Preprint version; final version available at ieeexplore.ieee.org
IROS (2015), pp: 1840-1845
Published by: IEEE

User Modelling for Personalised Dressing Assistance by Humanoid Robots

Yixing Gao, Hyung Jin Chang, Yiannis Demiris

***Abstract*— Assistive robots can improve the well-being of disabled or frail human users by reducing the burden that activities of daily living impose on them. To enable personalised assistance, such robots benefit from building a user-specific model, so that the assistance is customised to the particular set of user abilities. In this paper, we present an end-to-end approach for home-environment assistive humanoid robots to provide personalised assistance through a dressing application for users who have upper-body movement limitations. We use randomised decision forests to estimate the upper-body pose of users captured by a top-view depth camera, and model the movement space of upper-body joints using Gaussian mixture models. The movement space of each upper-body joint consists of regions with different reaching capabilities. We propose a method which is based on real-time upper-body pose and user models to plan robot motions for assistive dressing. We validate each part of our approach and test the whole system, allowing a Baxter humanoid robot to assist human to wear a sleeveless jacket.**

I. Introduction

Assistive robots in home environments have gained significant popularity, not only because of the increasingly sophisticated manufacturing of robots and the rapid developments in artificial intelligence, but also due to a huge potential to reduce the need for human labour in daily care, especially considering the ageing problem [1], [2], [3], [4], [5]. However, people vary significantly in their skill sets, culture, habits, behaviours *etc.* and these factors affect their choices and preferences in human-robot interactions. For a widespread use of home-environment assistive robots in the future, the ability to provide personalised assistance has become one of the key issues for such robots.

Although recent studies have enabled assistive robots to perform some daily tasks in home environments, for instance, cleaning rooms [6] and cooking [7], assistive dressing remains a challenging task for robots [8], [9] and very little effort has been put on this particular problem. Through reinforcement learning, a dual-arm robot managed to wear a T-shirt for a soft human mannequin with fixed poses [10], [11]. However, this method might be difficult to generalise when faced with complicated dynamic human poses. Particularly since the reinforcement learning component requires multiple trials and errors, which might put the user at risk. Thus, we are encouraged by the thought that such problems would be solved better through studying user preferences and building user models.

The authors are with the Personal Robotics Lab, Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom {y.gao12,hj.chang,y.demiris}@imperial.ac.uk

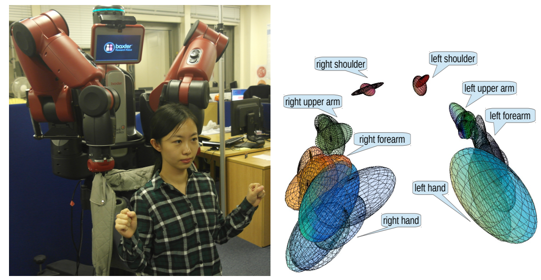


Fig. 1: The proposed approach enables a Baxter humanoid robot to provide personalised assistance according to user models and real-time upper-body pose to assist human to dress. User models are represented as a mixture of Gaussian for the movement space of each upper-body joint.

Some methods which learn human preferences in human-robot interactions focus on object hand-over scenarios. In [12] and [13], through user studies, they aim to analyse robot hand-over configurations which users preferred most. However, poses of users, which directly affects robot motions, are not considered in these methods. Especially for users with upper body movement limitations or elderly people, their reaching abilities and pose preferences require more attention. Furthermore, due to limitations of current methods, home-environment assistive robots are not able to accomplish more complicated tasks, such as providing personalised dressing assistance.

In this paper, we present an end-to-end approach to model the movement space of human upper-body joints and enable a humanoid robot to provide personalised dressing assistance. Human upper-body pose is recognised with a top-view depth sensor using randomised decision forests [14]. To model the movement space of each upper-body joint, an unsupervised expectation-minimisation (EM) algorithm [15] is used to learn Gaussian mixture models (GMMs). We first test the parts of upper-body pose recognition and user modelling methods with four different user motion datasets, then we test the whole system on two healthy users by enabling a Baxter robot to provide user-specific assistance for them to put on a sleeveless jacket separately. Robot assistive motions are planned according to real-time upper-body pose and user modelling knowledge.

