Vision-Based Online Adaptation of Motion Primitives to Dynamic Surfaces: Application to an Interactive Robotic Wiping Task

Athanasios C. Dometios[1], You Zhou[2], Xanthi S. Papageorgiou[1], Costas S. Tzafestas[1], and Tamim Asfour[2]

***Abstract*—Elderly or disabled people usually need augmented nursing attention both in home and clinical environments, especially to perform bathing activities. The development of an assistive robotic bath system, which constitutes a central motivation of this work, would increase the independence and safety of this procedure, ameliorating in this way the everyday life for this group of people. In general terms, the main goal of this work is to enable natural, physical human-robot interaction, involving human-friendly and user-adaptive on-line robot motion planning and interaction control. For this purpose, we employ imitation learning using a leader-follower framework called Coordinate Change Dynamic Movement Primitives (CC-DMP), in order to incorporate the expertise of professional carers for bathing sequences. In this paper, we propose a vision-based washing system, combining CC-DMP framework with a perception based controller, to adapt the motion of robot's end-effector on moving and deformable surfaces, such as a human body-part. The controller guarantees globally uniformly asymptotic convergence to the leader movement primitive, while ensuring avoidance of restricted areas, such as sensitive skin body areas. We experimentally tested our approach on a setup including the humanoid robot ARMAR-III and a Kinect v2 camera. The robot executes motions learned from the publicly available KIT whole-body human motion database, achieving good tracking performance in challenging interactive task scenarios.**

***Index Terms*—Motion and Path Planning; Learning and Adaptive Systems; Human-Centered Robotics.**

I. Introduction

MOST advanced countries tend to be aging societies, with the percentage of people with special needs for nursing attention being already significant and due to grow. Health care experts are called to support these people during the performance of Personal Care Activities such as showering, dressing and eating [1], inducing great financial burden both to the families and the insurance systems. During the last years health care technology is developing towards assistive and adaptable robotic systems designed for both in-house and clinical environments, aiming at supporting disabled and elderly people with special needs in terms of Personal Care Activities. There have been very interesting developments in this field, such as the Oasis seated shower system [2] and Poseidon robotics care system [3], with either static physical interaction [4], or mobile solutions [5].

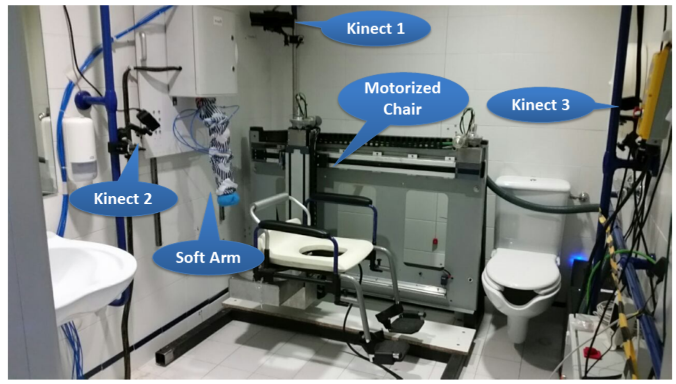

Fig. 1: The I-Support system installed in a clinical environment (Santa Lucia Foundation, Rome, Italy). This system constitutes of a soft robotic manipulator (courtesy of Sant'Anna School of Advanced Studies), a motorized chair (courtesy of Robotnik Automation, Valencia, Spain) and three Microsoft Kinect cameras.

However, body care (showering or bathing) is among the first activities of daily living (ADLs) which incommode an elderly's life [6], since it is a demanding procedure in terms of effort and body flexibility. In this context, an assistive robotic application involving direct human-robot physical contact (such as showering) is way more demanding in terms of safety than other assistive robotic systems. These high safety standards include operation of the robot on curved and deformable human body parts in a dynamic environment, since unexpected body-part motion may occur during the robot's operation. In addition, the showering task should be executed in a human-friendly way in terms both of motion and force exertion on each body part, in order to increase the comfort of an elderly user. Therefore, proper washing motions for each task should be learned by demonstration of health care experts.

Learning a complicated motion by demonstration includes choosing appropriate motion representations and corresponding learning strategies [7]–[9]. Previous works based on the motion primitives for learning robot's interactive motions can be found in [10]–[13]. As a popular method, Dynamic Move-

Manuscript received: September, 10, 2017; Revised December, 12, 2017; Accepted January, 15, 2018.

