



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/155658/>

Version: Published Version

Article:

Jevtić, A., Flores Valle, A., Alenyà, G. et al. (2019) Personalized robot assistant for support in dressing. *IEEE Transactions on Cognitive and Developmental Systems*, 11 (3). pp. 363-374. ISSN: 2379-8920

<https://doi.org/10.1109/TCDS.2018.2817283>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Personalized Robot Assistant for Support in Dressing

Aleksandar Jevtić¹, Andrés Flores Valle, Guillem Alenyà, *Member, IEEE*, Greg Chance, Praminda Caleb-Solly, Sanja Dogramadzi, and Carme Torras, *Senior Member, IEEE*

Abstract—Robot-assisted dressing is performed in close physical interaction with users who may have a wide range of physical characteristics and abilities. Design of user adaptive and personalized robots in this context is still indicating limited, or no consideration, of specific user-related issues. This paper describes the development of a multimodal robotic system for a specific dressing scenario—putting on a shoe, where users’ personalized inputs contribute to a much improved task success rate. We have developed: 1) user tracking, gesture recognition, and posture recognition algorithms relying on images provided by a depth camera; 2) a shoe recognition algorithm from RGB and depth images; and 3) speech recognition and text-to-speech algorithms implemented to allow verbal interaction between the robot and user. The interaction is further enhanced by calibrated recognition of the users’ pointing gestures and adjusted robot’s shoe delivery position. A series of shoe fitting experiments have been performed on two groups of users, with and without previous robot personalization, to assess how it affects the interaction performance. Our results show that the shoe fitting task with the personalized robot is completed in shorter time, with a smaller number of user commands, and reduced workload.

Index Terms—Assistive robots, multimodal human–robot interaction (HRI), robot personalization.

I. INTRODUCTION

BY 2050, the world population is expected to increase by 2 to 4 billion people [1]. This growth will have a profound demographic consequence: while in 2000, 10% of the world’s population was over 60 years old, by 2050 this proportion will be more than doubled. Some studies report that more than half of the people 75 years or older who suffer from age-related physical and cognitive impairment need assistance

Manuscript received September 29, 2017; revised February 23, 2018; accepted March 14, 2018. Date of publication March 20, 2018; date of current version September 9, 2019. This work was supported in part by the I-DRESS Project under Grant EU CHIST-ERA 2014, in part by the Programme Acciones de Programacion Conjunta Internacional 2015 of the Spanish Ministry of Economy, Industry and Competitiveness under Project PCIN-2015-147, in part by the U.K. EPSRC under Project EP/N021703/1, and in part by the Spanish State Research Agency through the María de Maeztu Seal of Excellence to IRI under Project MDM-2016-0656. (*Corresponding author: Aleksandar Jevtić.*)

A. Jevtić, A. Flores Valle, G. Alenyà, and C. Torras were with the Institut de Robòtica i Informàtica Industrial, CSIC-UPC, 08028 Barcelona, Spain (e-mail: ajevtic@iri.upc.edu; anflores@iri.upc.edu; galenyà@iri.upc.edu; torras@iri.upc.edu).

G. Chance, P. Caleb-Solly, and S. Dogramadzi were with the Bristol Robotics Laboratory, University of the West of England, Bristol BS16 1QY, U.K. (e-mail: greg.chance@uwe.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCDS.2018.2817283



Fig. 1. Assisted-dressing scenario with a WAM robot: use case of putting on a shoe.

with activities of daily living (ADL) [2]. Assistive technologies can improve the life quality for both older adults and their caregivers [3]. Assistive robots, in particular, can help patients with recovery and allow prolonged independent living, while compensating for increased costs of care and lack of nursing staff [4].

The main goal of this paper is development of an autonomous robot that provides personalized assistance to a user in performing a dressing task. In this context, the considered dressing task consists in comfortably putting on a shoe which has been selected by the user. The experiments were designed to evaluate robot performance and user workload under different conditions. The user is assumed to have reduced mobility, partial control over legs, and is in a seated position as shown in Fig. 1. The user should be able to interact with the robot through a number of modalities. This will allow the robot to be adaptable to situations where a single modality is insufficient, e.g., asking the robot to pick up “the black shoes” where there are several choices. Ambiguity may be reduced with the addition of the gesture modality, in this case pointing. The interpretation of the pointing gesture may be difficult due to the context of the situation. Pointing to an object relatively nearby compared to one further away may result in a different arm pose (e.g., elbow bent or straight,

TABLE I
SUMMARY OF THE MOST RELEVANT STUDIES IN ASSISTED DRESSING

Publication	Application	Modalities	Adaptation method	User tests
T. Tamei et al., 2011 [6]	T-shirt dressing	Vision	N/A	No
N. Koganti et al., 2013 [7]	Cloth state estimation	Vision	N/A	No
N. Koganti et al., 2014 [8]	T-shirt dressing	Vision	N/A	No
T. Matsubara et al., 2013 [10]	Cloth estimation	Vision	N/A	No
A. Colomé et al., 2015 [12]	Scarf dressing	Force	N/A	No
G. Chance et al., 2016 [13]	Jacket dressing	Vision, speech, accelerometer	N/A	No
K. Yamazaki et al., 2016 [16]	Trousers dressing	Vision, force	N/A	Yes
K. Yamazaki et al., 2013 [23]	Cloth state estimation	Vision	N/A	No
Y. Gao et al., 2015 [14]	Jacket dressing	Vision	Model of body motion space	Yes
J. P. Costeira et al., 2015 [15]	Hat dressing	Vision	Learning of user motion constraints	Yes
E. Pignat et al., 2017 [17]	Jacket and shoe dressing	Vision	Modeling user behavior with HMM	No

hand rotated) and for this reason specific calibration, i.e., robot personalization, is required.

Natural human–robot interaction (HRI) requires successful recognition of the user’s and robot’s intentions [5]. In the shoe fitting task, the successful interaction is based on continuous tracking of the shoe and the user. The contribution of this paper is twofold. First, a multimodal robotic system for support in dressing was developed. Several vision- and speech-based modalities have been developed to deal with the user’s and robot’s intentions in real time. Second, we proposed a robot personalization method to evaluate the ability of the developed multimodal robot to adapt to an individual user. The personalization focused on reducing user workload and frustration, especially important for users with reduced mobility.

A. Relevant Work

Assisted dressing is receiving increased attention in the robotics community. Earlier studies evaluated assisted dressing on a mannequin with a dual-arm robot [6], [7]. The robot was able to pull a T-shirt over the mannequin’s head while tracking the position of the collar and sleeves. In [8], the work of the same authors was extended to include learning of the mannequin-cloth relationship. Successful manipulation of some types of garments depends on accurate estimation of their state [9]. To get a better insight into the interaction between the robot and nonrigid garments, some authors proposed to perform a dressing task on a dual-arm robot, by putting the robot arms into the corresponding sleeves of a T-shirt [10].

An important aspect of HRI is safety, where the adaptation to users can be studied from the aspect of user’s limitations in avoiding events that can lead to discomfort or injuries [11]. Still, most of the studies on safety in robot-assisted dressing have not included tests with users and were limited to experiments on a mannequin. Some proposed solutions employ learning techniques to teach a compliant robot arm to wrap a scarf around a mannequin’s neck [12] or detect failures in jacket dressing [13]. The proposed scenarios with a mannequin have a limited utility for real-world applications because the mannequin’s position is always fixed. The obtained results are difficult to generalize when applied to human motion.

Adaptation to users is of great importance for acceptance of the robots, not least for persons with reduced mobility. Gao *et al.* [14] proposed building of a unique model that

defines user’s mobility space. A different approach of personalized assistance was proposed in [15], where the robot and user take turns when moving to compensate for the user’s mobility limitations. Although some level of adaptation was achieved in these studies, no perception of the garment state was considered. Recent work by Yamazaki *et al.* [16] included both garment state estimation and personalized assistance for users, allowing a humanoid robot to assist users with putting on a pair of trousers. The personalized assistance was incorporated into the robot’s motion planning, taking into account visual feedback of the trousers and the size of the user’s legs. Pignat and Calinon [17] applied learning by demonstration to provide personalized assistance with dressing. The authors used hidden semi-Markov models to encode sensory and motor information necessary to perform both time-dependent and independent dressing task segments.