II. Related Work

For human-robot interactions, Fong *et al.* [16] argue that user modelling is useful for socially interactive robots to adapt their behaviours to accommodate to different users. Various methods have been proposed to build user models,

Preprint version; final version available at ieeexplore.ieee.org
IROS (2015), pp: 1840-1845
Published by: IEEE

for instance, a stereotype approach is used in [17] to model users in a dialogue interaction by defining three subgroups of user stereotypes with different attributes. Tapus and Mataric [18] investigate the role of user personalities with a hands-off assistive robot during post-stroke rehabilitation therapy. For a companion robot approaching a seated person in a helping context, Dautenhahn *et al.* [19] study user preferences for comfortable approach directions by considering factors such as gender differences, age and handedness. Apart from robot approaching, Cakmak *et al.* [12] present a user study on human preferences of robot hand-over configurations, using a simulated kinematics model of human to collect information on user preferences. Besides, spatial reasoning of users, such as user visibility and arm comfort, is considered as an important factor in [13] for object hand-over tasks. However, the above methods cannot learn human preferences in more complicated scenarios, such as assistive dressing.

Existing methods on assistive dressing focus more on robot trajectory learning and lower the impact of user poses and comfortableness. For instance, a dual-arm robot learned how to dress a soft human mannequin with a T-shirt through reinforcement learning, using the topological relationships between the T-shirt and mannequin [10]. Instead of using a motion capturing system to detect the hem of the T-shirt, Koganti *et al.* [11] used a depth sensor to detect the coloured hem and estimate human-cloth topological relationship. However, in both studies it is assumed that the arms of the mannequin are inside the sleeves of the T-shirt while the hem of the T-shirt is held by the robot, which simplifies the whole dressing process. Furthermore, with the method of reinforcement learning, there exists a potential safety issue for real users which is not solved in [10] and [11].

Human pose recognition plays an important role in many human-robot interactions. In [10], markers, which are attached on both the mannequin and T-shirt, are detected by a motion capture system. While some research uses fiducial markers [20], colour information or motion capture systems [21] to recognise human pose, some develop their own algorithms which can only be used for specific experimental settings [22]. There have been important recent advances in real-time human pose estimation using a single depth camera [14], which can be used during human robot interactions as well. Experiments have shown that the method in [14] performs extremely well in real time tests. Luo *et al.* uses a similar method to recognise human back pose for robot massage applications [23]. However, it is worth noting that this study is concerned with a static non-articulated object (i.e. user backs).

III. Proposed Approach

In this paper, we propose an end-to-end approach for home-environment assistive humanoid robots to personally assist users with upper body limitations to dress by modelling the movement space of upper-body joints. To recognise the human upper-body pose, we adopt one of the best performing whole-body pose estimation methods [14] and adapt it to a top-view depth camera attached to a humanoid robot. In section III-A, we discuss how the randomised decision forests algorithm is applied to estimate human upper-body pose from a single depth image. Modelling the movement space of upper-body joints using GMMs is shown in section III-B. In section III-C, we show how to plan robot motions for personalised assistive dressing.

A. Real-time human upper-body pose recognition

Considering the dressing need, only the human upper-body pose is concerned. In this paper, we define eight body parts for an upper body which are left/right (L/R) shoulder, L/R upper arm, L/R forearm and L/R hand. We omit the human head since we are more concerned with arm movements.

First of all, labelled training images for a user should be collected. Apart from moving the human back, a user is allowed to move both arms freely within the working range of the camera while pair-wise pixel-aligned RGB and depth images are recorded. The user wears clothing with eight different colours on the upper body during training data collection, which facilitates the following segmentation process using colour information.