This paper was recommended for publication by Editor Prof. Nancy Amato upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by the European Union's Horizon 2020 research and innovation programme under the project "I-SUPPORT" with grant agreement No 643666.

[1]A. Dometios, X. Papageorgiou and C. Tzafestas are with School of Electrical and Computer Engineering, National Technical University of Athens, Greece. {athdom,xpapag}@mail.ntua.gr {ktzaf}@cs.ntua.gr

[2]Y. Zhou and T. Asfour are with Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Germany {you.zhou,asfour}@kit.edu

The first two authors contributed equally to this work and their names are alphabetically ordered.

Digital Object Identifier (DOI): see top of this page.

ment Primitive (DMP) [7], [14] with its simple parametrized formulation supports not only imitation learning but also reinforcement learning. Considering its simplicity and potential extension, in this work, we choose DMP to encode washing movement primitives. Based on the DMP formulation, we developed in our previous work [15] a leader-follower framework to take also the user's movement into consideration. Furthermore, in order to realize proper robot's motion and human-friendly interaction according to clinical requirements, we learn washing actions demonstrated by professional carers, which are recorded in KIT whole-body motion database [16].

The spectrum of applications, which involve a robotic manipulator executing surface interactive tasks with the environment using visual feedback, is wide. Examples are from automotive industry, in which industrial manipulators are responsible for transferring or spraying actions on car parts [17] or service robotic interacting with household environment [18]. Famous surgical robotic platforms also follow a master-slave design and consist of manipulators tele-operated by surgeons with a visual aspect of the scene. Another interesting applications of semi-autonomy in cardiac surgery is the beating heart motion compensation [19].

However, close and tight physical interaction with a human being is a much more delicate task and requires advanced perception capabilities. The progress of RGB-D sensors together with sophisticated computer vision algorithms have increased radically the perception abilities of robotic systems. In particular, approaches based on Deep Learning techniques have presented very detailed results on human perception and specifically body-part segmentation. In [20], [21], pixel-level human body parts semantic segmentation is presented, whereas in [22] their sparse pose is calculated with close to real-time computational performance. These recent results, have motivated the development of a vision-based motion controller in [23], which uses a Navigation Function approach to guarantee the proper execution of the task within the body part limits and to achieve the adaptation of simple motions on moving, deformable surfaces.

In this paper, we propose a vision based washing system, which integrates the leader-follower framework of motion primitives (CC-DMP) with a vision-based controller to adapt reference path of a robot's end-effector and allow the execution of washing actions (e.g. pouring water, scrubbing) on moving, curved and deformable surfaces, like human body-parts. This system incorporates clinical carer's expertise by producing motions which are learned by demonstration, using data from the publicly available KIT whole-body motion database. Moreover, the perception based controller uses navigation functions to guarantee globally uniformly asymptotic convergence to the leader movement primitive and obstacle area (e.g. skin injuries) avoidance. In addition, on-line motion adaptation from depth data realizes user motion compensation and local surface estimation, which allows the regulation of the distance and the orientation of the robot's end-effector perpendicularly to the surface. This regulation enables the execution of both contactless actions (such as water pouring) and actions involving physical contact with indirect (open-loop) application of forces (such as wiping or scrubbing).

In our experiments, we use humanoid robot ARMAR-III developed at KIT [24], which applies washing actions on a planar whiteboard and over a male subject's back region. We choose this robot with known kinematic model and mature low-level controller to provide an efficient proof-of-concept demonstration and show-case the developed methodology. In the context of the I-Support project, a soft arm robotic manipulator is under development and will be used for validation experiments in the near future. Nevertheless, by using a humanoid robot in our current experiments we further showcase the diversity of our method and its potential application in the household service robot of the future.

II. Problem Statement

The I-Support robotic shower system, which is currently in development and initial validation stage (Fig. 1), aims to support both elderly and people with mobility disabilities during showering activities, i.e. pouring water, soaping, body part scrubbing, etc. The degree of automation will vary according to the user's preferences and disability level. In Fig. 1, system's basic parts are presented. The robotic system provides elderly showering abilities enhancement, the motorized chair ensures the safe transition of the user in the shower room and three Microsoft Kinect sensors are used for user all-around visual perception and Human-Robot Interaction (HRI) applications.