Most of the early work on robot-assisted dressing relied on vision as the primary interaction modality, as summarized in Table I. Recent studies included additional modalities, such as haptics to improve the interaction with the user [18]–[22]. The evaluation of such systems focused on robot performance without considering the direct user input for robot personalization, hence limiting the scalability of such systems in applications with people. In the work presented in this paper, a robotic system was developed that exploits speech-based and vision-based interaction modalities to successfully assist a user with a dressing task, and can be customized to the particular set of user abilities and needs through direct input from the user. The results provide a proof-of-concept for the I-DRESS project,¹ which aims to develop a multimodal robotic system equipped with a wide range of sensors and safety features to provide proactive assistance with dressing to users with limited mobility.

II. METHODOLOGY

In the context of an assisted-dressing task in which a robot assists the user in putting on a shoe, every person would have a particular way of interacting. The multimodal approach developed in this research enables the system to learn and respond to individual anthropometrics, speech, and gestures commands resulting in personalized interaction with a user. The development of the robot assistant for support in

¹The I-DRESS project: <https://i-dress-project.eu/>.

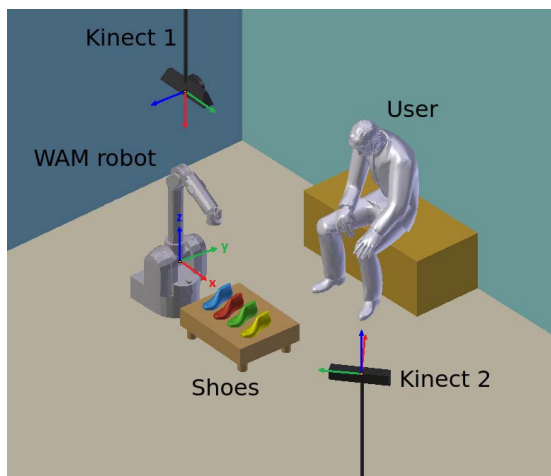


Fig. 2. Dressing scenario.

dressing required integration of several hardware and software components. The robot features several vision-based and speech-based modalities for interaction with the user.

A. Task Description

The application scenario consists of a user's daily activity of putting on a shoe in a seated position. The target users are persons with reduced mobility, with partial control of their legs, i.e., having a certain level of difficulty in lifting their legs and moving their feet. The user may choose from a set of shoes using speech or a combination of speech with pointing gestures to form so-called deictic expressions. The robot's task is to grasp the requested shoe and position and hold it in an appropriate position in front of the user so that they can comfortably place their foot inside.

In the first instance, experiments were performed to compare task efficiency using single or combined interaction modalities. Each participant performed two experiments. In the first experiment, only speech could be used to request a shoe; in the second, the participants were asked to combine the speech with the pointing gesture into deictic commands. The experiments were performed in the laboratory environment, and a graphical model of the scenario is shown in Fig. 2.

B. Hardware

The central part of the system is a Barrett's 7-DOF WAM robotic arm equipped with an in-house developed gripper shown in Fig. 1. The gripper has four fingers, which are controlled by a servo motor [see Fig. 3(a)]. A set of cros-type of shoes commonly used by patients in hospitals was also used in this scenario. Each shoe has a ribbon attached that is grasped by the four fingers before the shoe can be moved to the user [see Fig. 3(b)]. The ribbons are marked with four different color markers for easier recognition [see Fig. 3(c)].

Visual input is provided by two Microsoft Kinect cameras, an XBOX 360 and a Kinect One, which will be referred to as Kinect 1 and Kinect 2, respectively. The depth and RGB images from the Kinect 1 are used to recognize the colors and

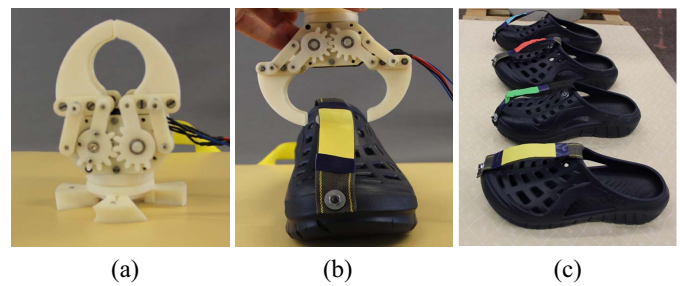


Fig. 3. Gripper and cros shoes used in the experiments. (a) Gripper, (b) ribbon attached for easier grasping, and (c) color markers used for recognition of both the shoes and their grasping points.

locations of the shoes markers. User tracking, posture, and gesture recognition rely on depth images provided by the Kinect 2, while the audio input from its integrated 4-microphone array operating at 48 kHz is used for speech recognition and sound localization. The cameras were connected to different personal computers (PCs) and showed no noticeable interference during operation, which can sometimes occur when using two cameras. One of the reasons for no noticeable interference may be different orientation of the two cameras: the Kinect 1 was facing downward, while the Kinect 2 was facing the user. Also, some studies suggest that use of different technologies to compute depth may reduce interference in a dual-camera setup: while Kinect 1 computes alterations in the IR light pattern it projects, Kinect 2 computes the IR rays time of flight.

The integration of hardware and algorithms was performed in robot operating system (ROS). Three PCs run the entire system. A PC running Ubuntu 12.04 LTS 64-bit, powered by an Intel quad-core Q9550 CPU @2.83 GHz×4 with 8 GB of RAM was used to run most of the implemented algorithms and to connect the Kinect 1 camera. The second PC running Ubuntu 12.04 LTS 64-bit powered by an Intel Core i5-2400 CPU @3.10 GHz×4 and 4 GB of RAM was used to control the WAM robot and the gripper, having all the necessary drivers installed. The third PC running Windows 8.1 Pro 64-bit, powered by an Intel Core i7 X990 @3.47 GHz and 2.80 GHz and 16 GB of RAM, processed the speech recognition and user tracking data obtained using the Kinect for Windows SDK 2.0 library. The three PCs communicated via laboratory Ethernet.

C. Algorithms

Vision and speech were used as inputs for development of several modalities for HRI, but also for the interaction of the robot with the environment (e.g., recognition of the shoes). Some authors associate modalities with the type of perception, e.g., vision, sound, etc., however, we use a more detailed definition of modality as a channel for a certain type of message between the user and the robot, such as posture, gesture, etc., which can be developed from the same sensory input, such as vision. Verbal interaction between the user and the robot was implemented through speech recognition and speech synthesis algorithms. Visual interaction consisted of user tracking, pointing recognition, and posture recognition. An additional modality was deictic expression recognition that combined

TABLE II
SPOKEN UTTERANCES WITH ASSOCIATED
SEMANTIC TAGS AND ACTIONS

Spoken utterance	Semantic tag	Robot action
“begin”	start	waits for user to select a shoe
“calibrate pointing”	pointing	starts calibration
“take this shoe”	take	picks up the shoe pointed at
“take the {color} shoe”	{color}	picks up the {color} shoe
“dress me”	dress	waits for dressing posture
“move {direction}”	{direction}	moves end-effector {direction}
“stop”	stop	stops moving
“that’s ok”	ok	releases shoe from gripper
“abort”	abort	returns to home position

speech and pointing recognition. Finally, adaptation to users, or personalization, consists of calibrating each person’s pointing gesture and adjusting the robot’s position to suit the user ergonomically. User personalization method is described in Section III.

1) *Speech Recognition*: Speech was used for bidirectional communication between the user and the robot. Through speech recognition the robot was able to understand user’s voice commands to start or finish the task, correct its behavior or learn user preferences. The implementation of the speech-recognition algorithm was made through the Microsoft Speech Platform SDK 11 engine, which transcribes spoken utterances to text. A grammar model was created in XML-format to define the utterances specific to assisted-dressing scenario. Each utterance was associated with a semantic tag, which was retrieved when the utterance was recognized. A set of the utterances and associated semantic tags used in the experiments is given in Table II.