After collecting RGB and depth images for training, we extract foreground pixels by filtering the head and the background, which are shown in Fig. 2b and Fig. 2c. To generate labels for every pixel remaining in RGB images, we first calculate sample colours in the L*a*b* colour space for each piece of cloth, where L* is the luminosity layer, a* and b* are chromaticity layers. Then, we classify each pixel using the nearest neighbour rule. Due to the noise caused by light, the classification result of every image is further improved by image erosion and image dilation. The segmentation result of one depth image is shown in Fig. 2d. Because every pixel in a RGB image is aligned to a pixel in its corresponding depth image, the labels generated from RGB images can be used for labelling depth images.

A randomised decision forest [14] is a multitude of decision trees which consist of split and leaf nodes. A split node can be seen as a weak classifier which contains the information of a selected feature and a threshold while a leaf node contains the information of the probabilities that this leaf node belongs to a certain body part.

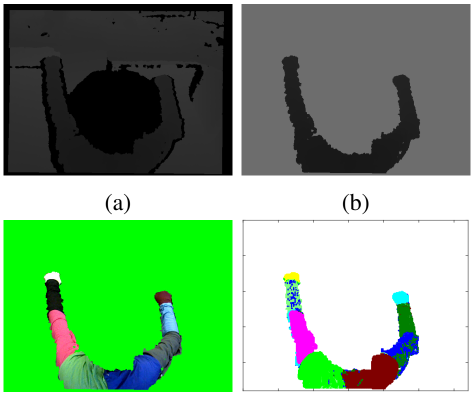
For a given pixel n in a depth image, we choose the same depth comparison features as in [14]:

$$f_\theta(I,n) = d_I(n + \frac{u}{d_I(n)}) - d_I(n + \frac{v}{d_I(n)}) \qquad (1)$$

where $d_I(n)$ is the depth value of pixel n, I indicates the specific image that pixel n comes from and $\theta = (u, v)$, where u and v are offsets. The offsets u and v are normalised by the depth of pixel n to ensure that the features are depth invariant. This feature calculates the depth differences between two neighbour pixels of n. If a neighbour pixel lies outside the bounds of the image or on the background, the depth of this neighbour pixel is set to a large positive constant value.

We follow the same training steps as [14] for each tree. Through training each tree, a pair of offsets u and v and a

Preprint version; final version available at ieeexplore.ieee.org
IROS (2015), pp: 1840-1845
Published by: IEEE



(a) (b) (c) (d)

Fig. 2: Data preprocessing. (a) shows an original depth image. (b) and (c) show the depth and RGB image after filtering head and background. The user wears clothing with eight different colours and (d) shows generated body labels for the training depth image using a colour-based segmentation method.

threshold τ are learned for every split node and a distribution over body part labels is stored in every leaf node. In the testing phase, each pixel from a filtered new depth image is classified by a learned tree model. Specifically, every pixel traverses the tree starting at the root and finally reaches a leaf node after repeatedly evaluating equation (1) and branching left or right by weak classifiers.

After classifying foreground pixels of a depth image in the testing phase, we extract the mean point of each body part to represent the position of its corresponding joint. We calculate mean values of 2D coordinates for each body part after filtering outliers, and then convert them to 3D coordinates using depth information in the camera coordinates. For the assistive robot to work with the user pose, we convert every joint position from the camera coordinates to the robot coordinates according to their spatial relationships.

B. User modelling

Our target population are users with upper body movement limitations and we aim to enable assistive robots to provide not only personalised help but also long-term adaptive assistance. For this purpose, we model the movement space of each upper-body joint.

As we discussed earlier, from any depth image, we can get 3D joint information of the user in the robot coordinates. For user modelling purpose, a sequence of N depth images is used. We define the joint set of a single depth image as $J_i = \{J_i^1, J_i^2, ..., J_i^M\}$ where M is the total number of upper-body joints, and i indicates the depth image. For each joint, $J_i^m = [x_i^m, y_i^m, z_i^m]$ where $m \in \{1, ..., M\}$ and x, y, z are 3D coordinates of the joint. The user joints information are 3D points in the space and are not informative enough for an assistive robot to know limitations of the user body. Considering that the working space of each joint is quite different, we use GMMs to model distributions of each joint movement. We define $\mathcal{J}^m = \{J_1^m, J_2^m, ..., J_N^m\}$, where $\mathcal{J}^m$ represents the set of joint m from N depth images. The Gaussian mixture distribution of $\mathcal{J}^m$ is in the form

$$p(\mathcal{J}^m) = \sum_{k=1}^{K^m} \pi_k^m \mathcal{N}\left(\mathcal{J}^m | \mu_k^m, \Sigma_k^m\right) \tag{2}$$

