Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population

A. Zlatintsi[a,*], A. C. Dometios[a], N. Kardaris[a], I. Rodomagoulakis[a], P. Koutras[a], X. Papageorgiou[a], P. Maragos[a], C. S. Tzafestas[a], P. Vartholomeos[b], K. Hauer[c], C. Werner[c], R. Annicchiarico[d], M. G. Lombardi[d], F. Adriano[d], T. Asfour[e], A. M. Sabatini[f], C. Laschi[f], M. Cianchetti[f], A. Güler[g,1], I. Kokkinos[h,1], B. Klein[i], and R. López[j]

[a] *Institute of Communication and Computer Systems (ICCS), National Technical University of Athens (NTUA), Greece*
[b] *OMEGA Technology, Athens, Greece*
[c] *Bethanien Hospital, Germany*
[d] *Fondazione Santa Lucia, Rome, Italy*
[e] *Karlsruhe Institute of Technology (KIT), Germany*
[f] *Scuola Superiore Sant'Anna (SSSA), Pisa, Italy*
[g] *Department of Computing, Imperial College London, UK*
[h] *Department of Computer Science, University College London, UK*
[i] *Frankfurt University of Applied Sciences, Frankfurt am Main, Germany*
[j] *Robotnik Automation, SLL, Valencia, Spain*

Abstract

In this paper we present a prototype integrated robotic system, the I-Support bathing robot, that aims at supporting new aspects of assisted daily-living activities on a real-life scenario. The paper focuses on describing and evaluating key novel technological features of the system, with the emphasis on cognitive human-robot interaction modules and their evaluation through a series of clinical validation studies. The I-Support project on its whole has envisioned the development of an innovative, modular, ICT-supported service robotic system that assists frail seniors to safely and independently complete an entire sequence of physically and cognitively demanding bathing tasks, such as properly washing their back and their lower limbs. A variety of innovative technologies have been researched and a set of advanced modules of sensing, cognition, actuation and control have been developed and seamlessly integrated to enable the system to adapt to the target population abilities. These technologies include: human activity monitoring and recognition, adaptation of a motorized chair for safe transfer of the elderly in and out the bathing cabin, a context awareness system that provides full environmental awareness, as well as a prototype soft robotic arm and a set of user-adaptive robot motion planning and control algorithms. This paper focuses in particular on the multimodal action recognition system, developed to monitor, analyze and predict user actions with a high level of accuracy and detail in real-time, which are then interpreted as robotic tasks. In the same framework, the analysis of human actions that have become available through the project's multimodal audio-gestural dataset, has led to the successful modelling of Human-Robot Communication, achieving an effective and natural interaction between users and the assistive robotic platform. In order to evaluate the I-Support system, two multinational validation studies were conducted under realistic operating conditions in two clinical pilot sites. Some of the findings of these studies are presented and analysed in the paper, showing good results in terms of: (i) high acceptability regarding the system usability by this particularly challenging target group, the elderly end-users, and (ii) overall task effectiveness of the system in different operating modes.

Keywords: Human-Robot Communication, Assistive Human-Robot Interaction (HRI), Bathing robot, Multimodal dataset, Audio-gestural command recognition, Online validation with elderly users

1. Introduction

Advanced countries with well organized and modern health care systems tend to become aging societies, according to World Health Organization's research on health and ageing [1]. The percentage of population with special needs for nursing attention (including people with disabilities) is significant and due to grow. Health care experts are called to support these people during the performance of Activities of Daily Living (**ADLs**) such as dressing, eating and showering, inducing great financial burden both to the families [2] and the caregivers [3]. Great research effort has been spent over the last decades [4, 5, 6, 7] on studying and classifying the functional disabilities of older adults and associating the latter with basic factors of morbidity and mortality.

Personal care (showering or bathing), which is crucial for a person's hygiene, is included among the first ADLs, which incommode an elderly's life [7] and ADL difficulties in bathing or showering represent the strongest predictor of subsequent institutionalization in older adults [8]. Older adults require assistance in bathing or showering more frequently than for any other ADL [9]. As bathing is a highly intimate ADL, the wish

*Corresponding author: Athanasia Zlatintsi, School of ECE, National Technical Univ. of Athens (NTUA), 15773 Athens,Greece, e-mail: nzlat@cs.ntua.gr
[1] During this work I. Kokkinos and R. A. Güler were working at the French Institute for Research in Computer Science and Automation (INRIA), France.

for independence from personal bathing assistance of caregivers as long as possible is, however, not unusual in older adults [10]. An assistive bathing robot may thus represent an opportunity for older adults with bathing disability to take care of themselves in bathing, to reserve their privacy, and reduce the burden of caregivers.

During the last decades, an enormous number of socially interactive robots have been developed constituting the field of Human-Robot Interaction (HRI) an actual motivating challenge. The robotics society is attempting to tackle this challenge of unattended nursing by developing flexible and modular assistive devices that aim to cover the needs for support of everyday tasks involved in the caring of frail people in both in-house [11] and clinical environments. Specifically, devices intended for this purpose involve either static physical interaction [12, 13, 14], or are mounted on mobile platforms [15, 16]. The focus of these solutions lies on a specific body part (e.g., the head) and the support of disabled people on the performance of ADLs with rigid manipulators. Furthermore, the literature reveals two commercial solutions presented in [17, 18], which provide a safe, independent showering experience and are equipped with soaping and water rinsing system. On the downside, both of these solutions completely lack physical interaction with the user and therefore lack some basic functionalities of the bathing sequence such as scrubbing and wiping the senior.

Direct physical interaction with frail seniors raises a multitude of issues, including safety, reliability for human-robot interaction and adaptability to the user's needs and preferences. Moreover, human body parts are curved and deformable and their size and shape differs a lot from one person to another. Unexpected body-part motion may also occur during the operation of the robot, increasing the risk of undesirable and possibly harmful contact between the human and the robot. Therefore, there are augmented human perception requirements, not only in terms of sensorial information adequacy and accuracy but also in terms of perception algorithms.

The first requirement cannot be addressed with simple proximity sensors and monocular cameras, since both of these solutions give poor sensorial feedback. Additionally, cameras intended for visual capturing with RGB data during the shower process, could raise some ethical issues in terms of privacy and comfort, especially when an elderly person's mental health deteriorates. On the contrary, the use of cheap depth or stereo vision cameras can provide rich information with good accuracy [19, 20] and fulfill the requirements of a great variety of applications [21], without necessarily the use of RGB data. The second requirement has actually led HRI to extend into other research areas [22]. One such area concerns the development of multimodal perception interfaces, required to facilitate natural human-robot communication, including not only visual sensors, but also audio or inertial sensors [23]. The concept behind these algorithms is to design interaction techniques that will enhance the communication making it natural and intuitive, enabling robots to understand, interact and respond to human intentions intelligently. For a review, we refer the reader to [24, 25, 26, 27, 28, 29].

Recently, researchers in computer vision proposed some approaches based on Deep Learning (DL) techniques, which have presented very detailed results on human perception and specifically body-part segmentation. The first fully-convolutional neural network (CNN) implementing semantic segmentation is [30] and many more works [31, 32, 33, 34] followed with detailed results, which are able to segment the human body parts at pixel level or get a sparse pose of the body parts [35, 36] with close to real-time performance.

Direct interaction of a robotic device with the environment is a research subject that the robotics society is addressing for many years. But the interaction with a human being is a much more delicate action and is considered risky to be executed by a rigid robotic manipulator, even if it is equipped with the most sophisticated force/impedance control schemes. On the other hand, the advantage of soft robots [37] lies on their inherent or structural compliance, which gives them the ability to actively interact with the environment and undergo large deformations [38]. The term soft robotics is not only used to state that the devices are made of soft materials, but also to underline the shift from robots with rigid links (even hyper redundant ones [39]) to bio-inspired continuum robots. Many continuum manipulators have already been presented [40] with different types of actuation, e.g., tendon based [41, 42] or a combination with pneumatic chambers [43, 44, 45]. The repertoire of actuation is not only important for the motion dexterity and shape [46] of a soft robot but also for stiffening [47, 48] and compliance, two properties that are crucial especially for physical interaction with a human.

Contributions and overview

The I-Support project envisioned, and during its course accomplished, the development and integration of an innovative, modular, ICT-supported robotic system that supports frail older adults' motion abilities. It successfully assists them to safely and independently complete various physically and cognitively demanding bathing tasks, such as properly washing their back and their lower limbs. The main contributions of the work presented in this paper relate to the development and seamless integration of novel cognitive human-robot interaction technologies and to the evaluation of these technologies, as individual modules as well as an integrated assistive robotic platform as a whole, through a series of clinical validation studies in realistic scenarios and under real operating conditions. These technologies include: human activity monitoring and recognition, adaptation of a motorized chair for safe transfer of the elderly in and out the bathing cabin, a context awareness system that provides full environmental awareness, as well as a prototype soft robotic arm and a set of user-adaptive robot motion planning and control algorithms. Key features of this assistive robotic platform, described in the paper, are supported by our state-of-the-art pipeline for multimodal modeling and learning which aims to enhance the human-robot communication making it natural, intuitive and easy to use, addressing aspects of smart assistive HRI. An important contribution of the work presented in this paper also concerns a new dataset that includes audio commands, gestures, which is an integral part of human communi-

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020. DOI. https://doi.org/10.1016/j.robot.2020.103451.

cation [49], and co-speech gesturing data, which is still quite limited in HRI [50], as well as a suite of tools used for data acquisition. The end goal of our work is to support and maximize safety, reliability and self-confidence for the frail users, as well as to enable intuitive and transparent HRI which incorporates reliable user intention recognition and real-time adaptation of the reactive and supportive robotics system's behavior. To that end, two multinational validation studies were conducted in two European clinical pilot sites for the evaluation of the various functionalities of the I-Support system in the bathroom environment of the selected care facilities. The validation studies included elderly subjects and were carried out under realistic conditions, showing good results and high acceptability by this particularly challenging target group.

The remainder of this paper is structured as follows: in Sec. 2 we describe the overall architecture of the I-Support system. In Sec. 3 we go into more detail regarding the developed online audio-gestural recognition system for human-robot communication, while in Sec. 4 we describe the adaptive robotic motion planning method. The dataset collected during the course of the project for modelling and offline evaluation is presented in Sec. 5. Finally, in Sec. 6 two multinational validation studies conducted in two European hospitals are described, showcasing promising results both regarding the system performance but also the user satisfaction.

2. System Architecture

The I-Support system is depicted in Fig. 1(b), presenting an installation in a clinical bathroom environment. The safety and operational requirements emerging from two use cases were taken into account. These use cases include two demanding washing tasks in terms of mobility and force exertion for the elderly, i.e. washing the back and the legs (lower limbs). Next, we describe the main modules of the system, namely: the motorized chair, the human-robot interaction module and the sensors used for communication, the context awareness system, the soft robotic arm and the overall software and process architecture.

2.1. Motorized chair

A motorized chair has been employed inside the shower to effectively assist the older adults during sit-to-stand and stand-to-sit tasks and for safely transfer from the exterior to the interior space of the shower cabin. The design that was followed was to adapt a commercially available chair due to the reduced costs in comparison to custom designs. The selected chair was adapted in accordance in order to provide 3DOF motion. More specifically, a lateral translation adjusts the proximity of the user to the robot and a vertical translation regulates the height of the chair from the ground. Additionally, the rotational DOF around its axis allows for easy accessibility in different bathroom environments.

2.2. Human-Robot Interaction

For the purpose of human-robot communication and perception, the I-Support system was equipped with Kinect V2 RGB-D cameras. These sensors are frequently used for visual analysis in assistive robotics [51], [52], [53] as they are inexpensive, reliable and simple to waterproof. They are also easy to mount on different surfaces and integrate software-wise. Three Kinect sensors were placed in the bathroom as shown in Fig. 1(b). This multi-view setup was designed so as to be able to deal with the two main technological tasks of the I-Support project, namely: a) audio-gestural command and action recognition and b) body pose estimation for robotic manipulation, considering various constraints (i.e. the size of the bath cabin, the size and the placement of the chair and the soft-arm robot base). For body pose estimation, it is required to have two different data streams for the analysis of the tasks under investigation (i.e. washing the legs and washing the back), which are RGB and depth. These two streams should be registered, in other words they have to be calibrated and aligned, to allow simultaneous processing. For the second task, i.e. (audio-)gestural and action recognition, it is essential to have the RGB stream in High Definition (HD) format, in order to retain the region of interest (i.e. hand gestures or face) with acceptable resolution.