2) *Speech Synthesis*: Robot feedback is an important aspect of HRI as it allows the user to understand the robot’s current state and actions. It is used to inform the user about the progress of the dressing task and necessary actions; for example, after a shoe is picked up, the dressing assistance will not continue until the user extends the foot toward the robot. Robot verbal feedback is also used to confirm whether a user command was correctly recognized, which contributes to user safety but also allows a timely intervention by the user in order to correct the robot’s behavior. A text-to-speech algorithm was implemented in Python, and relies on the *gTTS* package using the Google’s Text-to-Speech API. The algorithm takes a text string as input and converts it into a speech transcription in *mp3* format reproduced by the speakers. Similarly to speech recognition, a vocabulary of utterances was defined specific to the assisted-dressing scenario. Examples of the utterances are: “ready to help,” “taking the {color} shoe,” “please, approach,” etc.

3) *User Tracking*: The ability to track and follow user’s body parts, such as a foot or a hand, is necessary to perform the proposed assisted-dressing task. Microsoft Kinect SDK provides tracking of 25 body joints, with their position and orientation, at a 10 Hz frame rate [24]. Specifically, tracking of the position of the foot and the orientation of the knee–ankle axis were implemented for a proper positioning of the shoe (see Fig. 8), but also to ensure collision avoidance and to keep the interaction safe.

4) *Pointing Recognition*: The use of pointing gestures for robot control proved to be an accepted way of interaction for inexperienced users [25]. Various pointing recognition methods have been proposed in literature, which were tailored according to system’s sensing abilities, e.g., finger tracking [26], or task requirements, e.g., distance of the pointing target [27]. Our previous studies showed that the pointing recognition using the position of the elbow and wrist joints can successfully be applied to robot control in close HRI [28], [29]. The user-tracking algorithm described in Section II-C3 provides the position of the arm joints in real-time, hence it was possible to implement the same method in the current study.

The estimation of the user pointing target was applied in combination with speech to form deictic expressions, which allowed more diverse and intuitive interaction with the user. For example, the user could point to a desired shoe while saying “take this shoe!” and the shoe closest to the pointing target would be selected, as shown in Fig. 7(a). Even though the reference to the color using speech seems to be easier and simpler when distinguishing the shoes, the pointing gesture is likely to provide a more reliable alternative solution in real life situations when the colors might not be descriptive enough to discriminate different objects; for example, there may be more than one pair of shoes of the same color, or the user may not remember the exact name of the color, etc.

The computation of the pointing target was performed in the robot frame of reference. Let $\mathbf{p}_e = (x_e, y_e, z_e)$ be the position of the user’s elbow and $\mathbf{p}_w = (x_w, y_w, z_w)$ the position of the user’s wrist, both obtained from the Kinect 2 applying the user skeleton-tracking algorithm. The pointing direction is computed as a straight line

$$\mathbf{s} = \mathbf{p}_e + \lambda(\mathbf{p}_w - \mathbf{p}_e) \quad (1)$$

where $\lambda \in \mathfrak{R}$. In the proposed dressing scenario, the shoes are placed on a platform that is parallel to the ground floor at the constant height, $z = h$. After substituting this value in (1), the pointing target, $\mathbf{p}_t = (x_t, y_t, z_t)$, which is found at the intersection of the pointing line with the shoe plane is given by

$$\begin{aligned} x_t &= x_e + \frac{h - z_e}{z_w - z_e}(x_w - x_e) \\ y_t &= y_e + \frac{h - z_e}{z_w - z_e}(y_w - y_e) \\ z_t &= h. \end{aligned} \quad (2)$$

Finally, let $S = \{\text{blue, red, green, yellow}\}$ be a set of the available shoes on the platform, and \mathbf{p}_s , $s \in S$, their respective locations that are obtained with the shoe-recognition algorithm described later in this section. The target shoe $s_t \in S$ is selected as the closest one from the pointing target

$$s_t = \arg \min_{s \in S} (|\mathbf{p}_t - \mathbf{p}_s|). \quad (3)$$

A graphical representation of shoe selection is shown in Fig. 7(a), where for demonstration purposes the blue shoe was selected by the user.

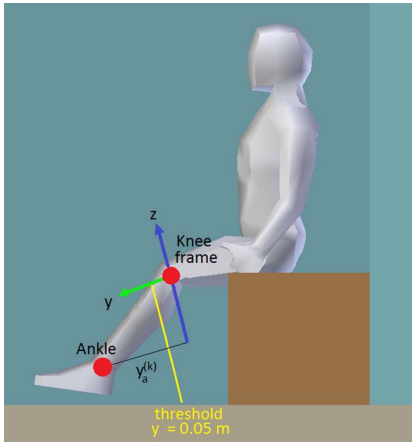


Fig. 4. Dressing posture recognition using relative ankle and knee positions.

5) *Posture Recognition*: Posture recognition was developed to detect the user's readiness to be dressed after the shoe selection phase. The algorithm is able to recognize when the user's right leg is extended toward the robot by analyzing the position of the knee and the ankle joints. The leg is considered to be extended when the ankle joint passes the perpendicular axis of the femur bone by more than 0.05 m respect to the knee joint, which is shown in Fig. 4.

Posture recognition evaluates the user's intention to be dressed but it also contributes to user safety. The algorithm was running at 10 Hz, and a threshold was used to detect the change in posture. If the user withdraws the foot, the robot returns to the home position and waits for the next instruction. The posture recognition algorithm is executed after the shoe selection phase, only when the user verbally confirms intention to be dressed by saying "dress me."

6) *Shoe Recognition*: For the proposed assisted-dressing scenario, shoe manipulation was simplified by attaching a ribbon to the top of a shoe so that the gripper can grasp the shoe from above [see Fig. 3(b)]. The ribbons were of size 3 cm \times 17 cm, with rectangular 3 cm \times 6 cm color markers placed in the central segment of the ribbon. The recognition of the markers was implemented using the OpenCV image-processing library that takes both RGB and depth images provided by the Kinect 1 to compute the color and position of different segments in the image. The Kinect 1 was mounted above the shoe platform providing a top view of the shoes. The experimental set consisted of four shoes marked with blue, green, red, and yellow markers shown in Fig. 5.

The RGB images obtained with the Kinect 1 were first converted to HSV format. The colors in the image were clustered according to their HSV values and their centroids were computed. The HSV values of the markers used in the experiments were obtained from the test sample images and their ranges are given in Table III. Depth images obtained with the Kinect 1 were used to compute the coordinates of the markers' centroids in the camera reference system. The positions of the markers were transformed to the robot reference system and set as the corresponding shoe's gripping points. It is important to note that the algorithm was executed each time the user requested a

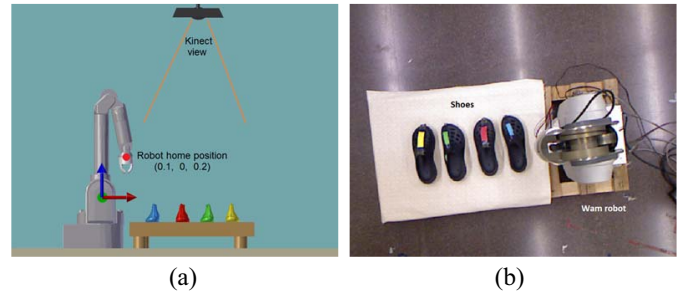


Fig. 5. Shoe recognition. (a) Simulated experimental setup and (b) real-world view from the Kinect camera.

TABLE III
HSV VALUES RANGE FOR SHOE MARKER RECOGNITION

	Hue	Saturation	Value
blue	[97, 110]	[70, 150]	[100, 255]
red	[167, 179]	[150, 220]	[100, 255]
green	[48, 58]	[70, 130]	[100, 255]
yellow	[20, 30]	[110, 170]	[100, 255]

shoe from the robot. The marker positions were used to define the shoes gripping points, but also to inform the user if the requested shoe had already been picked up and is no longer available on the platform. The described implementation made the system more robust to unexpected user behavior.