We adopt the unsupervised EM learning algorithm in [15] to estimate parameters of each Gaussian model. Given $\mathcal{J}^m$, the minimum and maximum number of mixture components, this algorithm outputs the best selected number of components K^m, the obtained mixture probabilities π_k^m, the estimates of the means μ_k^m and covariances Σ_k^m of the components.

In order to quantify the working space for joint m, we define π_k^m as the flexibility weight. For joint m, if $\pi_k^m \geq \frac{1}{K^m}$, then the working space of the k^{th} Gaussian model is defined as *reaching-without-difficulty* region. If $\pi_k^m < \frac{1}{K^m}$, then the working space of the k^{th} Gaussian model is defined as *reaching-with-difficulty* region.

C. Personalised dressing assistance

We propose an intuitive motion planning method for humanoid assistive robots to plan arm movements. This is inspired by real scenarios where a human assistant helps another person to wear a sleeveless jacket. A pre-defined set of goal positions are sent successively to the robot to execute, where exact values of these goal positions are determined dynamically according to real-time human upper-body pose and user models. The inverse kinematics problems [24] are solved by MoveIt!, the robot operating system (ROS) motion planning library [25]. Table I shows to which positions robot grippers move at each step. Human upper-body pose is recognised at the beginning and the calculated joint positions in the robot coordinates are the joint positions used in Table I.

During assistive dressing, the robot uses two grippers to hold the shoulder areas of a sleeveless jacket respectively, thus the positions of the robot grippers also represent the positions of the jacket shoulders. The robot holds the sleeveless jacket behind the user, which is the first step in Table I. The reason that the positions for the robot grippers are named left or right shoulder *area* is because the robot grippers should keep a short distance with the user shoulders. We assume that the right arm of the user has less moveability and the robot helps the user wear the right arm first. In the second step, the robot moves its right gripper to the right hand area of the user and its left gripper to the right shoulder area of the user. The offset in the second step depends on both

TABLE I: Robot motion planning procedures

STEP	ROBOT LEFT GRIPPER	ROBOT RIGHT GRIPPER
1	Left shoulder area	Right shoulder area
2	Right shoulder area	Right hand + *user-specific offset*
3	Towards left shoulder	Right forearm
4	Towards left shoulder	Right upper arm
5	Towards left shoulder	Right shoulder
6	Left shoulder area	No movement

Preprint version; final version available at ieeexplore.ieee.org
IROS (2015), pp: 1840-1845
Published by: IEEE

the current position of the right hand of the user and the movement space of the right hand. In the robot coordinates, the current position of the right hand of the user is defined as $p_{rh} = [x_{rh}, y_{rh}, z_{rh}]^T$ and the estimates of the means of right hand are $\mu_k^m = [\mu_{kx}^m, \mu_{ky}^m, \mu_{kz}^m]^T$ where m indicates the index for the right hand. We define the Euclidean distance between p_{rh} and μ_k^m as:

$$d_k^m = \sqrt{(x_{rh} - \mu_{kx}^m)^2 + (y_{rh} - \mu_{ky}^m)^2 + (z_{rh} - \mu_{kz}^m)^2} \quad (3)$$

With this information, the robot can put the jacket in different positions for different purposes. For instance, the robot can leave the jacket in a reaching-without-difficulty region for the right hand of the user to assist him/her easily wear the jacket with little challenges. Moreover, the robot can leave the jacket in a reaching-with-difficulty region with more challenges for the user. The reason we emphasise these challenges in assistive dressing is because our assistive robot should not simply act as a dressing machine which plans its actions using only the information of user poses, but we expect it to become a personalised assistant which can not only help dress the user but also provide rehabilitation aids for people with upper-body movement limitations. We believe that our users should not become completely dependent on the dressing help from assistive robots, where users may gradually lose their physical abilities.