To tackle the above mentioned tasks, we have employed three Kinect V2 sensors that can provide all required visual and audio information. These sensors can capture Full HD RGB video at 30fps (frames per second), while the depth information is recorded using the infrared camera embedded in the Kinect in standard resolution. The color stream is captured in BGRA format of total 32bpp (bits per pixel) resulting in an uncompressed image of about 8MB, while for the depth information the same format of total 16bpp is employed (ca. 800KB). For the audio/spoken command recognition task we experimented with the audio stream that can be captured by the built-in multi-array microphone of the Kinect sensors. Specifically, each Kinect incorporates an array of four individual microphones and the raw audio information can be captured at 16000 Hz, with 32-bit resolution. Figure 2 shows examples of the data streams that are acquired from the Kinect sensors.

In the early stages of system integration process the placement of the Kinect sensors inside the bathroom was a challenging task. The constraints of the camera placement were the capturing of all significant tasks involving HRI, while at the same time satisfying the space limitations imposed by a conventional bathroom in nursing homes. After extensively experimenting with all possible positions the resulting sensor set-up was able to capture the necessary information for both tasks, i.e. information of the user's back or legs for robotic manipulation, as well as the hand gestures performed for communication with the robot. Two of the sensors (Kinect sensors 1 and 2) were placed inside the bath cabin, in order to capture the legs and the back of the user during the different tasks, while a third camera was placed outside the cabin, in order to capture the gestures performed by the user during the task washing the back.

Specifically, during the task washing the legs Kinect 2 recorded the user's legs (including registered RGB and depth in SD res-

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

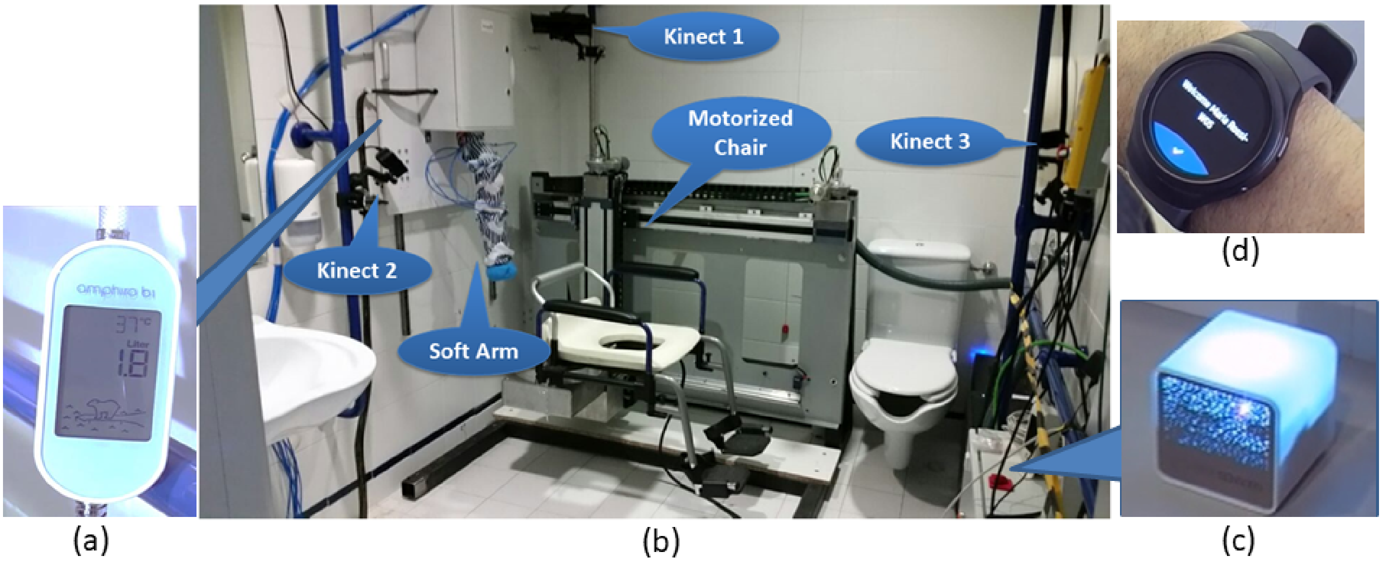


Figure 1: Installation of the I-Support system in clinical environment for experimental validation. The devices constituting the overall system are presented. (a) Amphiro b1 water flow and temperature sensor. (b) General aspect of the system showing the Motorized Chair, the Soft Robotic Arm and the installation of the Kinect sensors (for audio-gestural communication). (c) Air temperature, humidity and illumination sensors by CubeSensors. (d) Smartwatch for user identification and activity tracking.

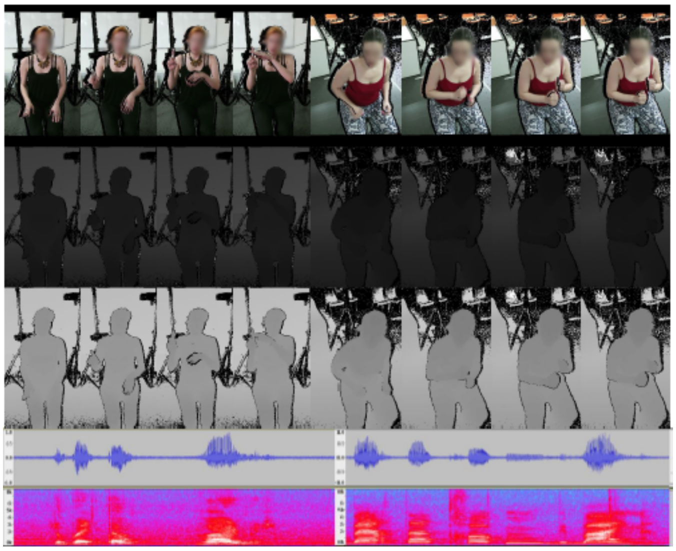


Figure 2: Data streams acquired by Kinect 1 and 3: RGB (top), depth (2nd row) and log-depth (3rd row) frames from a selection of gestures ("Temperature Up", "Scrub Legs"), accompanied by the corresponding German spoken commands waveforms (4th row) and spectrograms (bottom row) "Roberta, Wärmer", "Roberta, Wasch die Beine".

olution), for body pose estimation and visual tracking of the robot; while Kinect 1 was used by the audio-gestural and action recognition module. Except for the streams in SD resolution, sensor 1 also recorded the color stream (RGB) in Full HD resolution. In this configuration no video information was captured by Kinect 3. During the task washing the back, Kinect 1 recorded the back of the user (captured information includes RGB and depth in SD resolution), while Kinect 3 recorded the color stream, used for gesture recognition, in Full HD. In this case, Kinect 2 was not required for capturing any data.

2.3. *Context Awareness System (CAS)*

In order to achieve a proper operational flow during a showering task, it is important for the system not only to be able to perceive and communicate with the user, but also to have full environmental awareness. The latter can be achieved with the aid of extra sensorial data coming from different types of sensors. To begin with, Amphiro sensors, depicted in Fig. 1(a), can provide useful data to the system regarding water flow and temperature. Additionally, air temperature, humidity and illumination are environmental conditions indicative for the user's safety and comfort. These values are obtained by sensors constructed by CubeSensors, shown in Fig. 1(c).

A smartwatch similar to the one presented in Fig. 1(d) is integrated for user identification and activity tracking purposes [54]. The identification process includes a data acquisition step in which the user's personal data (e.g. gender, age, size) and preferences (e.g. showering duration, scrubbing patterns, body parts to avoid etc.) are stored in a database. The identification step implemented with the smartwatch includes a bar-code scanning, through which the personalization data are passed to the system and are taken into account for the bathing procedure. Furthermore, activity tracking algorithms are employed to recognize falls or inactivity time periods, which are indicative for emergency situations. The above mentioned data are available not only to the I-Support system but also to the nursing staff via an Android application, for monitoring and acting while emergency situations emerge.

2.4. *Soft Robotic Arm*

A soft-arm has been developed to assist elderly people in bathing tasks. Soft manipulators can be considered intrinsically safe thanks to the actuation technologies they are made of. One of their main features is their compliant body that can deform passively to adapt to environment changes, thus reducing

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

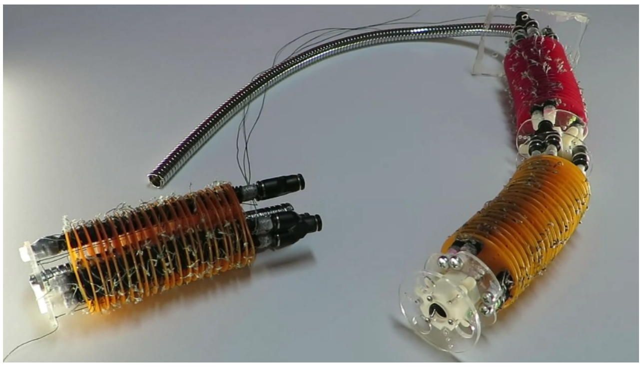


Figure 3: Construction stage of the Soft Robotic Manipulator. The modular design provides the flexibility to modify the robot's characteristics according to the motion requirements.

the complexity of active control. The technical requirements, in terms of desired reachable workspace and expected movements have been guaranteed by a hybrid actuation strategy, as summarized in [43]. Here, we briefly recap the arrangement of the actuators in the modules (Fig. 3). Two different actuation technologies are combined based on their working principle, in order to achieve the desired motion behavior:

- Flexible fluidic actuators, which enable elongation and omni-directional bending, exhibiting low accuracy;
- Tendon-driven mechanisms, which can shorten the mechanism and provide (redundant) omni-directional bending with high movement resolution.

As a result, the system is capable of elongation/contraction and bending movements, exploring a workspace compatible with different activation patterns, while specific antagonistic activation sequences provide stiffness capabilities as well.

2.5. *Software Architecture and Operational Flow*

In order to accommodate the hardware devices mentioned above along with the required real-time processing of the data, we have designed and implemented a system that uses Linux operating system in order to be able to handle multiple and fully synchronized Kinect sensors in modular configurations. With this architecture, the acquisition and the data processing is made using the Robot Operating System (ROS) using a different Linux machine for each camera. For the ROS setup we have used a master-client network approach, where a master computer system controls all the other hosts where the various sensors are connected.

Integration of the software nodes was accomplished in three stages: (i) unit testing of ROS free software (i.e. mathematical functions, non-ROS libs etc.), (ii) unit testing of ROS nodes and creation of ROS packages, and (iii) integration of all packages and testing that they all work together according to specs (topics, services, actions, loop rates). It was demonstrated that ROS-network set up and data connectivity function successfully. The process for remote launching of the ROS nodes was almost entirely automated through nested launch files.

2.5.1. *Operational flow and Finite State Machine (FSM)*

The operational flow of I-Support is composed by the robotic task sequence of the washing activities. It also includes the initialization, the termination and the error handling actions of the I-Support system. The robotic task sequence for the showering process has been derived by conducting a survey about the senior users' preferences on shower-activities, based on questionnaires that were answered by the seniors and by the caregivers. The outcome of the survey indicated that the critical body areas that should be washed by I-Support system are the back of the user, the private parts, and the legs of the user. For the pilot studies the consortium decided to test and validate the washing of the legs and of the back of the user. Furthermore, the information gathered by the end-users indicated that the washing activities should be broken down into *washing, rinsing* and *scrubbing*. Also, it indicated that the Seniors would like to use a small set of commands that will allow them to intuitively interact with the robotic system and synthesize a complete sequence of shower activities including actions such as start, stop, pause and repeat the procedure. To this end, the following minimal set of audio-gestural commands was defined: $\{wash-back, wash-legs, scrub-back, scrub-legs, rinse-back, rinse-legs, start, halt, stop, repeat\}$.