7) *Robot Motion Planning*: Shoe grasping and positioning to enable comfortable insertion of the foot by the user required accurate robot movement. To reach a desired point in robot's workspace, the end-effector directional points provided in Cartesian space were transformed into robot joints positions that satisfy the constraints implemented through an inverse kinematics algorithm [30]. The robot operated in a compliant mode to ensure user safety. Predefined positions of the robot's end-effector were associated with different robot states. In the *home position* shown in Fig. 5, the robot waited for the user to initiate the task. After receiving a requests to pick up a shoe, it computed the position of the shoe marker and verified that the selected shoe was reachable. To ensure successful grasping and avoid collision with other shoes, the robot gripper was guided through a set of predefined directional points above the selected shoe's marker. After the user's request to be dressed, the robot delivered the shoe to the *delivery position* (see Fig. 8), at a safe distance from the user's right foot. This distance was empirically obtained from the test trials. It was computed with respect to the position of the user's ankle in the knee frame of reference, at $d_{xy} = 0.4$ m in the xy plane taking into account the orientation of the right leg along the knee-ankle axis, and $d_z = 0.5$ m in the z -axis. The adjustment of the delivery position was a part of robot personalization method described in Section III-B.

The robot was capable of adjusting the delivery position by following the user's foot, which consisted in maintaining the distance and adjusting the orientation of the gripper. The preliminary tests showed that the recognition of the foot orientation was unreliable. For this reason, the axis passing through the ankle and knee joints was used as a reference. Let

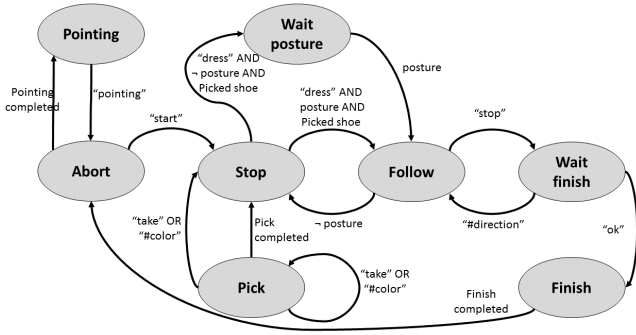


Fig. 6. Decision-making module diagram.

$\mathbf{p}_a = (x_a, y_a, z_a)$ and $\mathbf{p}_k = (x_k, y_k, z_k)$ be the positions of the user's ankle and knee, respectively. The position of the ankle with respect to the knee $\mathbf{p}_a^{(k)} = (x_a^{(k)}, y_a^{(k)}, z_a^{(k)})$ is computed as

$$\mathbf{p}_a^{(k)} = \mathbf{p}_a - \mathbf{p}_k. \quad (4)$$

The angle between the knee and ankle with respect to the robot's x -axis is then given by

$$\beta = \tan^{-1}\left(\frac{y_a^{(k)}}{x_a^{(k)}}\right) + \begin{cases} \pi, & \text{if } (x_a^{(k)} < 0) \text{ and } (y_a^{(k)} > 0) \\ -\pi, & \text{if } (x_a^{(k)} < 0) \text{ and } (y_a^{(k)} < 0) \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

And for the case of $x_a^{(k)} = 0$

$$\beta = \begin{cases} \pi/2, & \text{if } (y_a^{(k)} > 0) \\ -\pi/2, & \text{if } (y_a^{(k)} < 0) \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

By knowing the angle and distances in the xy plane and z -axis, the robot end-effector position $\mathbf{p}_r = (x_r, y_r, z_r)$ can now be computed

$$\begin{aligned} x_r &= x_a + d_{xy} \cos \beta \\ y_r &= y_a + d_{xy} \sin \beta \\ z_r &= z_a + d_z. \end{aligned} \quad (7)$$

The position is continuously updated allowing the robot to follow the user's foot, while keeping a predefined distance for safety.

8) *Decision-Making Module*: The decision-making module is implemented as a finite-state machine, as shown in the diagram in Fig. 6. It integrates all the above-described algorithms, and defines the robot behavior with eight possible states: 1) abort; 2) stop; 3) pick; 4) wait posture; 5) follow; 6) wait finish; 7) finish; and 8) pointing. Transitions between the states are evoked by the interaction events detected by any of the interaction modalities, and these events are also shown in the diagram. In case of inconsistent user input, the robot remains in the current state and via spoken feedback informs the user about the issue and requests a new input.

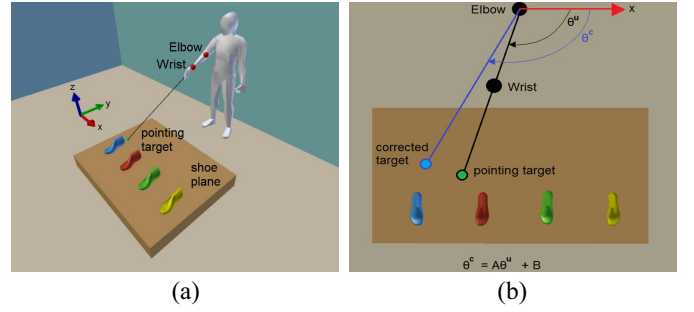


Fig. 7. Computation of the pointing target for the blue shoe: the user angle, θ^u , is computed as the angle of the elbow-wrist axis in the robot frame of reference. The corrected angle, θ^c , is computed using a linear fitting function whose parameters A and B are obtained during the pointing calibration procedure.

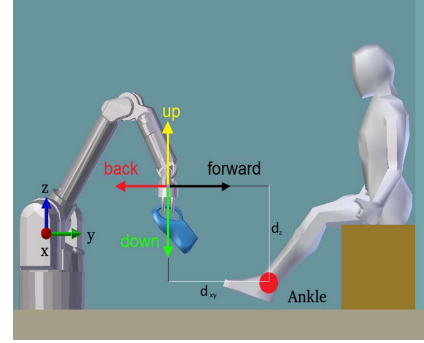


Fig. 8. Robot shoe delivery requires tracking and following of the user's ankle in real-time. During personalization, the user can use voice commands to adjust the position of the robot.

III. ROBOT PERSONALIZATION

To develop a personalized robot dressing assistant, a method consisting of user pointing calibration and robot position adjustment was proposed. Pointing calibration improves the accuracy of the pointing recognition during shoe selection, while the robot position adjustment allows the users to modify the shoe delivery position for a better comfort. This is especially important for users with mobility issues who may perform pointing and foot positioning differently, in accordance with their limitations.

A. Pointing Calibration

Pointing is performed differently by each user, and the estimation of the pointing target may largely differ from the one that is perceived by the user. For this reason, a pointing calibration algorithm was proposed that compensates the user's pointing error and takes into account specific task requirements. Preliminary experiments showed user consistency in pointing. It is important to note that users were in a seated position that restricted their pointing gesture, which in the proposed scenario ensured successful repeatability of the pointing action. The calibration procedure is initiated by the user and it is described in Algorithm 1. It can be performed as many times as needed, although for this paper it was performed only once before the assisted-dressing experiment.

Algorithm 1 Pointing Calibration

```

1:  $\{A, B\} \leftarrow \{1, 0\}$  // default parameters values
2: if “pointing” then // user says “calibrate pointing”
3: shoes  $\leftarrow \{\text{blue, red, green, yellow}\}$ 
4: user_angles  $\leftarrow \{\}$ 
5: corrected_angles  $\leftarrow \{\}$ 
6: for all shoe in shoes do // user points to a shoe
7:  $\theta^u \leftarrow \text{get\_pointing\_angle}()$ 
8:  $\theta^c \leftarrow \text{get\_shoe\_angle}(\text{shoe})$ 
9: user_angles  $\leftarrow \text{append } \theta^u$ 
10: corrected_angles  $\leftarrow \text{append } \theta^c$  // robot says “OK”
11: end for
12:  $\{A, B\} \leftarrow \text{linear\_fitting}(\text{user\_angles, corrected\_angles})$ 
13: end if
14:  $\theta^c \leftarrow A\theta^u + B$  // applying correction

```

During calibration, the robot asks the user to point to all four shoes in a predefined order. The user points to each shoe and confirms the pointing by voice. The robot stores the pointing target associated with its corresponding shoe, and confirms this to the user. For each target, the algorithm computes two angles in the robot frame of reference: 1) the user angle, θ_s^u and 2) the corrected angle, θ_s^c , $s \in \{\text{blue, red, green, yellow}\}$, as shown in Fig. 7(b). The user angle is computed from the straight line passing through the elbow and wrist joints and the robot’s x -axis; similarly, the corrected angle is computed as the straight line connecting the elbow and the shoe s and the x -axis. The values obtained in preliminary trials suggested a close-to-linear relationship between the two sets of angles, θ_s^c and θ_s^u

$$\theta^c = A\theta^u + B. \quad (8)$$

During experiments, individual user’s pointing calibration results, i.e., the four values of θ_s^c and θ_s^u obtained for four colored markers, were used to compute the parameters A and B of the linear fitting function. Let $\mathbf{p}_u = (x_u, y_u, z_u)$ be the difference between the wrist and elbow positions, $\mathbf{p}_u = \mathbf{p}_w - \mathbf{p}_e$, and $\mathbf{p}_c = (x_c, y_c, z_c)$ the difference between the shoe position and the elbow, $\mathbf{p}_c = \mathbf{p}_s - \mathbf{p}_e$. The user angle, $\theta_s^u(x_u, y_u)$, and corrected angle, $\theta_s^c(x_c, y_c)$, are then computed the same as in (5) and (6), by substituting $x_a^{(k)}$ and $y_a^{(k)}$ with x_u and x_c and y_e , respectively.