In this experiment, we let the robot put the jacket in a *reaching-without-difficulty* region for the right hand of the user. The chosen user-specific offset in step 2 is d_l^m, where

$$l = \arg\min_k (\pi_k^m d_k^m) \quad \text{s.t. } \pi_k^m \geq \frac{1}{K^m} \quad (4)$$

We assume that with right arm movement limitations, the changes between the new positions of right upper arm and forearm after step 2 and their starting positions are very small, so that the robot can still plan its rest actions based on the starting positions of these two joints. Another reason for this assumption is because the view of the depth sensor is blocked by robot arms after they start moving, so that human upper-body pose could not be detected.

In step 3, the robot uses its right gripper to pull the jacket along with the right forearm of the user while its left gripper starts to move towards the left shoulder of the user. The robot arms should always keep a short distance with each other to avoid any potential self-collisions. In step 4, the robot moves its right gripper to the right upper arm of the user, and then to the right shoulder of the user to finish dressing the right arm. After the right part of the jacket is worn on the user body, the jacket adds constraints on the movement of the robot left gripper. Thus in the final step, the robot simply moves its left gripper to the left shoulder area of the user and the user needs to pull back the left arm to wear the left part of the jacket.

IV. Experimental Evaluation

We use a Xtion PRO camera which provides RGB and depth images at the frame rate of 30Hz and frame resolutions of 640×480 pixels. The assistive humanoid robot we use in the dressing application is a Baxter robot built by Rethink Robotics. To observe upper body behaviours of users from a top view, we mounted the Xtion PRO on top of the head of Baxter (see Fig. 1).

We first validate how our pose estimation and user modelling methods work with four sets of user data collected from a healthy person. Then we test the whole system by demonstrating how the Baxter robot makes use of real-time upper-body pose and user modelling information to assist human to wear a sleeveless jacket.

A. Evaluating the performance of the pose estimation and user modelling methods

We collected $3,000$ pairs of RGB and depth training images from a user which covered random arm movements without occlusions. We used 30, 300, $1,000$ and $3,000$ training images to train tree models with depth 10 and depth 20 separately. A depth 20 tree is deep enough for the amount of our training images. For each depth image, we randomly selected 504 example pixels where every 63 pixels came from one body part. For each node in the tree model during training, we randomly generated 500 candidate features and 10 candidate thresholds per feature.

For the testing purpose of upper-body pose estimation and user modelling, we let the user perform four sets of different arm movements and we selected 500 images from each set as testing images. We instructed the healthy subject to pretend to move the right arm with limitations and the four datasets were collected as below:

- *Motion A:* Arms move in left and right directions
- *Motion B:* Arms move in forward and backward directions
- *Motion C:* Arms move in up and down directions
- *Motion D:* Arms draw circles vertically

The same colour based segmentation method in Section III-A was used to generate ground-truth body labels for testing images, then we calculated the accuracy of classification results from different trained tree models, which are shown in Table II.

It can be seen that while the depth of the tree model keeps the same, the classification accuracy increases as the number of training images increases. The classification accuracy is

TABLE II: Classification accuracy of testing images with different tree models

Number of training images	Depth of tree model	Classification accuracy
30	20	42.78%
300	20	63.83%
1000	20	78.20%
3000	20	85.53%
30	10	37.86%
300	10	59.45%
1000	10	73.28%
3000	10	78.36%

Preprint version; final version available at ieeexplore.ieee.org
IROS (2015), pp: 1840-1845
Published by: IEEE

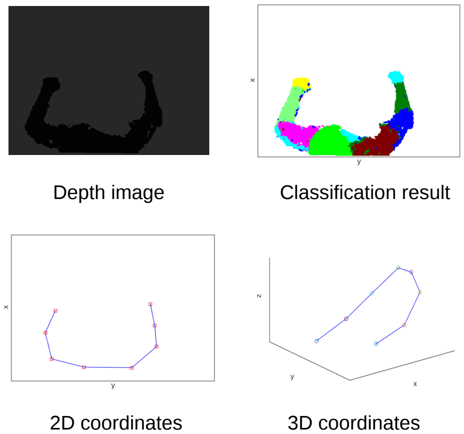


Fig. 3: Extracting 3D joint coordinates. From a testing depth image, we first classify every pixel with a trained randomised decision tree model. Then we calculate 2D coordinates of each body part after filtering outliers and finally we convert 2D coordinates to 3D coordinates by using depth information.