The entire I-Support process is modeled as a sequence of states and is supervised by a finite state machine (FSM) which has been developed and modeled as a directed graph. Each node of the graph corresponds to a state (for example washing, rinsing, scrubbing), and each directed edge corresponds to an event that triggered a change of state and optionally some associated action (for example the user requests through audio-gestural commands to repeat or stop the washing). The FSM manages the showering process by monitoring the progress of each state and the generation of the state sequence until the termination of the process. The actual sequence of states is not fixed, instead it is flexible and varies depending on the external triggers. The triggering events are generated either by the HRI module (see Sec. 3), or by the robotic motion planner or by the Chair-Controller; based on the triggering events different states sequences can be synthesized by the end-user. For example, an indicative set of state sequence is: IDLE, CHAIR-IN, DOUSING, SCRUBBING, SCRUBBING-PAUSED, RINSING, WASHING-ENDED, CHAIR-MOVING-OUT. In between these states (which represent washing activities) there are intermediate states that represent robot kinematic transitions to a starting pose required for the initiation of the next activity. The described sequence is depicted in Fig. 4. Any emergency situations are handled by the emergency stop state, to which all states can transition. The state washing-ended is also accessed by all states in order to facilitate handling of errors, such as the inability of the user to complete the washing procedure or the abrupt wish to terminate the process.

The FSM was implemented in ROS using the Smach package, which is a powerful and scalable Python-based library for hierarchical state machines. The Smach package does not depend on ROS and can be used in any Python project. The executive Smach stack however provides seamless integration

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

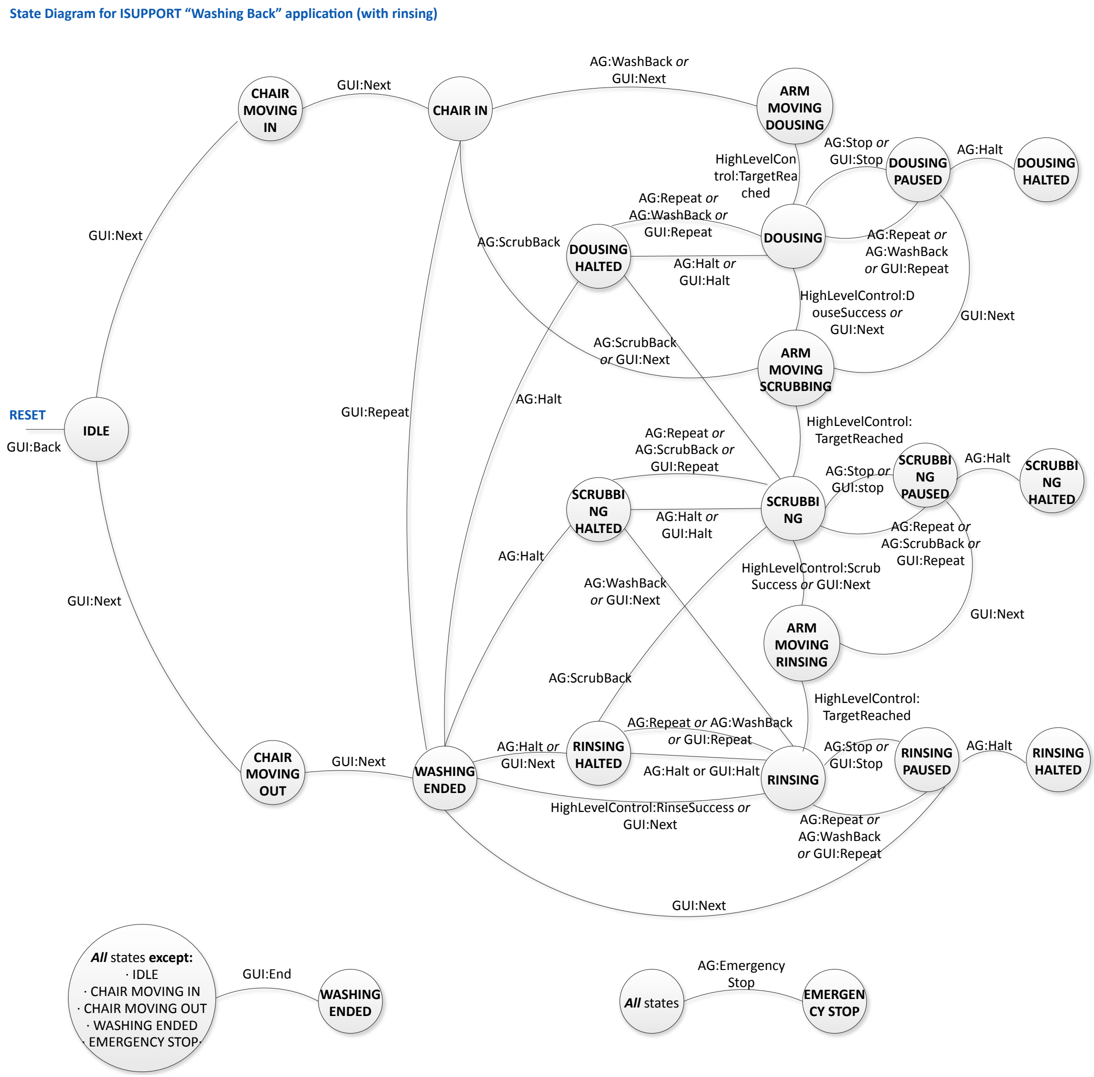


Figure 4: Example of I-Support FSM state diagram for washing back application.

with ROS, including smooth actionlib integration and a Smach viewer to visualize and introspect state machines. Implementation wise, all states are stored in the generic state machine container. They are initialized and when triggered by an event they execute the action code stored in the state. Each has a number of possible outcomes associated with it. An outcome is a user-defined string that describes how a state finishes. The transition to the next state is specified based on the outcome of the previous state.

The Finite State Machine (FSM) was personalized, thus, depending on the user, it could be automatically modified in order to meet the particular needs and requirements (i.e. size, gender, preferences, medical condition). This was achieved by interfacing the FSM with the Context Awareness System (CAS) and the personalised information that was stored in the I-Support system database. The FSM also performed the launching of the ROS nodes (depending on the current state) and the monitoring (running, success, fail) of each active node.

2.6. *Safety validation*

I-Support is a hardware device that comes into physical contact with humans, either on purpose during scrubbing and washing tasks or accidentally as it describes trajectories in the shower workspace. Furthermore, it is an electrical device that operates in a humid and wet environment subjected to jets of water. It might operate in a healthcare environment, such as a hospital, where the levels of noise and electromagnetic noise should be kept at a minimum. Hence, it was necessary to take actions for thorough hazard mitigation and for a comprehensive safety validation of the I-Support system during the design process.

The methodology that was adopted to address the safety issues involved the following steps: (i) analysis of the international safety standards and regulations and selection of the most relevant standard, (ii) classification of the I-Support device according to that standard, (iii) hazard analysis of the I-Support system (identification of hazards, identification of causes associated with hazards, determination of degree of risk), (iv) hardware and software design decisions to mitigate hazards (such us upper limits on acceleration and speed, minimization of weight, use of soft materials, DC voltages below 25V, EC compliance, emergency stops, controlled system shutdown, etc.), (v) testing and validation of the I-Support components and of the integrated system by an external certified safety engineer, (vi) safety validation of the complete I-Support installation in the operating environment where the pilot studies took place.

The ISO 13482 was selected as the most relevant safety guideline to be followed throughout the project development. The I-Support robotic solution was classified as a personal care robot, restraint-free assistance robot (to help provide more ease and comfort in daily life for independent living). Hence, the comprehensive description of the hazard-risk analysis, the subsequent design decisions, and the official safety validation results all complied with the ISO 13482 safety requirements. The hazard analysis, risk analysis and design decisions are all presented in detail in [D2.3, Annex 1, http://www.i-support-project.eu/dissemination/]. Finally, an external safety officer was employed to perform systematic tests and inspections and verify that the hardware and software implementation for all components, as well as for the overall installation in the healthcare environment comply with the ISO 13482 standards and safety requirements.

3. Audio-Gestural HRI

Gesture recognition is a visual task that can aid in human-robot interaction (HRI) and relies on similar visual processing methods as in action classification and pose estimation. For the I-Support system, we implemented a simple rule for gesture recognition that involves the recognition of a vocabulary of gestures, which serve the role of visual non-verbal commands. For robustness we supplement gesture recognition with the additional perceptual task of spoken command recognition. The two modalities can work either independently or in fusion for enhanced performance. Within this context of assistive robotics, and by exploring state-of-the-art approaches from automatic speech recognition and visual action recognition, we have developed an intelligent interface that multimodally recognizes actions and commands by fusing the unimodal information streams to obtain the optimum multimodal hypothesis.

3.1. Visual processing

Gesture recognition allows the interaction of the elderly users with the robotic platform through a predefined set of gestural commands. For this task, we have employed state-of-the-art computer vision approaches for feature extraction, encoding, and classification. Our gesture and action classification pipeline, see Fig. 5, employs Dense Trajectories [55] along with the popular Bag-of-Visual-Words (BoVW) framework. The main concept consists of sampling feature points n from each video frame on a regular grid and tracking them through time based on optical flow. Specifically, the employed descriptors are: the Trajectory descriptor, HOG [56], HOF [57] and Motion Boundary Histograms (MBH) [56]. As depicted in Fig. 6, non-linear transformation of depth using logarithm (log-depth) enhances edges related to hand movements and leads to richer dense trajectories on the regions of interest, close to the result obtained using the RGB stream.

The features were encoded using BoVW and were assigned to $K = 4000$ clusters forming the representation of each video. Afterwards, each trajectory was assigned to the closest visual word and a histogram of visual word occurrences was computed. For classification non-linear SVMs were adopted using the χ^2 kernel [56], and different descriptors were combined in a multichannel approach accomplishing an accuracy of 81% and 84% for the tasks washing the legs and back, respectively. Since multiclass classification problems were considered, an one-against-all approach was followed and the class with the highest score was selected accomplishing classification results of up to 83% and 85% for the two tasks.

3.2. Audio processing

In addition to gesture command recognition, we have also developed a spoken command recognition module [58] (Fig. 7) that detects and recognizes commands provided by the user freely, at any time, among other speech and non-speech events, possibly infected by environmental noise and reverberation. The employed features for acoustic modeling were 39 Mel-Frequency Cepstral Coefficients (MFCCs) with their first- and second-order derivatives extracted every 30ms from overlapping windows of 25ms duration. We target robustness via a) denoising of the far-field signals by delay-and-sum beamforming, b) global Maximum Likelihood Linear Regression (MLLR) adaptation of the acoustic models to the speaker-microphone channels of the targeted environment, and c) combined command detection/recognition. In order to build our models we performed offline classification experiments of pre-segmented commands, based on a task-dependent grammar of 23 German spoken commands, accompishing accuracies of 76% and 68% for the two tasks. For a more detailed analysis and further results regarding the offline experiments of the audio-gestural HRI module we refer the reader to [59].

4. Real-time Robotic motion Adaptation

The execution of the robotic tasks, instructed by the user or the carer via HRI commands, relies on a detailed and adaptive robotic motion planning method. This method was developed during the project and is based on enriched visual information. In particular, Point-Cloud data provided by the Kinect sensors, which are mounted on the shower room as shown in Fig. 1 (b), are used as an input together with accurate body part recognition performed either as a pixel-wise [31] area coverage of each body part, or by using more structural [60] skeleton information as depicted in Fig. 8. The latter deep learning methods highly enhance the environment perception skills of the system and are able to provide robust input to the motion planning, remaining at the same time unaffected by the changes of the environment conditions (e.g. illumination) of different shower rooms.

The motion planning method initially proposed in [61] provides a solution to the problem of defining the motion behavior

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020. DOI. https://doi.org/10.1016/j.robot.2020.103451.

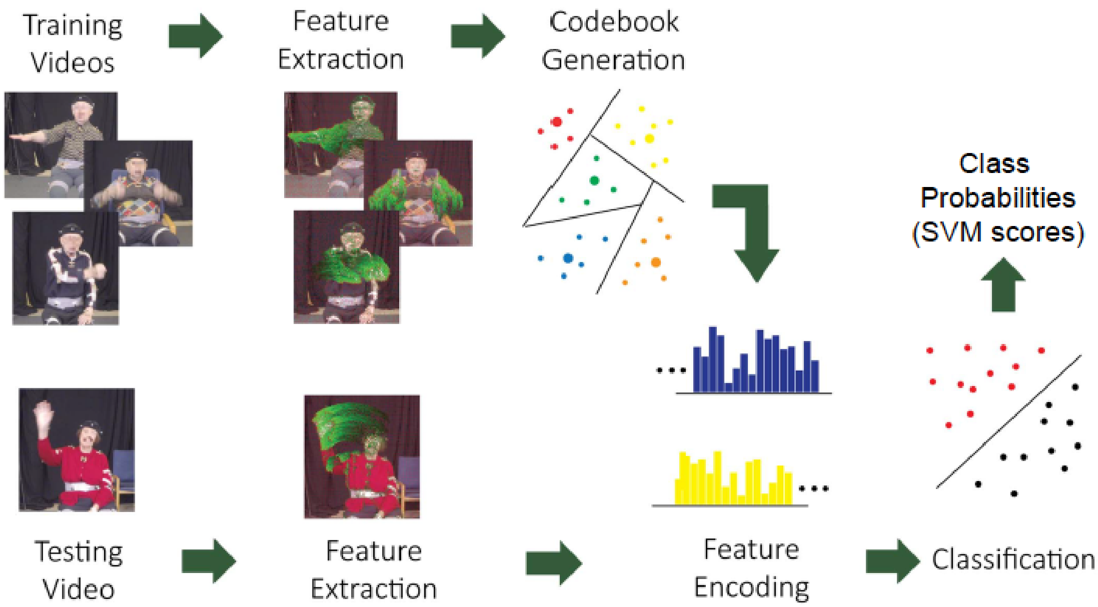


Figure 5: Visual gesture classification pipeline.