The parameters A and B can now be computed from these two sets of angles applying a linear regression model defined in (8). The same equation will be used to correct the user’s pointing angle during the experiments. The corrected pointing target in the shoe plane, $\mathbf{p}_t^c = (x_t^c, y_t^c, z_t^c)$, is then computed using the polar coordinates with the user’s elbow joint, \mathbf{p}_e , as the origin. The distance of the corrected pointing target is given by

$$d_t^c = \sqrt{(x_t^c - x_e)^2 + (y_t^c - y_e)^2}. \quad (9)$$

Finally, the corrected pointing target coordinates can be computed

$$\begin{aligned} x_t^c &= d_t^c \cos \theta^c \\ y_t^c &= d_t^c \sin \theta^c \\ z_t^c &= h \end{aligned} \quad (10)$$

Algorithm 2 Robot Position Adjustment

```

1: robot_pos  $\leftarrow \text{initial\_pos}$ 
2: while  $\neg$  “ok” do // user says “ok”
3: if direction then // user says direction
4: adjustment_direction  $\leftarrow \text{direction}$ 
5: while  $\neg$  “stop” do // user says “stop”
6: robot_pos  $\leftarrow \text{robot\_pos} + \text{adjustment\_direction}$ 
7: end while
8: end if
9: end while
10: initial_pos  $\leftarrow \text{robot\_pos}$ 

```

where h is the height of the shoe platform. It is important to note that the pointing calibration algorithm corrects the accuracy of the user, but not the precision. Hence, the efficiency of the pointing calibration depends on the individual user’s consistency in performing the pointing gestures.

In the experiments in which only the speech modality was used, no calibration was required, so the fitting parameters were set to $A = 1$ and $B = 0$, such that $\theta^u = \theta^c$. Hence, no correction of the pointing target was performed.

B. Robot Position Adjustment

A predefined shoe-delivery position may not fit all the users as it may require an additional effort to place the foot inside the shoe. To reduce the user workload, particularly the physical effort, a *robot position adjustment* algorithm is proposed. The algorithm takes user requests to modify the distance (in the xy plane) and the height (along the z -axis) of the robot end-effector from the ankle joint, as shown in Fig. 8. The following requests given by voice are defined: “move forward,” “move back,” “move up,” and “move down.” The procedure of the position adjustment is described in Algorithm 2. The robot modifies the end-effector position along the requested direction until the user says “stop.” The modification in any direction can be repeated until the user is satisfied with the final position and confirms it by saying “that’s ok,” or the end-effector reaches a safety limit ($d_{xy} = (0.2 \text{ m}, 0.6 \text{ m})$, $d_z = (0 \text{ m}, 0.5 \text{ m})$). The robot position adjustment can be performed as many times as needed, however, for the purpose of this paper it was performed only once.

Both pointing calibration and adjusted robot end-effector position were associated with a particular user and recorded for future dressing tasks, until changed again on user request. The Kinect 2 allows skeleton recognition and tracking of up to six users in the sensor’s field of view. Each user’s skeleton information has an associated *userID* that can be used to consistently recognize and track a specific user; in our case, this was the user closest to the robot.

IV. EXPERIMENTS

The developed autonomous robot dressing assistant was tested in experiments with users who had no experience in robotics. The robot’s task was to pick and deliver a shoe to the user’s right foot. However, in each trial the participants were required to repeat this task with the robot twice, in order to increase the level of difficulty to the one of the real dressing

task. The experiments were designed to evaluate performance and user workload under different conditions. The following sections describe the experimental setup, tasks, user profiles and evaluation metrics.

A. Experimental Setup

The proposed experimental setup consisted of a WAM robot with gripper, Kinect 1 and Kinect 2 cameras, and a platform on which the shoes were placed, as shown in Fig. 2. Two pair of shoes were used marked with four colors: 1) blue; 2) green; 3) red; and 4) yellow. The distance between the shoes was 0.2 m. The Kinect 1 was positioned above the platform facing downward to allow the visual recognition of the colored shoe markers. Its location in the robot reference system was $(x, y, z) = (0.38 \text{ m}, 0.07 \text{ m}, 1.16 \text{ m})$, and its orientation given by its Euler angles was $(\alpha, \beta, \gamma) = (139^\circ, 80^\circ, 37^\circ)$. The Kinect 2 was used to recognize speech, and track the user movements. It was placed in front of the user, at an angle that prevents occlusion of the foot by the WAM robot during dressing. Its position in the robot reference system was set to $(x, y, z) = (2.03 \text{ m}, 0.57 \text{ m}, 0.53 \text{ m})$ and its orientation in the Euler angles was $(\alpha, \beta, \gamma) = (0^\circ, 0^\circ, 121^\circ)$. The entire system was manually calibrated to minimize the robot positioning errors. The manual measurements were verified by visualizing the scenario in the ROS framework, through Rviz. The user was seated on a wheeled platform, allowing the distance from the robot to be adjusted. However, two constraints were considered: the user had to remain inside the detection range of the Kinect 2 camera and the right foot, when extended, had to be inside the robot's workspace.

The dressing task consists of the following steps, which were provided as instructions to the users involved in the experiments.

- 1) *Start*: The robot is in the home position and after the user's "begin" confirms with *ready to help*.
- 2) *Shoe Selection*: The user selects one of the available shoes, either by pointing to the shoe and saying *take this shoe* or using a voice command to specify the shoe's color, for instance "take the green shoe."
- 3) *Choice Correction*: If the robot picks up a wrong shoe, the user can correct it by repeating the first step.
- 4) *Shoe Delivery*: The dressing is initiated by the voice command *dress me*. The robot waits for the user to extend the right foot (the posture is recognized), after which it approaches the user's foot at a safe distance, taking into account the orientation of the user's ankle and knee joints (see details in Section II-C7).
- 5) *Finish*: The robot follows the user's foot while maintaining the safe distance until the user says *stop*. The user can now safely place the foot inside the shoe. The task finishes when the user says *that's ok*, after which the robot releases the shoe from the gripper and returns to the home position.

Fine shoe fitting by the robot may be added to finalize the dressing task, however, due to its complexity it is not considered in this paper but as a part of future work.

B. Participants

The robot assistant was evaluated in experiments with 12 participants (8 males and 4 females) of similar educational level (six electrical engineers, three computer scientist, two chemist, and one biologist) and age (between 22 and 29), with no experience in robotics. The goal of the experiments was to assist the participant with selecting and putting on a shoe. To add complexity to the task, the participants were asked to select two shoes from the set, the blue and green one, to complete the task. The difficulty of choosing each shoe depended on its distance from the user and the pointing angle required to select it, so for a fair comparison, all the users were asked to choose the same shoes.

To evaluate the effect of personalization on robot performance and user workload, the participants were divided into two groups of six participants, each group consisting of two female and four male participants. The participants from the Group 1 performed the task with the default robot setup, i.e., without personalization. The participants from the Group 2 were asked to perform the pointing calibration and robot position adjustment (described in Section III) before performing the dressing task. In both groups, the order of experiments was changed for subgroups of three participants for counter-balancing.