the highest, at 85.53%, when the number of training images is $3,000$ and the depth of tree model is 20. When the number of training images is the same, classification accuracy is smaller with depth 10 tree than with depth 20 tree.

For any testing image, we first classify each pixel and find the 2D coordinates of each joint. We calculate the mean position of each body part after filtering outliers and use this position as the joint position. Then, the 2D joint coordinates are transformed into 3D joint coordinates using depth information. The whole process is shown in Fig. 3.

For every testing dataset, we model the movement space of each upper-body joint in the camera coordinates and draw the results where the transparency of each Gaussian for a joint depends on the obtained mixture probability, which is shown in Fig. 4.

B. *Assisting users with dressing*

We test the whole system on two users by enabling a Baxter humanoid robot to provide personalised assistance to help each user wear a sleeveless jacket, where the robot makes use of real-time upper-body pose of subjects and user models. The assumed scenario is that our users have limitations with their right arm movement and thus their right arms are supposed to get more help from the assistive robot while dressing the jacket. We instructed our users to perform arm movements of Motion A which was described in Section IV-A and built their user models. We let the Baxter robot put the sleeveless jacket in a *reaching-without-difficulty* position for the right hand of each user, where this position was chosen according to equation (12). Then, each user was assisted by the robot following the procedures in Table I. The interval of Baxter moving its grippers from current positions to new positions was 3 seconds.

Fig. 5 shows sequential shots of the Baxter robot assisting two users to wear a sleeveless jacket individually[1]. For both

[1]The video results can be found at `http://www3.imperial.ac.uk/personalrobotics/videos_new`

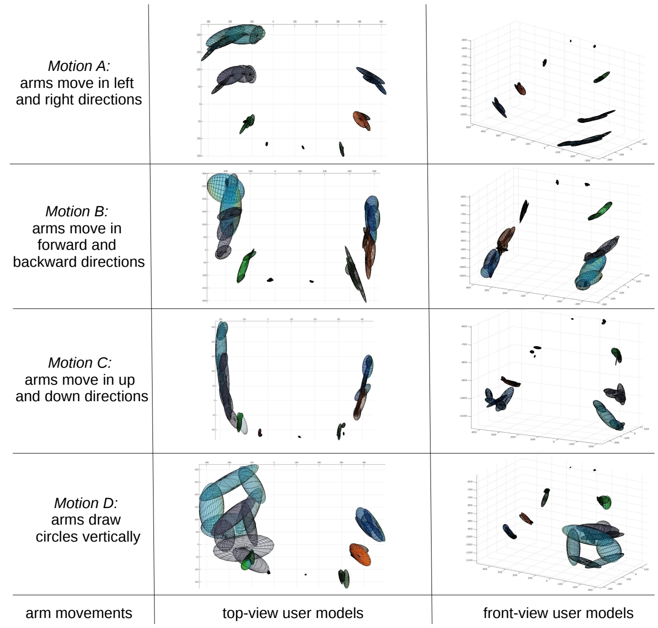


Fig. 4: User models for four sets of arm movement with different limitations. The movement space of each upper-body joint is modelled as a mixture of Gaussian. The transparency of each Gaussian depends on the obtained mixture probabilities using an unsupervised EM learning algorithm. This figure is better shown on screen.

users, the Baxter robot recognised their real-time upper-body pose successfully and used this information combined with user modelling knowledge to plan its arm trajectories. The robot moves two grippers slowly while dressing users and the planned trajectories of robot arms are guaranteed to be safe by adding orientation constraints to robot grippers and always keeping a minimum distance between two arms of the robot to avoid any potential robot self-collisions during trajectory planning. To further guarantee the safety of users, the whole dressing process is under careful supervision by researchers.