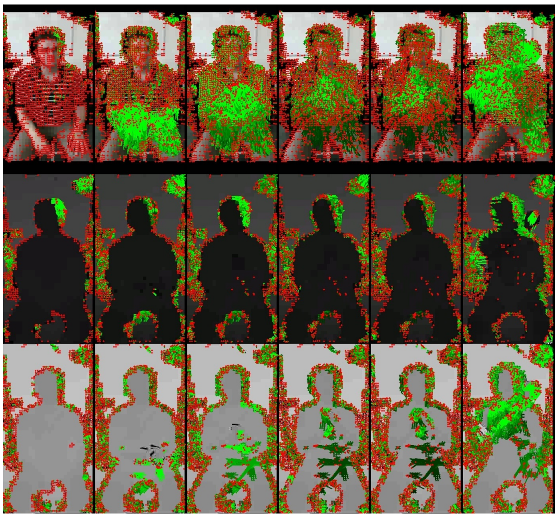


Figure 6: Comparison of dense trajectories extraction over the RGB (top), depth (middle) and log-depth (bottom) clips of gesture "Scrub Back".

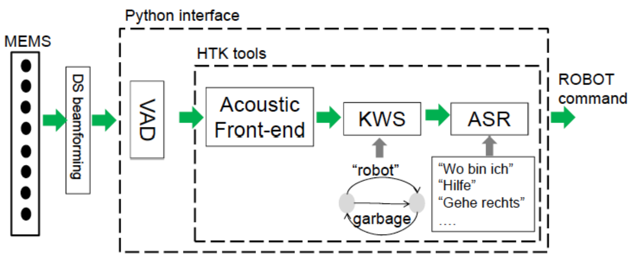


Figure 7: Spoken command recognition module for HRI, integrated in ROS, including always-listening mode and real time performance.

of a robotic manipulators end-effector, operating over a curved deformable surface (e.g. the users body part). Such surfaces

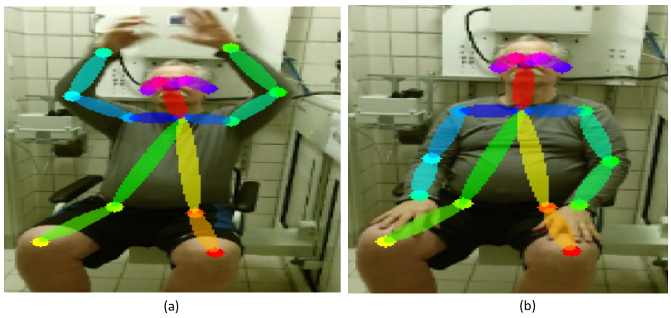


Figure 8: Example of the visual user perception algorithm in two instances: (a) hands-up and (b) relaxed. The visual information is used as an input for the motion planning method.

characterize the human body parts, which are systematically or randomly moving and deforming (e.g. due to users breathing motion). The main goal of the motion behavior task is the online calculation of the reference pose for the end-effector, in order for the robotic manipulator to be compliant with the body part and simultaneously execute predefined surface tasks (e.g. scrubbing the users back). Due to the motion control complexity of the soft robotic manipulator described in Sec. 2.4, the motion planning method is structured to be model free. Therefore, it can be adjusted to any robotic manipulator, provided that all the robot's workspace and velocity constraints are taken into account.

The two scenarios, i.e. washing or scrubbing the back or the legs, considered in the I-Support project are actually the most challenging for the elderly in terms of mobility limitations. These scenarios were considered during the development of the motion planning method as shown in Fig. 9. In the back washing scenario the area pixel-wise visual information is used, whereas in the legs washing scenario we use the skeleton human

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas,
P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour,
A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López,
"I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population",
Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

recognition. These recognition techniques are combined with the Point-Cloud information obtained from the Kinect sensors in order to achieve accurate reference pose estimation.

(a) (b)

Figure 9: Two scenarios were considered during the development of the motion planning method: (a) Back washing and (b) Legs washing. In (a) area pixel-wise visual information was used, whereas in (b) the skeleton information was adequate for the completion of washing tasks for the legs. The red spheres depict the result of the skeleton recognition for the right leg and the green sphere depicts the target position that the robot has to reach as a result of the motion planning method.

Another important aspect of the motion planning method is the straightforward combination with other motion planning techniques, which allow for the incorporation of imitation learning techniques. More specifically in [62] we integrated a leader-follower framework of motion primitives (CC-DMP) [63] with a vision-based motion planning method to adapt the reference path of a robots end-effector and allow the execution of washing actions. This system incorporates clinical carers expertise by producing motions, which are learned by demonstration, using data from the publicly available KIT whole-body motion database [64]. This integration accomplishes to make the robotic washing actions more human-friendly and more acceptable by the elderly users.

5. Audio-gestural data collection

In order to be able to build accurate models for learning in our audio-gestural communication module, we conducted a systematic collection of multisensory data, where various experiments were designed with the involvement and guidance of the clinical partners. During the whole process of designing the data collection experiments, the clinical partners continuously contributed with valuable feedback in order to take into account the specific capabilities and needs of the elderly end-users. The experiments contained multiple scenarios, e.g. for entering and exiting the shower, washing/scrubbing the back or the legs, for stopping or repeating a procedure, for changing the temperature of the water etc. Figure 10 shows some samples of the designed gestures for various bathing tasks.

The dataset includes data recordings in a more strict and guided context (where the commands are predefined) and recordings of freestyle audio-gestural commands while introducing various dialogue features for the interaction. Specifically, three different sessions for data collection were defined, namely recordings of: a) 33 audio-gestural commands, performed simultaneously, b) spoken commands only and c) gesture commands only. For the spoken commands we provided a short and a long

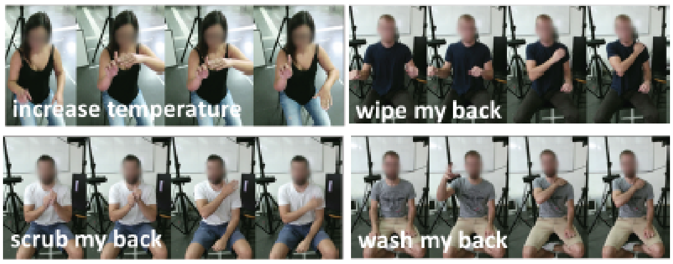


Figure 10: Sample gestures for the bathing HRI task.

version that is preceded by a system activation keyword, the female name “Roberta” that exists in both German and Italian language, where the actual validations with the elderly users were conducted. For the gesture commands we provided more than one gestures in order to include variability and thus, consider factors such as a) physiological aspects of the elderly, b) intuitiveness and naturalness, c) the cognitive capacity of the elderly as well as d) the design of a system that could recognize smaller or larger variations of the same command. Those commands, in a second phase, and for the validations with the elderly end-users were narrowed down, taking into account the results of the first recognition experiments (recognizability and discriminability by the machine algorithms employed) and after consulting the clinical partners regarding what is most suitable for the elderly end-users.

Data collection experiments were performed using the Kinect sensors integrated in a ROS environment and synchronized using software triggering, while an integrated annotation and acquisition web-interface that facilitates on-the-fly temporal ground-truth annotation for fast acquisition [65] was used. For carrying out the recordings supplementary material has been provided with the predefined spoken commands and the videos of the predefined gestures. For the freestyle experiments pictures have been assembled in order to guide the user so as to perform the various tasks using natural language and gestures.

5.1. Audio-Gestural Development Dataset

We have recorded visual data from 23 users (eight females and fifteen males, aged 23-35) while performing predefined gestures, and audio data from 8 users while uttering predefined spoken commands in German (the users were non-native German speakers, having only some beginner’s course). The total number of commands for each task was: 25 and 27 gesture commands for washing the legs and the back, respectively, and 23 spoken commands for the core bathing tasks, i.e. washing/scrubbing/wiping the back or legs, for changing base settings of the system, i.e. temperature, water flow and spontaneous/emergency commands. A background model was also recorded, including generic motions or gestures that are actually performed by humans during bathing; so as to be able to reject out-of-vocabulary gestures/motions as background actions.

5.2. Extended Audio-Gestural Dataset

The data have been collected from 12 native German speakers (three females and nine males, aged 18-30), having three

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

iterations/repetitions each. In more detail, the extended dataset includes:

- Calibration recording (checker-board pattern, no human subject).
- Human subject is given textual/visual description of the action to be performed (e.g. "instruct the system to dry the legs"); for both washing positions (back/legs).
- Human subject is shown a video of predefined gestures; for both washing positions.
- Background model including random "negative" gestures.
- Wash/wiping motions in wash-back/legs position, including four different wiping styles each (horizontal/vertical, small/big circles).
- Speech commands read in wash-back/legs position.

Complementary high-precision VICON MX motion capture data using 10 cameras have been recorded for a subset of the subjects, due to the efforts involved for calibrating the VICON system given the occlusions caused by the I-Support setup. Specifically, seven human subjects have been recorded in two recording sessions each, during the above-described recording protocol, yielding a total of 1.7 TB of data in 1636 Rosbags that are available from the *KIT Motion Database* for the purposes of the project (with VICON data being available for three of the subjects).

6. Experimental Evaluation

Two multinational validation studies were conducted in two European pilot sites: a) *Fondazione Santa Lucia* (FSL) Hospital in Rome, Italy and b) *Bethanien* Hospital in Heidelberg, Germany; including target population, outcomes and indicators to evaluate the I-Support system for addressing all evaluation activities. The intention of validating the system multinationally is to increase the value of the users' subjective assessments and especially to evaluate the acceptability of the system by people with different habits and social-ethical background. Both studies included the installation and the evaluation of the various functionalities of the I-Support system in the bathroom environment of the selected care facilities. For the conduction of the validations both pilot sites obtained approval from the Ethics Committee.

The studies were realized in two rounds at each pilot site:

1. The first pilot testing consisted of evaluation, in dry conditions, of well-defined functionalities of the first prototype of the bathing robot system and was intended to obtain early feedback. This evaluation round focused on a stable but limited intelligence prototype that included human-robot interaction (HRI) using audio-gestural commands. Feedback received from this testing round was used into the design and the development process.
2. The second pilot testing round consisted of mainly summative evaluation activities and focused on the fully functional and intelligent system, thus the showering activities with water (i.e. rinsing, scrubbing), including the shared control functionality of the I-Support system, the integrated learning and cognition strategies, as well as the full context awareness. Additionally, a water pouring scenario was also evaluated examining various operation modes and also the end-users' general satisfaction and preferences for various controllers for the robotic motion.

The user group of the I-Support bathing robot is defined as persons with difficulties in bathing activities, as evaluated by the bathing item of the Barthel Index (0 pt. = "patient can use a bath tub, a shower, or take a complete sponge bath only with assistance or supervision from another person") [66], which represents a clinically well-established index to evaluate an individual's functional disabilities in ADLs. According to this user group definition, participants of the evaluation studies only included persons with difficulty in their ability to perform bathing activities on their own.

The two use cases evaluated included the interaction of the I-Support system with two regions of the body:

- *Distal region* that comprises lower thighs from knee joints downwards and feet. Note that for washing these body parts, the user has to bend forward with a high risk of losing postural control.
- *Back region* expanding from the cervical spine to the tailbone. In this case, it is practically impossible for the users to reach their back.