To study the effect of robot personalization on the type of interaction modality, both groups of participants performed two experiments. In the first experiment, only the use of voice commands was allowed in selecting the shoes, while in the second experiment a combination of pointing and speech (deictic expression) was required to make a selection. Each experiment consisted of five trials, in each of which the user was asked to select and put on two shoes.

C. Evaluation

Several metrics were used to evaluate the performance of the robot and the workload of the participants. The quantitative metrics used to evaluate the performance were the task success, task completion time, and number of corrections. Task success is defined by

$$S_i = N_i/2 \cdot 100\%, \forall i = 1 \dots 10 \quad (11)$$

where N_i represents the number of successfully delivered shoes, and i is the number of the trial. Task completion time is defined as the overall duration of a single trial. The number of corrections refers to the number of times the participant must repeat the request to the robot because it grasped the wrong shoe.

For a qualitative evaluation, the participants were asked to fill in the raw NASA-TLX questionnaire after each experiment. The questionnaire evaluates six dimensions of user workload: 1) mental demand; 2) physical demand; 3) temporal demand; 4) performance; 5) effort; and 6) frustration, values from 0 to 100 [31]. The overall workload is computed as the average of the above-mentioned six dimensions.

A mixed ANOVA test was conducted using the personalization condition as a between-subject factor, and the interaction modality as a within-subject factor divided in two levels

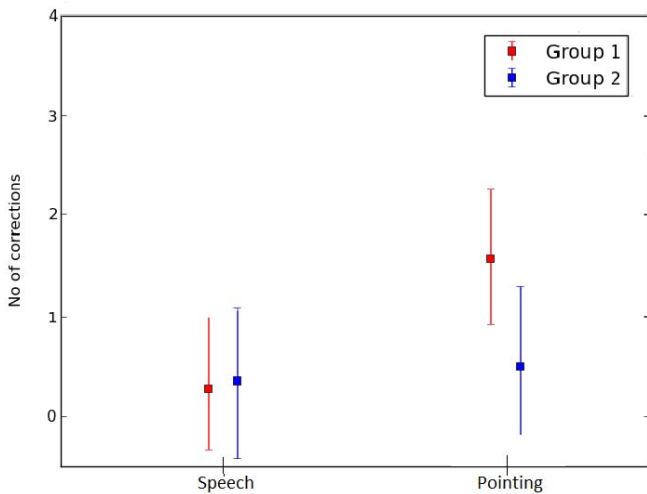


Fig. 9. Effect of the interaction modality and robot personalization on the number of corrections during shoe selection. Error bars represent the standard deviation of the mean. *Note:* Group 1—without personalization and Group 2—with personalization.

(speech and pointing/deictic). Statistical significance was computed for all the above-mentioned performance metrics. The results were considered significant for $p \leq 0.05$.

V. RESULTS AND DISCUSSION

For a total of 120 trials performed by 12 participants, 97.5% were successfully accomplished. In three trials that were classified as failures, the participants successfully guided the robot during the shoe selection and delivery, but failed to firmly place their foot inside the shoe, which resulted in the shoe being dropped on the ground. This suggests that the task was relatively easy to perform regardless of the interaction modality used to perform the shoe selection, and whether the robot personalization was performed or not.

Nevertheless, both type of modality and robot personalization condition influenced the task performance. The results of the ANOVA test show that there was a statistically significant effect of the type of modality on the average number of corrections, $F(1, 10) = 5.022$, $p = 0.049$. Furthermore, the pointing calibration reduced the number of corrections in the Group 2 by 79.2% compared to the results obtained by the Group 1, as shown in Fig. 9. The difference between the groups was statistically significant as determined by the ANOVA test ($F(1, 10) = 10.011$, $p = 0.01$). In fact, the Group 2 reported a similar number of corrections for both modalities, meaning that after calibration, the use of pointing gestures was as accurate as speech.

The effects of the interaction modality and robot personalization on task completion time are shown in Fig. 10. It can be noted that the task completion time was approximately the same in both groups when the speech modality was used. However, as a result of personalization when the pointing modality was used, the Group 2 required on average 23.3% less time than the Group 1 to complete the task. It can also be noted that for the Group 2, the task completion time was similar regardless of the modality used. On the contrary, the

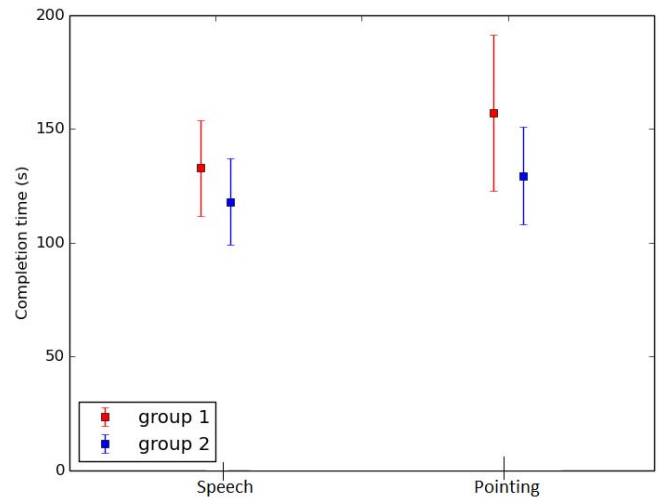


Fig. 10. Effect of the interaction modality and robot personalization on the task completion time. Error bars represent the standard deviation of the mean. *Note:* Group 1—without personalization, Group 2—with personalization.

Group 1 on average performed the task 24.2% slower with pointing than when the speech was used, indicating that pointing was less accurate without previous calibration. Although the ANOVA test results did not demonstrate statistically significant effect of the type of modality on the average task completion time, the effect of personalization was statistically significant, $F(1, 10) = 4.945$, $p = 0.05$.

The results for the six dimensions of user workload obtained with the NASA-TLX questionnaires are shown in Fig. 11. The type of modality had statistically significant effect on the user physical demand ($F(1, 10) = 5.248$, $p = 0.045$) and user performance ($F(1, 10) = 4.817$, $p = 0.053$). On average, the Group 2 who performed robot personalization experienced less overall workload than the Group 1: 3.2% when using speech, and 5.4% when pointing was used; however, the effect of personalization on user workload was not proved statistically significant by the ANOVA test. It should be noted though, that the user satisfaction analysis would be more reliable over a long-term interaction study that would also include a larger number of participants. For example, the pointing calibration and robot position adjustment may add both physical and mental demand to some users in a short experiment since they increase its complexity, but would prove beneficial over a longer period of interaction.

Though some of the results did not prove statistically significant, they are here presented to describe the behavioral trend of the participants. In comparison with the Group 1, the Group 2 experienced less physical demand (5.0% with pointing), temporal demand (5.0% with speech and 8.3% with pointing), and frustration (10.0% with pointing). The personalization performed by the Group 2 also led to a better performance (19.1% with speech and 8.3% with pointing). Although the personalization had no statistically significant effect on the level of user effort, it can be noted that the pointing modality required approximately 10% higher effort than speech, for both groups. Indeed, pointing was combined with speech to form deictic expressions, therefore, the final effort is expected to be higher.