V. Conclusions

We have presented an end-to-end approach to build user-specific models for home-environment humanoid robots to provide personalised dressing assistance. By mounting a depth camera on top of the head of a Baxter humanoid robot, we recognise the upper-body pose of users from a single depth image using randomised decision forests. From sequences of upper-body movements, the movement space of each upper-body joint is modelled as a mixture of Gaussian learned by an EM algorithm. We have demonstrated how user models are built with four sets of user data which have pretended limitations of arm movements. The experimental results show that our method of modelling upper-body joint movement of users combined with real-time human upper-body pose recognition enables a humanoid robot to provide personalised dressing assistance and our method has potential use in rehabilitation robotics and long-term human-robot interactions. We plan to extend our research in enabling

Preprint version; final version available at ieeexplore.ieee.org
IROS (2015), pp: 1840-1845
Published by: IEEE

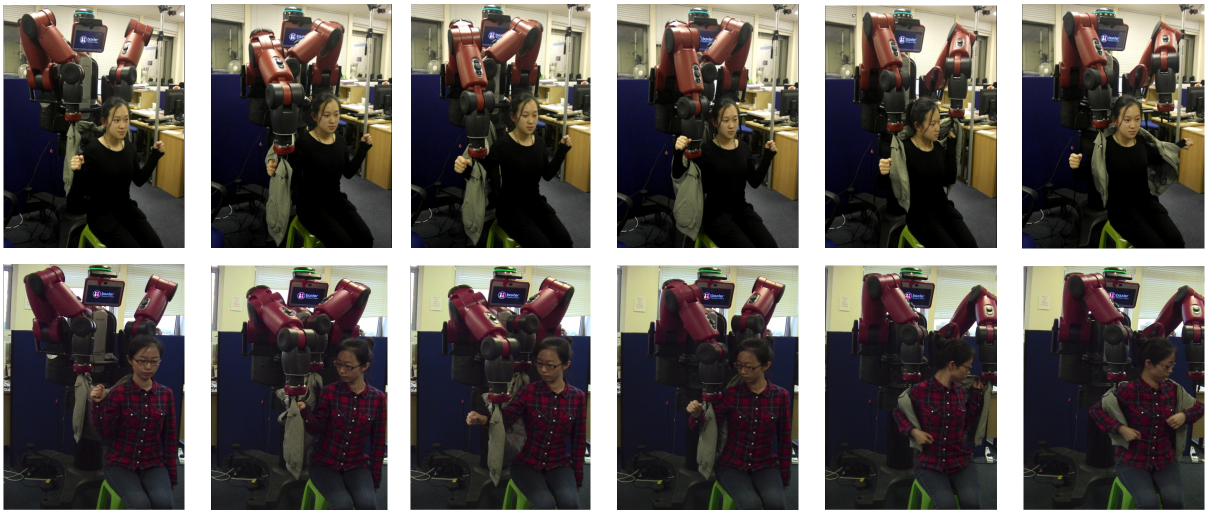


Fig. 5: Sequential shots of personalised assistive dressing. A Baxter robot assists two users to wear a sleeveless jacket individually. The robot motions are planned according to the user models and real-time upper-body pose.

humanoid robots to provide long-term adaptive assistance for users with upper body limitations by updating user models online.

ACKNOWLEDGMENT

This work was supported in part by the EU FP7 project WYSIWYD under Grant 612139. The authors gratefully acknowledge the help from all the members of the Imperial College Personal Robotics Lab.

REFERENCES

[1] P. Flandorfer, "Population ageing and socially assistive robots for elderly persons: the importance of sociodemographic factors for user acceptance," *International Journal of Population Research*, vol. 2012, 2012.

[2] A. Tapus, M. Maja, and B. Scassellatti, "The grand challenges in socially assistive robotics," *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, pp. N–A, 2007.

[3] J. Broekens, M. Heerink, and H. Rosendal, "Assistive social robots in elderly care: a review," *Gerontechnology*, vol. 8, no. 2, pp. 94–103, 2009.