During the validations, the Kinect sensors were installed in the bathrooms of the two hospital sites as shown in Fig. 1(b); incorporating the required adjustments regarding their positions and angles depending on the available space of each room in order to be able to monitor the elderly and the robotic soft arms. Based on this data the perception unit reconstructed a 3D model of the elderly and the robot and provided feedback to the system controller. Then the system controller generated motion control commands and tool commands and guided autonomously the soft arms to wash, scrub and rinse the specific body parts.

6.1. Multimodal Fusion and Online A-G Command Recognition System Integration

For the online validations and in order to evaluate the System Performance of the audio-gestural communication of the user with the robot, our online Audio-Gestural (A-G) multimodal action recognition system was used, developed in [65] using the Robotic Operating System (ROS), see Fig. 11. The online A-G system enables the interaction between the user and the robotic soft arms and thus, monitors, analyzes and predicts the user's actions, giving emphasis to the command-level speech and gesture recognition. Always-listening recognition is applied separately for spoken commands and gestures, combining at a second level their results. A late fusion scheme is used, where the individual multi-class results are combined, encoding

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

this way inter-modality agreement, using a simple rule that was found quite effective and fast among other approaches [58].

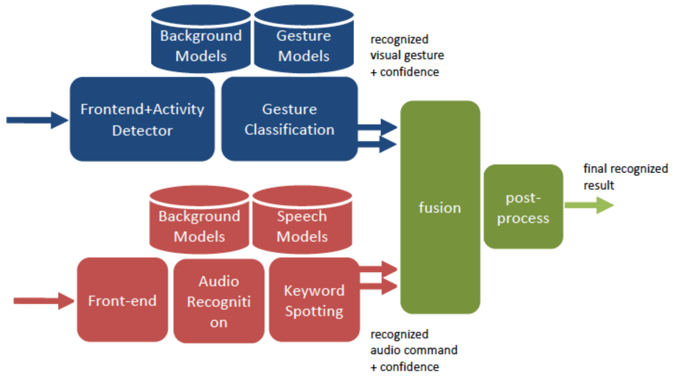


Figure 11: Online Audio-Gestural Command Recognition System.

The overall system comprises two separate sub-systems:

1. The *spoken command recognizer*, a single node that performs always-listening speech recognition of predefined command phrases that accompany the gestures.
2. The *gesture recognizer* consisting of two nodes: the *activity detector* that performs temporal localization of segments with visual activity and the *gesture classifier*.

The spoken command recognition module works as described in Sec. 3.2. The activity detector processes the RGB or the depth stream in a frame-by-frame basis and determines the existence of visual activity in the scene, using a thresholded activity "score" value. The output from both sub-systems are combined at the fusion node producing a single result. Specifically, the implemented fusion node receives the N-best results announced by the individual recognizers, checking for available results every 0.5 secs by receiving periodically messages in order to synchronize the recognizers. A waiting period T is also defined, during which the node waits to combine the incoming messages. If this period expires, it is assumed that one modality either may not have been activated by the user, or failed to detect the given input. In such cases, the node announces single-modality results.

The A-G recognition system's grammars (for both spoken and gesture command recognition) and functionalities were adapted to the specific bathing tasks, delivering recognition results as ROS messages to the system's finite state machine (FSM) that: (1) decided the action to be taken after each recognized command, (2) controlled the various modules and (3) managed the dialogue flow by producing the right audio feedback to the user, for more details see Sec. 2.5.

6.2. Experimental Setup and Protocol

The proposed experimental setups aimed to:

1. Objectively evaluate the effectiveness of the bathing systems individual modules. Therefore, the experimental protocols were designed to provide as much data as possible for statistical analysis, given the frailty of the users and their limitations.
2. Assess the acceptability of the overall system by the elderly subjects. To accomplish this, the experiments were conducted under realistic conditions and included typical procedures that the users might follow during their bathing routine with the I-Support system.
3. Assess the ability of the users to complete a bathing procedure, according to their preferences, using the HRI components of the system.
4. Assess the acceptability of the system by the caregiving personnel and its ability to monitor the effectiveness and safety of the bathing procedure.

6.2.1. Validation Round I

During the experiments, we simulated the two bathing scenarios at dry conditions, at the two pilot sites: 1) FSL Hospital and 2) Bethanien Hospital. Validation round I followed the same evaluation study designs in both pilot sites, in order to obtain comparable evaluation results. For the HRI experiments on each site, potential I-Support users were recruited based on the following main inclusion criteria: (1) dependency in bathing activities as assessed by the bathing item of the Barthel Index [66] and (2) no severe cognitive impairment as assessed by a score of >17 on the Mini-Mental State Examination (MMSE) [67]. Recruitment yielded a significant number of participants at both sites: 25 (mean age±SD: 67.4±8.9 years) and 29 (mean age±SD: 81.4±7.7 years), respectively. Table 1 shows an overview of the setups for the two validation rounds at the two pilot sites; where naive denotes elderly users having no experience with the system, but only a few minutes training prior to their interaction.

The **experimental protocol** exhibited a variety of 7 audio (A) or audio-gestural (A-G) commands, in Italian and in German (see Table 2), in sequences that would simulate realistic interaction flows for both tasks. Table 3 shows the sequence of the A and A-G commands as performed in both validation experiments. Prior to the actual testing phase, all commands were introduced to the participants by the clinical test administrator, while during the experiment the participants were guided on how to interact with the robot by showing the audio commands written on posters and the audio-gestural commands by performing them, instructing them to simply read or mimic them. The administrator could also intervene whenever the flow was changed unexpectedly after a system failure. Additionally, a technical supervisor handled the PCs and annotated on-the-fly the recognition results of the system.

6.2.2. Validation Round II

A second set of validation experiments were carried out at the same pilot sites, aiming at an enhanced user experience. In Bethanien hospital the experiments were also designed to accompany a clinical sub-study. For the purpose of these validation studies, some of the gestural commands were redesigned based on feedback from previous tests in order to become easier and more intuitive for the elderly users. To address the lack of data for the newly introduced gestures, a small scale data

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

		Validation Round I	Validation Round II
Heidelberg	# users	29 naive German-speaking patients	25 naive German-speaking patients
	mean age±SD	81.4±7.7 years	77.9±7.9 years
	MMSE (max. 30pt.)	25.3 ± 3.1	25.6 ± 3.1
	evaluated commands	7 A-G commands	7 A-G commands
	scenario	"Legs" and "Back" position	"Back" position Gesture-only and audio-gestural scenario
FSL	# users	25 naive Italian-speaking patients	25 naive Italian-speaking patients
	mean age±SD	67.4±8.9 years	69.4±7.5 years
	MMSE (max. 30pt.)	27.6 ± 2.5	24.3 ± 3.5
	evaluated commands	7 A-G commands	7 A-G commands
	scenario	"Legs" and "Back" position	"Legs" position

Table 1: Overview of the validation setups for the two rounds at the two pilot sites.

Vocabulary of A-G commands		
English	**Italian**	**German**
Wash legs	Lava le gambe	Wasch meine Beine
Wash back	Lava la schiena	Wasch meinen Rücken
Scrub back	Strofina la schiena	Trockne meinen Rücken
Stop (pause)	Basta	Stop
Repeat (continue)	Ripeti	Noch einmal
Halt	Fermati subito	Wir sind fertig

Table 2: Validation Round I: The audio-gestural commands that were included in the two bathing scenarios. All commands were preceded by the keyword "Roberta".

	Distal Region		Back Region	
ID	**Command**	**Modality**	**Command**	**Modality**
1	Wash Legs	A	Wash Back	A-G
2	Stop	A	Halt	A-G
3	Repeat	A	Scrub Back	A-G
4	Halt	A	Stop	A-G
5	Wash Legs	A-G	Repeat	A-G
6	Halt	A-G	Halt	A-G
7	Halt	A-G	Halt	A-G

Table 3: Validation Round I: The sequence of Audio (A) and Audio-Gestural (A-G) commands performed by the participants in the validation experiments.

collection with healthy subjects was carried out prior to the beginning of the studies.

Validation round II at FSL: 25 naive Italian-speaking patients tested the system (mean age±SD: 69.4±7.5 years). The **experimental protocol** in this case too included a set of 7 audio-gestural commands, for which Italian audio models were developed. The participants were seated in the "legs" position and the robot responded to each command with (a) audio feedback and (b) by simulating the appropriate action with the soft arm for the the commands: "Wash Legs", "Scrub Legs", "Stop", "Repeat" and "Halt". The exact scenario is shown in Table 4. The administrator, for each participant, filled in a report sheet according to the user's performance and the system's command recognition.

Validation round II at Bethanien: A set of 25 elderly patients were selected with mean age±SD: 77.9±7.9 years. The **experimental protocol** included 7 commands, shown in Table 5, to which each patient was briefly introduced. They were first asked to perform only the gestural part of the 7 commands (gesture-only (G) experiment) and then both the audio and the gestural part of the commands at the same time (audio-gestural (A-G) experiment).

During all experiments, the participant was seated in the "back" position. After a command was performed, a short break took place to give the I-Support system the opportunity to respond to the command. In case of successful recognition, the system responded with an appropriate audio response and soft-arm action. The maximum system response time was about 3 to 5 seconds for the G and the A-G experiments, respectively. If the system did not recognize the command correctly in this time interval or the command was performed incorrectly by the participant, the test administrator asked the participant to repeat the command once more (i.e. maximum 14 attempts for each experiment [2 attempts $\times$ 7 commands]).

6.3. Audio-Gestural Validation Experiments and Results

6.3.1. Validation Results Round I

Multimodal recognition was evaluated in terms of (1) Multimodal Command Recognition Rate (CRR): CRR = # of commands correctly recognized by the system / # of commands correctly performed by the user, (2) accuracy, and (3) user performance/learning rate, so as to measure the correlation of the system's performance with the user's experience. The above metrics were considered in order to correlate, as much as possible, the systems performance with the user's experience. User performance was evaluated by the percentage of well-performed commands relative to the total number of commands tested. A gestural command was rated as not well-performed when the movement of the gesture was not performed as intended, including movement errors that seriously affected the trajectory of the gesture. An audio-gestural command was rated as not well-performed if the gesture was not performed as intended (as described before) and/or if phrasing of the audio command was false. In this validation round, demanding acoustic and visual conditions were faced along with large variability in the performance of the A-G commands by the participants, constituting the recognition task rather challenging. Table 6 shows the obtained CRR (%) and accuracy results (%), which are up to 84% and 80% for both washing tasks, averaged across 29 (FSL) and 25 (Bethanien) users, respectively. The deviation between the two sites is due to the lower age and higher cognitive level of

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020. DOI. https://doi.org/10.1016/j.robot.2020.103451.

A-G command	**Soft-arm motion & feedback**
1. –	the soft-arm goes to a pre-defined initial position
2. Wash my legs = "Roberta, Lava la gambe"	audio feedback is provided and the soft-arm performs a washing movement
2. Increase temperature = "Roberta, Più fredda"	audio feedback is provided, the soft-arm continues the washing movement
3. Lower temperature = "Roberta Più calda"	audio feedback is provided, the soft-arm continues the washing movement
4. Stop = "Roberta, Fermati"	audio feedback is provided, the soft-arm's movement is paused
5. Scrub my legs = "Roberta, Strofina la gambe"	audio feedback is provided, the soft-arm performs a scrubbing movement
6. Stop = "Roberta, Fermati"	audio feedback is provided, the soft-arm's movement is paused
7. Repeat = "Roberta, Ancora"	audio feedback is provided, the soft-arm resumes the scrubbing movement
8. Halt = "Roberta, subito"	audio feedback is provided and the soft-arm goes to its rest position

Table 4: Validation Round II (FSL): Commands tested in the scenario on the audio-gestural human-robot interaction. Left column indicates the audio commands and the right column the audio feedback given by the system and the corresponding simulated action of the soft-arm.

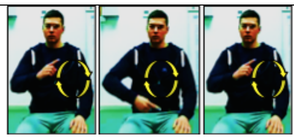
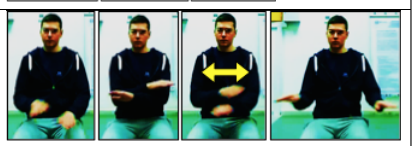
Audio Command	**Gesture Command**
1. Wash back = "Roberta, wasch meinen Rücken"	
2. Higher temperature = "Roberta, wärmer"	
3. Lower temperature = "Roberta, kälter"	
4. Scrub back = "Roberta, schrebbe meinen Rücken"	
5. Repeat = "Roberta, noch einmal"	
6. Stop = "Roberta, stop"	
7. Halt = "Roberta, wir sind fertig"	

Table 5: Validation round II (Bethanien): Commands tested on the audio-gestural human-robot interaction scenario. Left column indicates the audio commands and the right column the corresponding gesture for the audio-gestural commands.

the FSL patients, as well as the slight differences in lightning and cameras' placement.

