R. Silva, J. P. Costeira, F. S. Melo, and M. Veloso, "Personalized assistance for dressing users," in *International Conference on Social Robotics*. Springer, 2015, pp. 359–369.

[30] K. Yamazaki, R. Oya, K. Nagahama, and M. Inaba, "A method of state recognition of dressing clothes based on dynamic state matching," in *System Integration (SII), 2013 IEEE/SICE International Symposium on*. IEEE, 2013, pp. 406–411.

[31] K. Yamazaki, R. Oya, K. Nagahama, K. Okada, and M. Inaba, "Bottom dressing by a life-sized humanoid robot provided failure detection and recovery functions," in *System Integration (SII), 2014 IEEE/SICE International Symposium on*. IEEE, 2014, pp. 564–570.

[32] A. Toshev and C. Szegedy, "Deeppose: Human pose estimation via deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1653–1660.

[33] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," *arXiv preprint arXiv:1611.08050*, 2016.

[34] R. A. Güler, N. Neverova, and I. Kokkinos, "Densepose: Dense human pose estimation in the wild," *arXiv preprint arXiv:1802.00434*, 2018.

[35] M. Wang, X. Chen, W. Liu, C. Qian, L. Lin, and L. Ma, "Drpose3d: Depth ranking in 3d human pose estimation," *arXiv preprint arXiv:1805.08973*, 2018.

[36] U. Iqbal, A. Doering, H. Yasin, B. Krüger, A. Weber, and J. Gall, "A dual-source approach for 3d human pose estimation from single images," *Computer Vision and Image Understanding*, 2018.

[37] G. Chance, A. Jevtić, P. Caleb-Solly, G. Alenyà, C. Torras, and S. Dogramadzi, "elbows outpredictive tracking of partially occluded pose for robot-assisted dressing," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3598–3605, 2018.

[38] Z. Erickson, M. Collier, A. Kapusta, and C. C. Kemp, "Tracking human pose during robot-assisted dressing using single-axis capacitive proximity sensing," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2245–2252, 2018.

[39] C. G. Cifuentes, J. Issac, M. Wüthrich, S. Schaal, and J. Bohg, "Probabilistic articulated real-time tracking for robot manipulation," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 577–584, 2017.

[40] G. Izatt, G. Mirano, E. Adelson, and R. Tedrake, "Tracking objects with point clouds from vision and touch," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 4000–4007.

[41] J. Liu, D. Liu, J. Dauwels, and H. S. Seah, "3d human motion tracking by exemplar-based conditional particle filter," *Signal Processing*, vol. 110, pp. 164–177, 2015.

[42] W. Zhang, L. Shang, and A. B. Chan, "A robust likelihood function for 3d human pose tracking," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5374–5389, 2014.

[43] S. Sedai, M. Bennamoun, and D. Q. Huynh, "A gaussian process guided particle filter for tracking 3d human pose in video," *IEEE Transactions on Image Processing*, vol. 22, no. 11, pp. 4286–4300, 2013.

[44] I.-C. Chang and S.-Y. Lin, "3d human motion tracking based on a progressive particle filter," *Pattern Recognition*, vol. 43, no. 10, pp. 3621–3635, 2010.

[45] G. Canal, E. Pignat, G. Alenyà, S. Calinon, and C. Torras, "Joining high-level symbolic planning with low-level motion primitives in adaptive hri: application to dressing assistance," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–9.

[46] M. Fiore, A. Clodic, and R. Alami, "On planning and task achievement modalities for human-robot collaboration," in *Experimental Robotics*. Springer, 2016, pp. 293–306.

[47] G. Canal, G. Alenyà, and C. Torras, "Adapting robot task planning to user preferences: an assistive shoe dressing example," *Autonomous Robots*, pp. 1–14, 2018.

[48] A. Jevtić, A. F. Valle, G. Alenyà, G. Chance, P. Caleb-Solly, S. Dogramadzi, and C. Torras, "Personalized robot assistant for support in dressing," *IEEE Transactions on Cognitive and Developmental Systems*, 2018.

[49] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, M. Cakmak, and V. Micelli, "Toward seamless human-robot handovers," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 112–132, 2013.

[50] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and S. Kiesler, "Human preferences for robot-human hand-over configurations," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 1986–1993.

[51] F. C. Huang and J. L. Patton, "Movement distributions of stroke survivors exhibit distinct patterns that evolve with training," *Journal of neuroengineering and rehabilitation*, vol. 13, no. 1, p. 1, 2016.