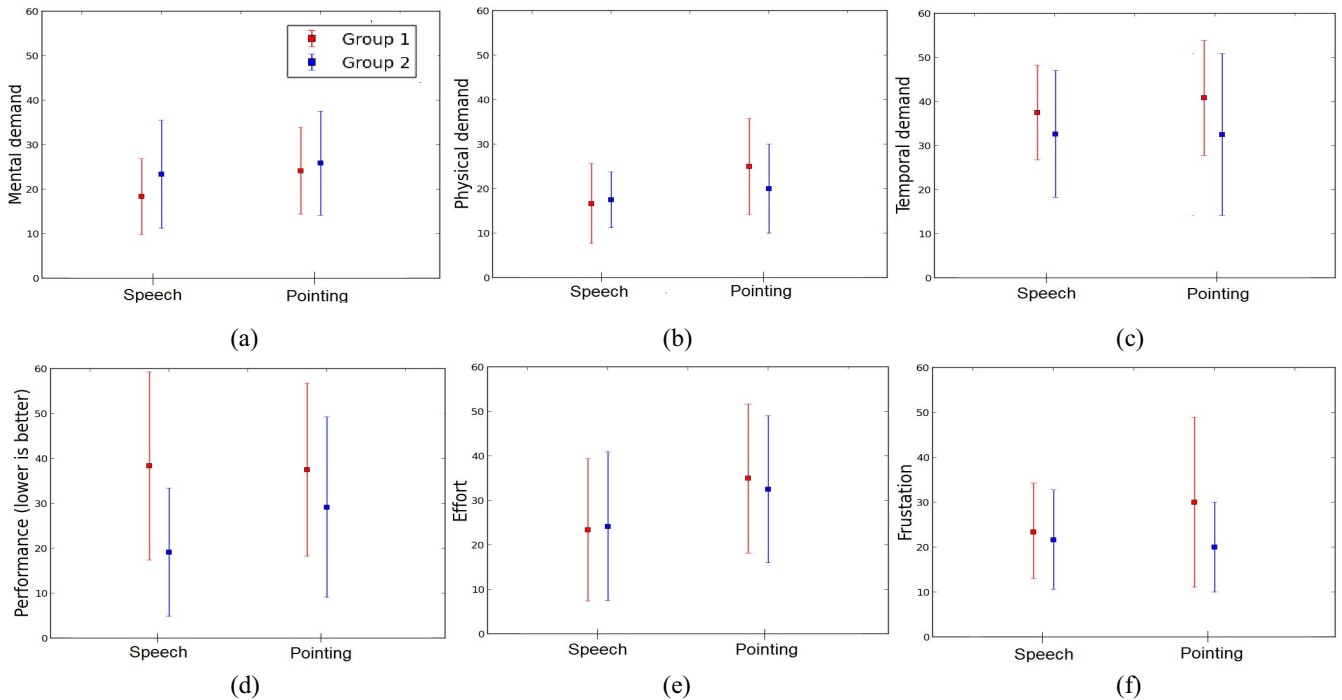


Fig. 11. Effect of the interaction modality and robot personalization on the six dimensions of workload. Note that here higher performance values indicate worse performance. Error bars represent the standard deviation of the mean. Note: Group 1—without personalization, Group 2—with personalization.

VI. CONCLUSION

Multiple modalities can add diversity and expressive power to HRI, but also result in a higher level of engagement that could positively impact the user’s level of concentration on the task and thus reduce errors or safety concerns from distraction, loss of interest or even boredom. This can be of high importance for the users that require assistance with the ADL such as dressing. For example, pointing can be used to make more precise requests if speech proves limited when choosing from a pile of similar shoes. A combination of modalities can have synergistic benefits, as in the case of deictic expressions. Also, more specifically, redundancy in the input to the system can improve accuracy. For example, speech-recognition in a noisy environment will be error prone.

In this paper, we exploited the concept of multimodality to develop personalized interaction with a robot assistant for support in dressing. The robot was able to adapt to the users’ individual requirements by performing pointing calibration and gripper position adjustment, which allowed more accurate shoe selection and more comfortable shoe positioning. It is important to note that the implementation of the robot personalization could be modified to improve its flexibility. First the system could adapt while performing the dressing task. For example, the user would be encouraged to point to a specific shoe or garment and vocalize the specific name. Given that the location of the user and the shoe are known in real time, the correction of the pointing target could be determined in this real scenario rather than a separate calibration routine. Second, a simple geometric model of the user could be implemented that adapted the correction angle with movement of the user or the garment, overcoming issue

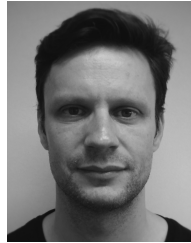
with linear mapping. However, in the scenario proposed in this paper, we are considering users in a seated position for which the linear mapping of the pointing targets proved suitable.

Even though adding modalities to the robotic system increases its complexity, in both system development and evaluation, our results showed that the robot was able to successfully perform the dressing task while reducing the overall user workload, as a result of personalization. Future work will include development of a framework that can intelligently manage the use of interaction modalities in each interaction event and transitions between them.

REFERENCES

- [1] J. E. Cohen, “Human population: The next half century,” *Science*, vol. 302, no. 5648, pp. 1172–1175, 2003.
- [2] G. A. Warshaw *et al.*, “Functional disability in the hospitalized elderly,” *J. Amer. Med. Assoc.*, vol. 248, no. 7, pp. 847–850, 1982.
- [3] M. E. Pollack, “Intelligent technology for an aging population: The use of AI to assist elders with cognitive impairment,” *AI Mag.*, vol. 26, no. 2, pp. 9–24, 2005.
- [4] H. Robinson, B. MacDonald, and E. Broadbent, “The role of healthcare robots for older people at home: A review,” *Int. J. Soc. Robot.*, vol. 6, no. 4, pp. 575–591, 2014.
- [5] S. Satake *et al.*, “How to approach humans?—Strategies for social robots to initiate interaction,” in *Proc. 4th ACM/IEEE Int. Conf. Human Robot Interact. (HRI)*, 2009, pp. 109–116.
- [6] T. Tamei, T. Matsubara, A. Rai, and T. Shibata, “Reinforcement learning of clothing assistance with a dual-arm robot,” in *Proc. 11th IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, Bled, Slovenia, 2011, pp. 733–738.
- [7] N. Koganti, T. Tamei, T. Matsubara, and T. Shibata, “Estimation of human cloth topological relationship using depth sensor for robotic clothing assistance,” in *Proc. Conf. Adv. Robot.*, Pune, India, 2013, pp. 1–6.

- [8] N. Koganti, T. Tameim, T. Matsubara, and T. Shibata, "Real-time estimation of human-cloth topological relationship using depth sensor for robotic clothing assistance," in *Proc. 23rd IEEE Int. Symp. Robot Human Interact. Commun. (RO-MAN)*, Edinburgh, U.K., 2014, pp. 124–129.
- [9] L. Twardon and H. Ritte, "Active boundary component models for robotic dressing assistance," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Daejeon, South Korea, 2016, pp. 2811–2818.
- [10] T. Matsubara, D. Shinohara, and M. Kidode, "Reinforcement learning of a motor skill for wearing a T-shirt using topology coordinates," *Adv. Robot.*, vol. 27, no. 7, pp. 513–524, 2013.
- [11] G. Chance, A. Jevtić, P. Caleb-Solly, and S. Dogramadzi, "A quantitative analysis of dressing dynamics for robotic dressing assistance," *Front. Robot. AI*, vol. 4, no. 13, pp. 1–14, 2017.
- [12] A. Colomé, A. Planells, and C. Torras, "A friction-model-based framework for reinforcement learning of robotic tasks in non-rigid environments," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Seattle, WA, USA, 2015, pp. 5649–5654.
- [13] 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 *Proc. 6th IEEE Int. Conf. Biomed. Robot. Biomechatronics (BioRob)*, Singapore, 2016, pp. 774–780.
- [14] Y. Gao, H. J. Chang, and Y. Demiris, "User modelling for personalised dressing assistance by humanoid robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Hamburg, Germany, 2015, pp. 1840–1845.
- [15] S. D. Klee *et al.*, "Personalized assistance for dressing users," in *Proc. 7th Int. Conf. Soc. Robot. (ICSR)*, vol. 9388. Paris, France, 2015, pp. 359–369.
- [16] K. Yamazaki, R. Oya, K. Nagahama, K. Okada, and M. Inaba, "Bottom dressing by a dual-arm robot using a clothing state estimation based on dynamic shape changes," *Int. J. Adv. Robot. Syst.*, vol. 13, no. 5, pp. 1–13, 2016.
- [17] E. Pignat and S. Calinon, "Learning adaptive dressing assistance from human demonstration," *Robot. Auton. Syst.*, vol. 93, pp. 61–75, Jul. 2017.
- [18] A. Kapusta *et al.*, "Data-driven haptic perception for robot-assisted dressing," in *Proc. 25th IEEE Int. Symp. Robot Human Interact. Commun. (RO-MAN)*, New York, NY, USA, 2016, pp. 451–458.
- [19] F. Zhang, A. Cully, and Y. Demiris, "Personalized robot-assisted dressing using user modeling in latent spaces," in *Proc. Int. Conf. Intell. Robots Syst. (IROS)*, Vancouver, BC, Canada, 2017, pp. 3603–3610.
- [20] Z. Erickson *et al.*, "What does the person feel? Learning to infer applied forces during robot-assisted dressing," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Singapore, 2017, pp. 6058–6065.
- [21] W. Yu *et al.*, "Haptic simulation for robot-assisted dressing," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Singapore, 2017, pp. 6044–6051.
- [22] Y. Gao, H. J. Chang, and Y. Demiris, "Iterative path optimisation for personalised dressing assistance using vision and force information," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Daejeon, South Korea, 2016, pp. 4398–4403.
- [23] K. Yamazaki, R. Oya, K. Nagahama, and M. Inaba, "A method of state recognition of dressing clothes based on dynamic state matching," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Kobe, Japan, 2013, pp. 406–411.
- [24] J. Shotton *et al.*, "Real-time human pose recognition in parts from single depth images," *Commun. ACM*, vol. 56, no. 1, pp. 116–124, 2013.
- [25] J. M. Beer *et al.*, "'Commanding your robot' older adults' preferences for methods of robot control," in *Proc. Human Factors Ergonomics Soc. Annu. Meeting*, vol. 56, 2012, pp. 1263–1267.
- [26] K. Nickel and R. Stiefelwagen, "Visual recognition of pointing gestures for human-robot interaction," *Image Vis. Comput.*, vol. 25, no. 12, pp. 1875–1884, 2007, doi: [10.1016/j.imavis.2005.12.020](https://doi.org/10.1016/j.imavis.2005.12.020).
- [27] D. Droeschel, J. Stückler, and S. Behnke, "Learning to interpret pointing gestures with a time-of-flight camera," in *Proc. 6th ACM/IEEE Int. Conf. Human Robot Interact. (HRI)*, Lausanne, Switzerland, 2011, pp. 481–488, doi: [10.1145/1957656.1957822](https://doi.org/10.1145/1957656.1957822).
- [28] A. Jevtić, G. Doisy, S. Bodiřoža, Y. Edan, and V. V. Hafner, "Human-robot interaction through 3D vision and force control," in *Proc. ACM/IEEE Int. Conf. Human Robot Interact.*, Bielefeld, Germany, 2014, p. 102.
- [29] A. Jevtić, G. Doisy, Y. Parmet, and Y. Edan, "Comparison of interaction modalities for mobile indoor robot guidance: Direct physical interaction, person following, and pointing control," *IEEE Trans. Human-Mach. Syst.*, vol. 45, no. 6, pp. 653–663, Dec. 2015.
- [30] A. Colomé and C. Torras, "Closed-loop inverse kinematics for redundant robots: Comparative assessment and two enhancements," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 2, pp. 944–955, Apr. 2015.
- [31] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, Dec. 1988.