Gesture recognition was expected more challenging while bathing the legs, due to occlusions of the hands with the robot and/or the chair. The users experienced only a limited amount of false alarms (3 in total as measured at FSL), which were con-

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

	System Performance %						User Performance %			
	CRR %			Accuracy %			Speech		Gestures	
	L	B	Av.	L	B	Av.	L	B	L	B
FSL	80	87	83.5	86	73	79.5	98	99	81	78
Bethanien	85	74	79.5	67	77	72	91	90	84	71

Table 6: Validation Round I Results: Average Audio-Gestural Command Recognition Results; System Performance (CRR%) and User Performance (%) averaged across 29 and 25 users at FSL and Bethanien Hospitals, respectively. (L stands for legs, B for back and Av. for average.

sidered annoying, since the system triggered a response without an "actual" input. Regarding the user performance, the participants performed successfully the spoken commands (over 98% and 90% accuracy for the two sites), while the average performance of gestures was satisfactory (between 70% to over 80% for the two tasks), considering the quick training provided by the administrator. Finally, we have to mention that the results of both modalities were somehow degraded, when the user performed simultaneously the A-G commands, due to increased cognitive load.

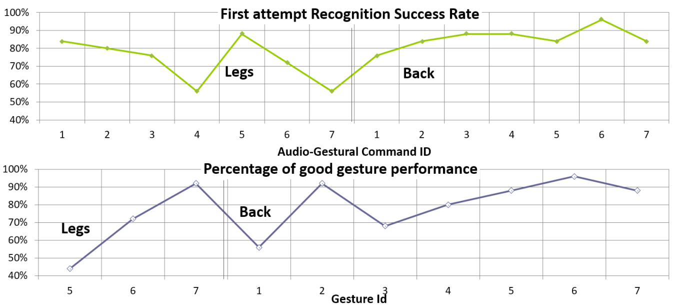


Figure 12: Validation Round I Results (FSL): Recognition statistics per command. The vertical red lines distinguish the command sequences (see the corresponding IDs in Table 3 for the tasks "washing the legs" and "washing the back".

Figure 12 shows indicative curves on how the users performed, on their first attempt, each gesture command after the short training, as measured for the two task at the FSL validation. We note that initially (gesture ID 5) the users either were not familiar with this type of communication or their concentration level was low, since they were performing only spoken commands up to that point. There was however a tendency of increased learning rate, meaning that during the experiments the users got more familiar with the multimodal commands and executed them more accurately, indicating the intuitiveness of this HRI modality. Especially for commands such as "Halt", which was repeated several times (ID 4,6,7) during the washing sequence the command performance of the user reached levels higher than 90%. This observation is highly important, since we can conclude that simple combinations of spoken and gestural commands are both memorable and suitable for an elderly user's communication with an assistive robotic system.

6.3.2. Validation Results Round II

During ***Validation round II at FSL*** we put the emphasis on the overall system integration, such as the control flow of the bathing procedure, where the system performance in CRR% was up to 83.5% for the legs position. The recognition performance of the individual audio and gesture modalities was 82.9% and 48.7% respectively, indicating, as in previous experiments, that the synergy between complementary modalities can actually enhance overall results.

Figure 13 reports results on User Performance and System Performance (% per command), i.e. the rates of well-performed commands and commands recognized by the system. As a second attempt was made when a command was not well performed and/or it was not recognized, for both User and System Performance the reported percentages refer to the rates of well-performed or recognized commands at the first attempts and when summed the first and the possible second attempts. Regarding user performance, since in this validation the commands were audio-gestural, by well-performed we consider that both the audio and the gestural command were performed as intended. Regarding system performance, percentages show the system's recognition regardless the users performance. Indeed, we noticed that although a command was not well-performed, the system was able to recognize it.

The results for user and system performance are quite different, thus deserving a careful discussion. As for user performance, the results show a wide range of rates of well-performed commands with the minimum value of 40% ("Wash Legs" command, 1st attempt) and the maximum value of 88.5% (second "Stop" command, 1st + 2nd attempt). Most of the times, the user had more difficulty in performing the gestural command than the audio one, thus most of times the second attempt was required due to insufficient performance. Interestingly, the "Stop" command, that is performed two times throughout the whole showering session, presents different rates between the first (on average 51%) and the second (on average 88.25%) attempt. This might suggest the unease of performing gestural commands and the need of repetition in order to better learn it. Regarding the system performance, the recognition rates shows a narrower range, with the minimum value of 72% (for the commands "Wash Legs", "Temperature Up", and "Repeat", 1st attempt) and a maximum value of 100% (for the commands "Scrub Legs" and first/second "Stop", 1st + 2nd attempt).

Overall System Usability: During Validation Round II, the participants completed the System Usability Scale (SUS) to evaluate their subjective perception of the overall usability of the I-Support bathing robot. The SUS is a well-established, reliable and valid 10–item scale, which can be quickly and easily administered to determine the user-perceived usability (effectiveness, efficacy, and satisfaction) of technical systems [68]. Its items are scored on a 5–point Likert-type scale ranging from "strongly agree" to "strongly disagree". The combined scores of the individual SUS items are converted into a total SUS score ranging from 0 to 100, with a higher score indicating better usability. SUS scores can be classified as "worst imaginable" (0–25 points), "poor" (25–39 points), "acceptable" (39–52 points), "good" (52–73 points), "excellent" (73–85 points), and "best imaginable" (85–100 points) perceived usability [69].

The SUS score across participants (n = 25) averaged 63.8±12.1 points, indicating an overall "good" usability of the I-Support system tested during the validation experiments. The SUS scores

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

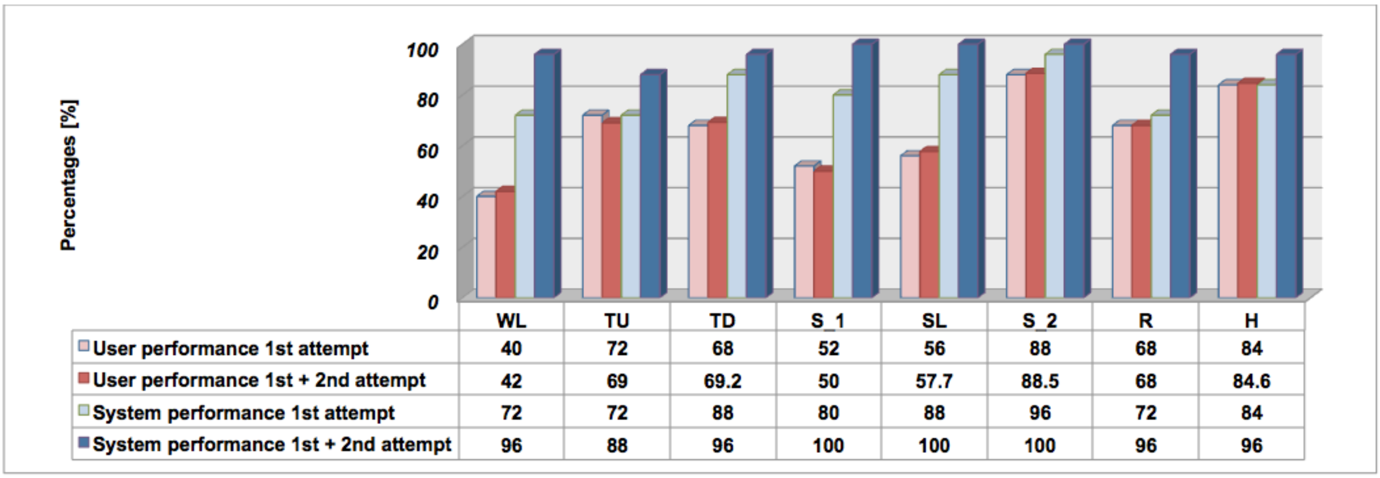


Figure 13: Validation Round II Results (FSL): User Performance (%) and System performance (%) rates per command, showing the percentage of the well-performed commands, i.e. commands performed as intended (User Performance) and the percentage of system recognitions (System Performance) for each of the following showering task commands: Wash Legs (WL), Temperature Up (TU), Temperature Down (TD), 1st Stop (S_1), Scrub Legs (SL), 2nd Stop (S_2), Repeat (R), Halt (H). Note: Audio-Gestural commands were considered as well-performed when both the audio and the gestural command were performed as intended.

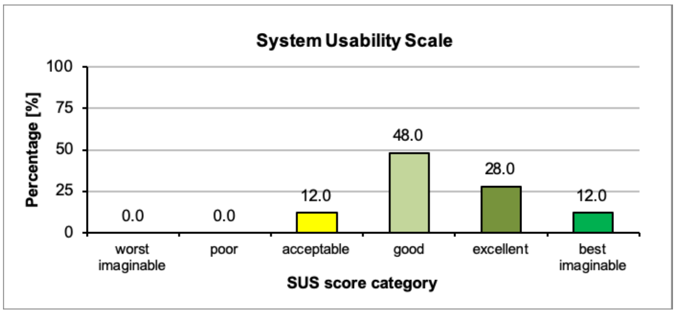


Figure 14: Percentage (%) distribution of participants' ratings in the different SUS score categories.

ranged from the minimum of 42.5 points to the maximum of 87.5 points, suggesting more or less a uniformity in the SUS results. However, when averaging the scores into the different rating categories, as can be seen in Fig. 14, most of the participants gave a rather positive feedback on the I-Support usability. Specifically, 12% of the participants rated the I-Support system as "acceptable", 48% as "good", 28% as excellent, and 12% as "best imaginable".

Validation results round II at Bethanien:

The user performance and system performance were rated and captured by standardized observation of the clinical test administrator using a report sheet for each modality. The user performance of the G and the A-G commands was rated using the same assessment strategy for both modalities. Table 7 shows a detailed analysis for the system performance, where we observe significantly higher CCR (%) results for the A-G experiment compared to the G-only experiment. The two esoteric columns show monomodal results for the audio and gesture modality separately obtained during the multimodal (A-G) scenario. In the last column, which shows the fusion (A-G) results, we observe that the use of both modalities can actually enhance the results. In our understanding and from the observations we made during the validation, we assume that the users probably paid much more attention to gestural than to audio commands.

	Gesture-only	**Audio-Gestural**		
		audio	**gesture**	**fusion**
Without training	59.6%	79.5%	48.0%	86.2%

Table 7: Validation Round II Results (Bethanien): Average System Performance results (CRR%) for the G-only and the A-G experiments; and monomodal results obtained during the multimodal A-G experiment.

Finally, Fig. 15 shows results for the system and the user performance per command, where we note that the per-gesture performance for the gesture-only experiment accomplished a CRR of up to 68.7%. Regarding the individual per command results, the system accomplished the maximum value of 83% for the "Wash" command, while the performance of the users yielded the maximum value of 66% for "Halt". In this case too, we observe that the system successfully recognized various commands that were not well-performed by the users (i.e. "Wash" 83% vs. 34% and "Repeat" 56% vs. 26% for System and User Performance, respectively), which indicates the significance of building good models for learning as well as designing models that are able to recognize smaller or larger variations of the same command.

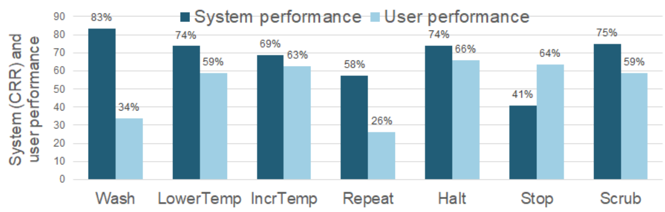


Figure 15: Validation Round II Results (Bethanien): System Performance (CRR%) and User Performance (%) results per command.

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

Water Pouring Scenario at Bethanien

The main goal of the water pouring scenario, conducted at Bethanien, was to study the ability of the elderly to control the showering process using different operation modes for the robotic soft-arm, which provide different amount of assistance during bathing the upper back region.