[52] Z. A. Wright, J. L. Patton, F. C. Huang, and E. Lazzaro, "Evaluation of force field training customized according to individual movement deficit patterns," in *Rehabilitation Robotics, IEEE International Conference on*. IEEE, 2015, pp. 193–198.

[53] Z. A. Wright, E. Lazzaro, K. O. Thielbar, J. L. Patton, and F. C. Huang, "Robot training with vector fields based on stroke survivors individual movement statistics," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 2, pp. 307–323, 2018.

[54] H. Wang, L. Wang, Y. Xiang, N. Zhao, X. Li, S. Chen, C. Lin, and G. Li, "Assessment of elbow spasticity with surface electromyography and mechanomyography based on support vector machine," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*. IEEE, 2017, pp. 3860–3863.

[55] M. Wüthrich, J. Bohg, D. Kappler, C. Pfreundt, and S. Schaal, "The coordinate particle filter-a novel particle filter for high dimensional systems," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2454–2461.

[56] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," in *International Conference on Artificial Neural Networks*. Springer, 1997, pp. 583–588.

[57] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.

[58] R. Urtasun, D. J. Fleet, and P. Fua, "3d people tracking with gaussian process dynamical models," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1. IEEE, 2006, pp. 238–245.

[59] J. M. Wang, D. J. Fleet, and A. Hertzmann, "Gaussian process dynamical models," in *NIPS*, vol. 18, 2005, p. 3.

[60] D. Kruse, R. J. Radke, and J. T. Wen, "Collaborative human-robot manipulation of highly deformable materials," in *Robotics and Automation, IEEE International Conference on*. IEEE, 2015, pp. 3782–3787.

[61] S. Chiaverini, G. Oriolo, and I. D. Walker, "Kinematically redundant manipulators," in *Springer handbook of robotics*. Springer, 2008, pp. 245–268.

[62] M. T. Mason, "Compliance and force control for computer controlled manipulators," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 11, no. 6, pp. 418–432, 1981.

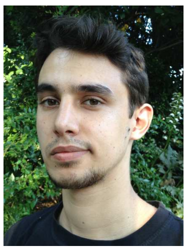
[63] S. R. Buss, "Introduction to inverse kinematics with jacobian transpose, pseudoinverse and damped least squares methods," *IEEE Journal of Robotics and Automation*, vol. 17, no. 1-19, p. 16, 2004.

[64] M. Titsias and N. D. Lawrence, "Bayesian gaussian process latent variable model," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 2010, pp. 844–851.

[65] J. Oksanen and T. Sarjakoski, "Error propagation analysis of dem-based drainage basin delineation," *International Journal of Remote Sensing*, vol. 26, no. 14, pp. 3085–3102, 2005.

[66] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Computer Vision (ICCV), 2017 IEEE International Conference on*. IEEE, 2017, pp. 2980–2988.

Fan Zhang received the B.Sc. degree in mechanical engineering and the M.Sc. degree in mechatronics engineering from Harbin Institute of Technology, China, in 2014 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Personal Robotics Laboratory, Electrical and Electronic Department, Imperial College London, U.K., under the supervision of Prof. Y. Demiris. His research interests involve learning algorithms applied to human-robot interaction. He received the 2016 IEEE International Conference on Mechatronics and Automation Best Student Paper Award.



Antoine Cully received the M.Sc. and the Ph.D. degrees in robotics and artificial intelligence from the Pierre et Marie Curie University of Paris, France, in 2012 and 2015, respectively, and the engineer degree from the School of Engineering PolytechParis UPMC, in 2012. His Ph.D. dissertation has received several Best-Thesis awards. Between 2015 and 2018, he was a postdoctoral researcher in the Personal Robotics Laboratory at Imperial College London. He is currently a Lecturer at the Department of Computing at Imperial College London. He has published several journal papers in prestigious journals including Nature, IEEE Transaction in Evolutionary Computation, and the International Journal of Robotics Research. His current research interests involve learning algorithms for autonomous robots, artificial cognitive systems, evolutionary robotics, with particular focus on learning methods for autonomous adaptation.

Yiannis Demiris (SM'03) received the B.Sc. (Hons.) and Ph.D. degrees from the Department of Artificial Intelligence, University of Edinburgh, U.K.

He is a Professor with the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K., where he is the Royal Academy of Engineering Chair In Emerging Technologies, and heads the Personal Robotics Laboratory. His current research interests include human-robot interaction, machine learning, user modeling, and assistive robotics. He has published over 150 journal and peer reviewed conference papers in the above areas. Prof. Demiris was a recipient of the Rector's Award for Teaching Excellence in 2012 and the FoE Award for Excellence in Engineering Education in 2012. He is a Fellow of IET, BCS, and Royal Statistical Society.