[4] J. Fasola and M. J. Mataric, "Using socially assistive human–robot interaction to motivate physical exercise for older adults," *Proceedings of the IEEE*, vol. 100, no. 8, pp. 2512–2526, 2012.

[5] C. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, C. Huijnen, H. van den Heuvel, A. van Berlo, A. Bley, and H.-M. Gross, "Realization and user evaluation of a companion robot for people with mild cognitive impairments," in *ICRA*, 2013, pp. 1153–1159.

[6] K. Yamazaki, R. Ueda, S. Nozawa, Y. Mori, T. Maki, N. Hatao, K. Okada, and M. Inaba, "System integration of a daily assistive robot and its application to tidying and cleaning rooms," in *IROS*, 2010, pp. 1365–1371.

[7] F. Gravot, A. Haneda, K. Okada, and M. Inaba, "Cooking for humanoid robot, a task that needs symbolic and geometric reasonings," in *ICRA*, 2006, pp. 462–467.

[8] A. Ramisa, G. Alenyà, F. Moreno-Noguer, and C. Torras, "Learning rgb-d descriptors of garment parts for informed robot grasping," *Engineering Applications of Artificial Intelligence*, vol. 35, pp. 246–258, 2014.

[9] A. Colomé, A. Planells, and C. Torras, "A friction-model-based framework for reinforcement learning of robotic tasks in non-rigid environments," *Proceedings of ICRA*, pp. 5649–5654, 2015.

[10] T. Tamei, T. Matsubara, A. Rai, and T. Shibata, "Reinforcement learning of clothing assistance with a dual-arm robot," in *Humanoids*, 2011, pp. 733–738.

[11] 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.

[12] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and S. Kiesler, "Human preferences for robot-human hand-over configurations," in *IROS*, 2011, pp. 1986–1993.

[13] E. A. Sisbot, L. F. Marin, and R. Alami, "Spatial reasoning for human robot interaction," in *IROS*, 2007, pp. 2281–2287.

[14] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, "Real-time human pose recognition in parts from single depth images," *Communications of the ACM*, vol. 56, no. 1, pp. 116–124, 2013.

[15] M. A. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 3, pp. 381–396, 2002.

[16] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and autonomous systems*, vol. 42, no. 3, pp. 143–166, 2003.

[17] T. Fong, C. Thorpe, and C. Baur, "Collaboration, dialogue, human-robot interaction," in *Robotics Research*. Springer, 2003, pp. 255–266.

[18] A. Tapus and M. J. Matarić, "User personality matching with a hands-off robot for post-stroke rehabilitation therapy," in *Experimental robotics*. Springer, 2008, pp. 165–175.

[19] K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, A. Sisbot, R. Alami, and T. Siméon, "How may i serve you?: a robot companion approaching a seated person in a helping context," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 2006, pp. 172–179.

[20] Y. Wu, Y. Su, and Y. Demiris, "A morphable template framework for robot learning by demonstration: Integrating one-shot and incremental learning approaches," *Robotics and Autonomous Systems*, 2014.

[21] K. Lee, Y. Su, T.-K. Kim, and Y. Demiris, "A syntactic approach to robot imitation learning using probabilistic activity grammars," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1323–1334, 2013.

[22] J. Fasola and M. Mataric, "A socially assistive robot exercise coach for the elderly," *Journal of Human-Robot Interaction*, vol. 2, no. 2, pp. 3–32, 2013.

[23] R. C. Luo, S. Y. Chen, K. Yeh, *et al.*, "Human body trajectory generation using point cloud data for robotics massage applications," in *ICRA*, 2014, pp. 5612–5617.

[24] Y. Nakamura and H. Hanafusa, "Inverse kinematic solutions with singularity robustness for robot manipulator control," *Journal of dynamic systems, measurement, and control*, vol. 108, no. 3, pp. 163–171, 1986.

[25] S. Chitta, I. Sucan, and S. Cousins, "Moveit!" *IEEE Robotics Automation Magazine*, vol. 19, no. 1, pp. 18–19, 2012.