Specifically, the ***three operation modes*** to be evaluated were:

- *Autonomous operation*: The soft-arm automatically executed the motions needed to provide water pouring for the full coverage of the upper back region. In this mode, the participant had *no control* of the robot motion.
- *Shared control*: The participant could use an input device to issue simple motion commands (i.e. left vs. right and up vs. down), which were translated to a number of (high-level) discrete commands for the I-Support control system (i.e. soft-arm moved left/right or up/down), while the system provided assistance in terms of audio signals indicating that: (1) the participant's command was recognized and (2) the command was successfully executed, meaning that the participant could issue the next motion command. Further assistance was provided in terms of ensuring that the upper back region was not exceeded (i.e. robot motions were restrained to remain within the standardized target body area, shown in Fig. 16). In this mode, the participant had predominant, but *not full control* over the robot motion.
- *Tele-manipulation*: The participant could issue motion commands (i.e. up vs. down and left vs. right) using the input device, similar to the shared control mode. In this mode, however, the system did not provide the audio signals for operating assistance, nor did it constrain the robot motion to the upper back region. Consequently, the participant had full control of the robot motion.

Training and Comparison of Input Devices: The main question during the first stage of this study was the user satisfaction and acceptability of the input device for the elderly users of the I-Support system. A motion tracking input method involving technologies, which are available in any smartwatch as presented in Sec. 2.3 versus a more typical button input device were examined. For the first input method a motion tracking hand-wearable device was constructed that was strapped on the external side of the palm (Fig. 17), containing an Inertial Measurement Unit (IMU) (including 3-axis accelerometers, 3-axis gyroscopes and 1 magnetometer), a micro-controller, and a bluetooth transmitter for wireless operation. The device could track the motion of the user's hand, which was transmitted to a central controller and those motions were then translated by the central controller into the desired motion command for the control of the soft-arm. A thimble with an embedded pressure sensor acted as an activation switch for the tracker device (Fig. 17(b)) and only when the user pressed the thimble, the tracker was activated and the motion of the hand was recorded and translated into the desired motion command. The second input device was a commercial waterproof computer keyboard, where the user could issue a motion command for the soft-arm to be controlled by pressing the appropriate arrow button (e.g. upward movement = up-arrow button).

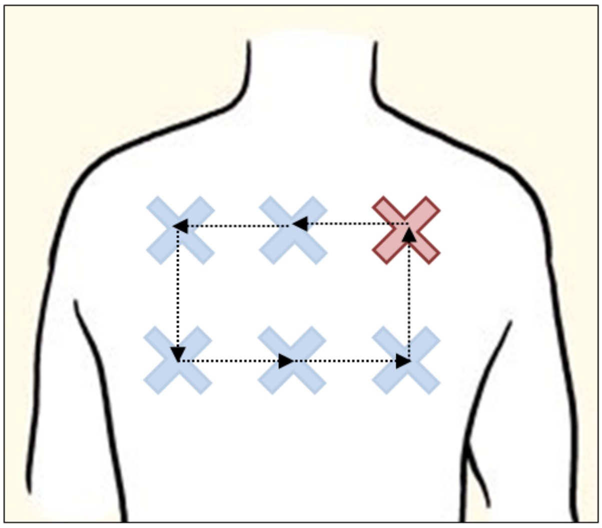


Figure 16: Standardized target body area (upper back region) with the six targets points for which the soft-arm provided water pouring. The red-colored cross represents the starting and final position, the black arrows indicate the path the water stream took on the participant's back region.

Figure 17: a) Motion tracking hand-wearable device. b) Thimble with the embedded pressure sensor for activation).

Both input methods were introduced to the participants as an option for controlling the soft-arm motion. Participants were initially trained in both options by a short computer game, which required to catch a red cube by a user-controlled green cube. If the red cube was caught then it would randomly jump to another field on the "game board" and the participant was instructed to catch it again as many times and as fast as he/she could for a time period of 1 min. The participants were asked, which option of the controller was found easier and thus would like to use in order to control the robot motion of the soft-arm during the water pouring scenario.

After the training on the two controllers, and independently of the participants cognitive status, all (100%) mentioned that (1) providing motion commands was much easier with the computer keyboard than with the motion tracking hand-wearable device and (2) they prefer to use the computer keyboard for controlling the soft-arm in the water pouring scenario. Therefore, the following water pouring scenario was performed in all participants with the use of the waterproof computer keyboard.

Water Pouring Experiments: The second stage of this

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020.
DOI. https://doi.org/10.1016/j.robot.2020.103451.

Operation Mode	Task Effectiveness [%] (mean ± SD)
Autonomous operation	100.0 ± 0.0
Shared control	79.4 ± 18.2
Tele-manipulation	64.4 ± 19.4

Table 8: Task effectiveness in the water pouring scenario with the different operation modes

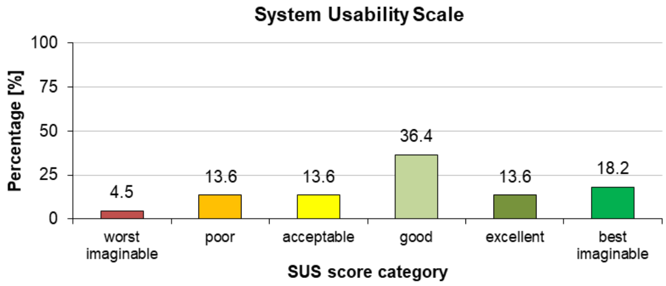


Figure 18: Percentage (%) distribution of participants' ratings in the different SUS score categories

study was conducted inside the showering cabin under real water pouring conditions. Initially, the participant (wearing a swimsuit or swimming trunks) was seated on the motorized chair and the water temperature was set according to his/her preferences. After that, the operation modes were tested in the following order: (1) autonomous operation, (2) shared control, and (3) tele-manipulation.

The test administrator explained that in the first autonomous operation (duration 1 min) the soft-arm would provide water pouring fully automatically for the upper back region following a predefined path, as shown in Fig. 16, with the starting point at the top right.

After the autonomous operation was completed, the participant was introduced into the shared control mode (duration 2 min). In this mode, the participant controls the motion of the soft-arm's water stream on his/her own using the controller chosen after the training game. In this mode the system would also provide an audio signal as previously described. In addition, it was explained that further assistance would be provided in terms of restricting the robot motion to the standardized target body area. Finally, the participant was instructed to cover the entire upper back region (i.e. all six target points) as fast as possible by using the shared control mode.

In the last round of testing, the participant used the tele-manipulation mode (duration 2 min), in which he/she also controlled the motion of the soft-arm's water stream on his/her own, using the same controller as before, trying to cover the entire back region as fast as possible. The participant had now full motion control over the soft-arm and the system did not provide any further assistance.

In the water pouring scenario, task effectiveness with the different operation modes was assessed by a measure of the area coverage, defined as the percentage of the predefined target body area covered with water during the standardized time period (e.g. 3 out of 6 target points covered with water corresponds to 50% coverage).

Maximum task effectiveness was achieved for all participants when the autonomous operation mode was used, indicating a very good and reliable system performance. Task effectiveness substantially decreased with the shared control mode and even more with the tele-manipulation mode, as compared to the fully autonomous mode (see Table 8). Fourteen participants (66.7%) in the shared control mode and 19 participants (90.5%) in the tele-manipulation mode did not achieve the maximum possible coverage.

Our results indicate that the autonomous operation mode for the robotic soft-arm of the bathing robot is highly effective and reliable in providing water pouring for a predefined body area. When giving participants more control over the soft-arm, task effectiveness gradually decreased with lower assistance provided by the bathing robot. These findings suggest that full system autonomy seems to provide a preferred mode of operation for this group of elderly population. A more detailed analysis on the differences in the task effectiveness (and also in the user satisfaction) with the different operation modes fits the scope of another publication focusing further on the findings of the clinical validation study [70].

Overall System Usability: Finally, the participants completed the System Usability Scale (SUS) to evaluate their subjective perception of the overall usability of the I-Support bath robot system.

The SUS score across participants that completed the "water pouring" scenario (n = 22) averaged 60.7±23.0 points, indicating an overall "good" usability of the I-Support system tested during the validation experiments. The SUS scores ranged from the minimum of 0 points to the maximum of 100 points, suggesting a large heterogeneity in the SUS results. However, when averaging the scores into the different rating categories (Fig. 18) most of the participants gave a rather positive feedback on the I-Support usability. More than 81.8% of the participants rated the I-Support system between "acceptable" to "best imaginable", while less than 18.2% rated the I-Support system between "poor" and "worst imaginable".

7. Conclusion

This paper presents the I-Support robotic platform, a human-centered robotic bathing system for smart assisted living, which provides assistance to the frail elderly population in order to safely and independently be able to complete an entire sequence of bathing tasks, such as washing their back and their lower limbs. To achieve this target advanced modules of cognition, sensing, context awareness and actuation have been developed, during the course of the project, and have been seamlessly integrated into the robotic assistance, including a functional soft-arm prototype and the adaptation of a cost-effective robotic chair. Additionally, we contribute a new multimodal audio-gestural dataset and a suite of tools used for data acquisition that have been used for the development and modeling of our high-accuracy, real-time, state-of-the-art multimodal action recognition module for the analysis, monitoring and prediction of user actions.

Please cite it as: A.Zlatintsi, A.C.Dometios, N.Kardaris, I.Rodomagoulakis, P.Koutras, X.Papageorgiou, P.Maragos, C.S.Tzafestas, P.Vartholomeos, K.Hauer, C.Werner, R.Annicchiarico, M.G.Lombardi, F.Adriano, T.Asfour, A.M.Sabatini, C.Laschi, M.Cianchetti, A.Güler, I.Kokkinos, B.Klein, and R.López, "I-SUPPORT: a Robotic Platform of an Assistive Bathing Robot for the Elderly Population", Robotics and Autonomous Systems, Volume 126, April 2020. DOI. https://doi.org/10.1016/j.robot.2020.103451.

We experimentally validated the I-Support system, in two clinical validation studies that were conducted in two European pilot sites showcasing really good system performance, achieving this way an effective and natural interaction and communication, through audio-gestural commands, between users and the assistive robotic platform. Regarding the task effectiveness in the water pouring scenario, the results indicate that the robotic soft-arm is highly effective and reliable and that full system autonomy is preferred by the elderly population. Finally, the two validation studies also proved high acceptability regarding the overall system usability by the elderly end-users.

Acknowledgment

This research work was supported by the EU under the project I-SUPPORT with grant H2020-643666.
The authors would like to thank Dr. Vincenzo Genovese and Mariangela Manti (SSSA), Holger Roßberg and Dr. Inga Schlömer (FRA-AUS), Dr. Vassilis Pitsikalis (NTUA) and all students/ researchers that were involved in the data collection process.