Aleksandar Jevtić received the B.S. degree in electrical engineering from the University of Belgrade, Belgrade, Serbia, in 2005, and the M.S. and Ph.D. degrees in computer science from the Technical University of Madrid, Madrid, Spain, in 2007 and 2011, respectively.

He was a holder of the Marie Curie and Beatrice de Pinós Post-Doctoral Fellowships. He is currently a Researcher with the Institute of Robotics and Industrial Informatics, CSIC-UPC, Barcelona, Spain. His current research interests include socially assistive robotics and HRI.



Andrés Flores Valle received the M.S. degree in artificial intelligence from the Technical University of Catalunya (UPC), Barcelona, Spain, in 2017, for the research he performed with the Institute of Robotics and Industrial Informatics, CSIC-UPC. He is currently pursuing the Ph.D. degree with the Center of Advanced European Studies and Research (CAESAR), Bonn, Germany

He is a member of Neural Circuits Research Group, CAESAR. His responsibilities included the implementation of the personal robot dressing assistant and the experimentation with users.



Guillem Alenyà (M'07) received the Ph.D. degree (Doctor Européen) from the Technical University of Catalunya, Barcelona, Spain, in 2007 with a research on robot navigation using active contours, which he partly developed at the Robosoft Company, Bidart, France, where he was supported by the EU-FP6 Marie-Curie Scholarship.

He holds a Researcher position with Spanish Scientific Research Council, Madrid, Spain. He has been a Visitor with KIT-Karlsruhe, Karlsruhe, Germany, in 2007, INRIA-Grenoble, Montbonnot-

Saint-Martin, France, in 2008, and Bristol Robotics Laboratory, University of the West of England, Bristol, U.K., in 2016. He has participated in numerous scientific projects involving image understanding, next-best-view, rule learning, and plan execution tasks. He is currently a local PI of the EU MSCA training network SOCRATES, a local co-PI of the EU project IMAGINE, and a Leader of numerous technological transfer projects. From 2009 to 2011, he served as the Head of the Robotics Department.



Greg Chance received the B.Sc. (Hons.) and Ph.D. degrees in physics from the University of Bath, Bath, U.K., in 2001 and 2005, respectively.

He is a Chartered Engineer and a member of the Institute of Physics. He is currently a Research Associate researching with the Bristol Robotics Laboratory, University of the West of England, Bristol, U.K. He has ten years industrial experience researching for Oxford Instruments, Abingdon, U.K., and BAE Systems, Farnborough, U.K.



Praminda Caleb-Solly received the graduation degree in electronic systems engineering from the University of Essex, Colchester, U.K., the M.Sc. degree in biomedical instrumentation engineering from Dundee University, Dundee, U.K., and the Ph.D. degree in interactive evolutionary computation from the University of the West of England, Bristol, U.K.

She is an Associate Professor with the Bristol Robotics Laboratory, University of the West of England, Bristol, U.K., and the Head of Electronics and Computer Systems at Designability, Bath, U.K., a not-for-profit organization designing assistive technology. She has over 18 years of research experience and her research portfolio comprises numerous national research council, government, and European funded projects. Her current portfolio includes projects on assistive robotics technology to support older adults with aging related impairments, intelligent data processing for smart home sensing, understanding children's use of powered mobility, and adaptive interfaces for driverless cars. She has also co-designed an app for self-management of anxiety, SAM. Her current research interests includes using participatory design methods to understand how to develop and evaluate multimodal interfaces that facilitate intuitive and safe interaction.



Sanja Dogramadzi received the graduation degree in mechanical and control engineering and the Ph.D. degree in medical robotics from the University of Newcastle, U.K.

She is leading Healthcare Robotics Group, Bristol Robotics Laboratory, University of the West of England, Bristol, U.K. She has led many research projects funded by the U.K. Councils and European Commission and is currently a Coordinator and the PI on H2020 funded Smartsurg Project, EPSRC Funded I-DRESS Project, and H2020 Funded Marie Curie ITN Socrates, and the Technical Lead for Innovate U.K. funded CHIRON Project. Her current research interests include design and control of surgical robots for minimally invasive fracture surgery and robot-assisted MIS, rehabilitation robots, physical assistance robots and safe, and ethical HRI.



Carme Torras (M'07–SM'11) received the M.Sc. degree in mathematics from the University of Barcelona, Barcelona, Spain, the M.Sc. degree in computer science from the University of Massachusetts at Amherst, Amherst, MA, USA, and the Ph.D. degree in computer science from the Technical University of Catalonia, Barcelona.

She is a Research Professor with the Spanish Scientific Research Council, and the Head of the Perception and Manipulation Group, Robotics Institute in Barcelona. She has published five books and about 300 papers in the areas of robot kinematics, computer vision, geometric reasoning, machine learning, and manipulation planning. She has supervised 18 Ph.D. theses and led 15 European projects, the latest being the Chist-Era project I-DRESS, and the H2020 project IMAGINE.

Dr. Torras was a recipient of the Narcís Monturiol Medal of the Generalitat de Catalunya in 2000, the ERC Advanced Grant for the project CLOTHILDE—CLOTH manipulation learning from demonstrations. She has been an Associate Vice-President for Publications of the IEEE Robotics and Automation Society (RAS), an Editor of the IEEE TRANSACTIONS ON ROBOTICS, and is currently an elected member of the Administrative Committee of IEEE RAS. She is an ECCAI Fellow in 2007, and a member of Academia Europaea in 2010 and the Royal Academy of Sciences and Arts of Barcelona in 2013. For more information: <http://www.iri.upc.edu/people/torras>.