References

[1] W. H. Organization, World report on ageing and health, World Health Organization, 2015.
[2] S. H. Zarit, K. E. Reever, J. Bach-Peterson, Relatives of the impaired elderly: correlates of feelings of burden, The gerontologist 20 (1980) 649–655.
[3] S. McFall, B. H. Miller, Caregiver burden and nursing home admission of frail elderly persons, Journal of Gerontology 47 (1992) 73–79.
[4] D. D. Dunlop, S. L. Hughes, L. M. Manheim, Disability in activities of daily living: patterns of change and a hierarchy of disability, American Journal of Public Health 87 (1997) 378–383.
[5] S. Katz, A. Ford, R. Moskowitz, B. Jackson, M. Jaffe, Studies of illness in the aged: The index of ADL: a standardized measure of biological and psychosocial function, JAMA 185 (1963) 914–919.
[6] M. Porta, A dictionary of epidemiology, Oxford University Press, 2014.
[7] J. C. Millán-Calenti, J. Tubío, S. Pita-Fernández, I. González-Abraldes, T. Lorenzo, T. Fernández-Arruty, A. Maseda, Prevalence of functional disability in activities of daily living (ADL), instrumental activities of daily living (ADL) and associated factors, as predictors of morbidity and mortality, Archives of Gerontology and Geriatrics 50 (2010) 306–310.
[8] S. Ahluwalia, T. Gill, D. Baker, T. Fried, Disaggregating activities of daily living limitations for predicting nursing home admission, Health Services Research 50 (2015) 560–578.
[9] J. Wiener, R. Hanley, R. Clark, J. Van Nostrand, Measuring the activities of daily living: Comparisons across national surveys, Journal of Gerontology 45 (1990) 229–237.
[10] S. Ahluwalia, T. Gill, D. Baker, T. Fried, Perspectives of older persons on bathing and bathing disability: A qualitative study, American Geriatrics Society 58 (2010) 450–456.
[11] C. Balaguer, A. Gimenez, A. Huete, A. Sabatini, M. Topping, G. Bolmsjo, The MATS robot: service climbing robot for personal assistance, IEEE Robotics & Automation Magazine 13 (2006) 51–58.
[12] T. Hirose, S. Fujioka, O. Mizuno, T. Nakamura, Development of hair-washing robot equipped with scrubbing fingers, in: Proc. Int'l Conf. on Robotics and Automation (ICRA-2012), pp. 1970–1975.
[13] Y. Tsumaki, T. Kon, A. Suginuma, K. Imada, A. Sekiguchi, D. Nenchev, H. Nakano, K. Hanada, Development of a skincare robot, in: Proc. Int'l Conf. on Robotics and Automation (ICRA-2008), pp. 2963–2968.
[14] M. Topping, An overview of the development of Handy 1, a rehabilitation robot to assist the severely disabled, Artificial Life and Robotics 4(4) (2000) 188–192.
[15] M. Hillman, K. Hagan, S. Hagan, J. Jepson, R. Orpwood, The weston wheelchair mounted assistive robot - the design story, Robotica 20 (2002) 125–132.
[16] B. J. F. Driessen, H. G. Evers, J. A. v Woerden, Manusa wheelchair-mounted rehabilitation robot, Proc. of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine 215 (2001) 285–290.
[17] ADL-Solutions, Oasis seated shower system, 2016. `http://www.adl-solutions.com/`.
[18] RoboticsCare, Robotics care poseidon, 2016. `http://www.roboticscare.com/robotics-care-poseidon/`.
[19] L. Yang, L. Zhang, H. Dong, A. Alelaiwi, A. El Saddik, Evaluating and improving the depth accuracy of Kinect for Windows v2, IEEE Sensors Journal 15 (2015) 4275–4285.
[20] K. Khoshelham, S. O. Elberink, Accuracy and resolution of Kinect depth data for indoor mapping applications, Sensors 12 (2012) 1437–1454.
[21] S. Pillai, S. Ramalingam, J. J. Leonard, High-performance and tunable stereo reconstruction, in: Proc. IEEE Int'l Conf. on Robotics and Automation (ICRA-2016), pp. 3188–3195.
[22] M. A. Goodrich, A. C. Schultz, Human-Robot Interaction: a survey, Found. Trends Human-Computer Interaction 1 (2007) 203–275.
[23] K. Davis, E. Owusuz, V. Bastaniy, L. Marcenaro, J. Hu, C. Regazzoni, L. Feijs, Activity recognition based on inertial sensors for ambient assisted living, in: Proc. Int'l Conf. on Information Fusion (FUSIO-2016).
[24] R. Kachouie, S. Sedighadeli, R. Khosla, M.-T. Chu, Socially assistive robots in elderly care: A mixed-method systematic literature review, Int'l Jour. Human-Computer Interaction 30 (2014) 369–393.
[25] D. Johnson, R. Cuijpers, J. Juola, E. Torta, M. Simonov, A. Frisiello, M. Bazzani, W. Yan, C. Weber, S. Wermter, N. Meins, J. Oberzaucher, P. Panek, G. Edelmayer, P. Mayer, C. Beck, Socially assistive robots, a comprehensive approach to extending independent living, Int'l Journal of Social Robotics 6 (2014) 195–211.
[26] E. Efthimiou, S.-E. Fotinea, T. Goulas, A.-L. Dimou, M. Koutsombogera, V. Pitsikalis, P. Maragos, C. Tzafestas, The MOBOT platform – showcasing multimodality in human-assistive robot interaction, in: Proc. Int.'l Conf. on Universal Access in Human-Computer Interaction, Vol. 9738 Lecture Notes in Computer Science, Springer Int'l Publishing, 2016, pp. 382–391.
[27] F. Rudzicz, R. Wang, M. Begum, A. Mihailidis, Speech interaction with personal assistive robots supporting aging at home for individuals with alzheimer's disease, ACM Trans. Access. Comput. 7 (2015) 1–22.
[28] A. Kötteritzsch, B. Weyers, Assistive technologies for older adults in urban areas: A literature review, Cognitive Computation 8 (2016) 299–317.
[29] A. Zlatintsi, I. Rodomagoulakis, V. Pitsikalis, P. Koutras, N. Kardaris, X. Papageorgiou, C. Tzafestas, P. Maragos, Social Human-Robot Interaction for the elderly: Two real-life use cases, in: Proc Int'l Conf. on Human-Robot Interaction (HRI–17), March 2017.
[30] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR-2015), pp. 3431–3440.
[31] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, arXiv preprint arXiv:1606.00915 (2016).
[32] L.-C. Chen, Y. Yang, J. Wang, W. Xu, A. L. Yuille, Attention to scale: Scale-aware semantic image segmentation, in: Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR-2016).
[33] F. Xia, P. Wang, L. Chen, A. L. Yuille, Zoom better to see clearer: Human part segmentation with auto zoom net, CoRR abs/1511.06881 (2015).
[34] X. Liang, X. Shen, J. Feng, L. Lin, S. Yan, Semantic object parsing with graph LSTM, in: Proc. European Conf. on Computer Vision, Springer, pp. 125–143, 2016.
[35] S.-E. Wei, V. Ramakrishna, T. Kanade, Y. Sheikh, Convolutional pose machines, in: Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR-2016), pp. 4724–4732.
[36] Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, Realtime multi-person 2D pose estimation using part affinity fields, arXiv preprint, arXiv:1611.08050 (2016).
[37] C. Laschi, B. Mazzolai, M. Cianchetti, Soft robotics: Technologies and systems pushing the boundaries of robot abilities, Science Robotics 1 (2016).
[38] L. Wang, F. Iida, Deformation in soft-matter robotics: A categorization and quantitative characterization, IEEE Robotics & Automation Magazine 22 (2015) 125–139.

[39] G. S. Chirikjian, J. W. Burdick, A hyper-redundant manipulator, IEEE Robotics & Automation Magazine 1 (1994) 22–29.

[40] I. D. Walker, Continuous backbone "continuum" robot manipulators, ISRN Robotics 2013 (2013).

[41] F. Renda, C. Laschi, A general mechanical model for tendon-driven continuum manipulators, in: Proc. Int'l Conf. on Robotics and Automation (ICRA-2012), pp. 3813–3818.

[42] M. D. Grissom, V. Chitrakaran, D. Dienno, M. Csencits, M. Pritts, B. Jones, W. McMahan, D. Dawson, C. Rahn, I. Walker, Design and experimental testing of the octarm soft robot manipulator, in: Defense and Security Symposium, Int'l Society for Optics and Photonics (2006), pp. 62301–62301.

[43] M. Manti, A. Pratesi, E. Falotico, M. Cianchetti, C. Laschi, Soft assistive robot for personal care of elderly people, in: Proc. IEEE Int'l Conf. on Biomedical Robotics and Biomechatronics (BioRob-2016), pp. 833–838.

[44] Y. Ansari, M. Manti, E. Falotico, Y. Mollard, M. Cianchetti, C. Laschi, Towards the development of a soft manipulator as an assistive robot for personal care of elderly people, Int'l Journal of Advanced Robotic Systems 14(2) (2017).

[45] C. Festo, Bionic motion robot, 2017. `https://www.festo.com/group/en/cms/12747.htm`.

[46] B. Bardou, P. Zanne, F. Nageotte, M. De Mathelin, Control of a multiple sections flexible endoscopic system, in: Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems (IROS-2010), pp. 2345–2350.

[47] M. Manti, V. Cacucciolo, M. Cianchetti, Stiffening in soft robotics: A review of the state of the art, IEEE Robotics & Automation Magazine 23 (2016) 93–106.

[48] M. Mahvash, P. E. Dupont, Stiffness control of a continuum manipulator in contact with a soft environment, in: Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems (IROS-2010), pp. 863–870.

[49] D. McNeill, Hand and Mind: What Gestures Reveal about Thought, Univ. of Chicago Press, 1996.

[50] J. Cassell, Computer vision in human-machine interaction, A framework for gesture generation and interpretation, Cambridge Univ. Press, 1998.

[51] T. L. Chen, M. Ciocarlie, S. Cousins, P. M. Grice, K. Hawkins, K. Hsiao, C. C. Kemp, C. King, D. A. Lazewatsky, A. E. Leeper, H. Nguyen, A. Paepcke, C. Pantofaru, W. D. Smart, L. Takayama, Robots for humanity: using assistive robotics to empower people with disabilities, IEEE Robotics & Automation Magazine 20 (2013) 30–39.

[52] D. Vasquez, P. Stein, J. Rios-Martinez, A. Escobedo, A. Spalanzani, C. Laugier, Human Aware Navigation for Assistive Robotics, Springer International Publishing, Heidelberg, pp. 449–462.

[53] Y. Gu, H. Do, Y. Ou, W. Sheng, Human gesture recognition through a Kinect sensor, in: IEEE Int'l Conf. on Robotics and Biomimetics (ROBIO-2012), pp. 1379–1384.

[54] V. Genovese, A. Mannini, A. M. Sabatini, Smartwatch step counter for slow and intermittent ambulation, IEEE Access 5 (2017) 13028–13037.

[55] H. Wang, C. Schmid, Action recognition with improved trajectories, in: Proc. IEEE Int'l Conf. on Computer Vision (ICCV-13).

[56] H. Wang, M. M. Ullah, A. Kläser, I. Laptev, C. Schmid, Evaluation of local spatio-temporal features for action recognition, in: Proc. British Machine Vision Conf. (BMVC-09).

[57] I. Laptev, M. Marszalek, C. Schmid, B. Rozenfeld, Learning realistic human actions from movies, in: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR-08).

[58] I. Rodomagoulakis, N. Kardaris, V. Pitsikalis, E. Mavroudi, A. Katsamanis, A. Tsiami, P. Maragos, Multimodal human action recognition in assistive human-robot interaction, in: Proc. Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP-16).

[59] A. Zlatintsi, I. Rodomagoulakis, P. Koutras, A. C. Dometios, V. Pitsikalis, C. S. Tzafestas, P. Maragos, Multimodal signal processing and learning aspects of Human-Robot Interaction for an assistive bathing robot, in: Proc. Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP-2018), Calgary, Canada.

[60] Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, Realtime multi-person 2D pose estimation using part affinity fields, in: Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR-2017).

[61] A. C. Dometios, X. S. Papageorgiou, A. Arvanitakis, C. S. Tzafestas, P. Maragos, Real-time end-effector motion behavior planning approach using on-line point-cloud data towards a user adaptive assistive bath robot, in: Proc. Int'l Conf. on Intelligent Robots and Systems (IROS-2017).

[62] A. C. Dometios, Y. Zhou, X. S. Papageorgiou, C. S. Tzafestas, T. Asfour, Vision-Based online adaptation of motion primitives to dynamic surfaces: Application to an interactive robotic wiping task, IEEE Robotics and Automation Letters 3 (2018) 1410–1417.

[63] Y. Zhou, M. Do, T. Asfour, Coordinate change dynamic movement primitivesa leader-follower approach, in: Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems (IROS-2016), pp. 5481–5488.

[64] C. Mandery, O. Terlemez, M. Do, N. Vahrenkamp, T. Asfour, Unifying representations and large-scale whole-body motion databases for studying human motion, IEEE Transactions on Robotics 32 (2016) 796–809.

[65] N. Kardaris, I. Rodomagoulakis, V. Pitsikalis, A. Arvanitakis, P. Maragos, A platform for building new human-computer interface systems that support online automatic recognition of audio-gestural commands, in: Proc. ACM on Multimedia Conf. (ACM-16).

[66] F. Mahoney, D. Barthel, Functional evaluation: The barthel index, Maryland State Medical Journal 14 (1965) 61–65.

[67] M. F. Folstein, S. E. Folstein, P. R. McHugh, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician, Journal Psychiatr. Research 12 (1975) 189–198.

[68] J. Brooke, SUS: a "quick and dirty" usability scale. Usability Evaluation in Industry, London, Taylor & Francis, 1986.

[69] A. Bangor, P. Kortum, J. Miller, Determining what individual SUS scores mean: Adding an adjective rating scale, Journal of Usability Studies 4 (2009) 114–123.

[70] C. Werner, A. Dometios, P. Maragos, C. Tzafestas, J. Bauer, K. Hauer, Evaluating the task effectiveness and user satisfaction with different operation modes for an assistive bathing robot in older adults with bathing disability, Under review Assist. Techn. (2019).