The motion adaptation problem of a robotic manipulator's end-effector on a moving, curved and deformable surface (e.g. user's body part), in a workspace equipped with a depth-camera, is considered. The core of this motion adaptation task is to calculate at each time step the reference pose for the end-effector of a robotic arm, which will let the robotic manipulator execute proper human friendly surface tasks (e.g. wiping the user's back) and at the same time to be compliant with this body part. We assume that the field of view of the depth camera includes the workspace of the robot and the obstacles in the workspace are visible from the camera perspective. Obstacle areas may regard restricted areas, either on the user's body part subject to washing (e.g. local injury) or on other body parts, which may interfere to the robot's motion (e.g. the hands of the user) and should be avoided during the washing sequence.

The boundaries of each body-part can be found on the image plane with simple color filters or more robust semantic segmentation techniques [22], which are based on Deep Learning. These boundaries will differ from user to user, therefore adaptability to different users is also a very important feature of the system.

Moreover, each human has unlike preferences and needs during the washing sequence. It is crucial for the user to feel comfortable and safe during the operation of the system. Therefore, proper and human friendly washing motions for each subtask should be learned by demonstrations of health care experts. This procedure might raise some requirements for each task, in terms of execution time and motion complexity. However, decomposition into simpler primitive motions (e.g. periodic and discrete) is necessary for a robotic device for technical reasons. The fusion of such primitive actions with different parameters (e.g. duration, amplitude e.t.c.) can reproduce more delicate and human-friendly actions.

In the next section, we will briefly provide the preliminaries of our works, based on which, we develop an integrated perception-based motion planning and interactive control system, which is able to incorporate the recent advances of visual human perception algorithms (in particular on-line segmentation and reconstruction of human body parts) and can simultaneously, in the context of the envisaged application, imitate and execute proper washing actions.

III. Preliminaries

A complete perception-based washing system consists of two main parts, a vision-based controller and an adaptable motion representation. The former one enables the system to perceive and handle the change of the environment, especially in our scenario the moving, curved and deformable washing surface. The latter one introduces the possibility of imitation learning and reinforcement learning by demonstration of health-care experts and generates adaptable washing actions.

A. Perception-Based Motion Planning

The basic goal of the perception based motion planning is to calculate on the fly the leader reference pose, around which each learned washing motion will be applied. This approach commences with the planning of the leader movement primitive's path on a fixed 2D **"Canonical" space**, which is spatially normalized, as depicted with blue color in Fig. 2. This space can be considered as a canvas on which any path can be inscribed, in order for the robot to be able to navigate on any part of the surface that needs to be washed (e.g. the back of the user). This path is followed by using a controller $\mathbf{U} = H\left(-\nabla\varphi\right)$, where H is a function as defined in [23], and φ is a navigation function of the form:

$$\varphi(\mathbf{q},t) = \frac{\gamma(\mathbf{q},t)}{\left[\gamma^{\kappa}(\mathbf{q},t) + \beta(\mathbf{q},t)\right]^{1/\kappa}} \tag{1}$$

where $\kappa > 0$, γ is the distance to the 2D leader DMP, $\beta(\mathbf{q})$ is the product of obstacle areas described as functions resulting from visual feedback, $\mathbf{q}$ is the position vector in "Canonical" space and $\mathbf{U}$ is the vector of velocity inputs. Globally uniformly asymptotic convergence to the leader path is guaranteed within the body-part and proved in [23].

Adaptation of the controller's result on the operating surface is implemented with two bijective transformations. In particular, at each time step one point from the "Canonical" space is transformed in the visually segmented boundaries of each body part in Image space using affine transformation T_1 (rotation and anisotropic scaling). In this approach we define the **Image space** (IM) as the subspace of $\mathbb{R}^2_+$ with boundaries imposed by the resolution of the depth camera (e.g. Kinect camera's basic resolution is 512 x 424).

The latter step of the adaptation includes the transformation T_2 of the point on the Image space to the **Task space** (V), i.e. the $\mathbb{R}^3_+$ subspace which lies within the camera field of view (FOV). Basic assumption of this approach is that the body part which will be washed lies within the Task space

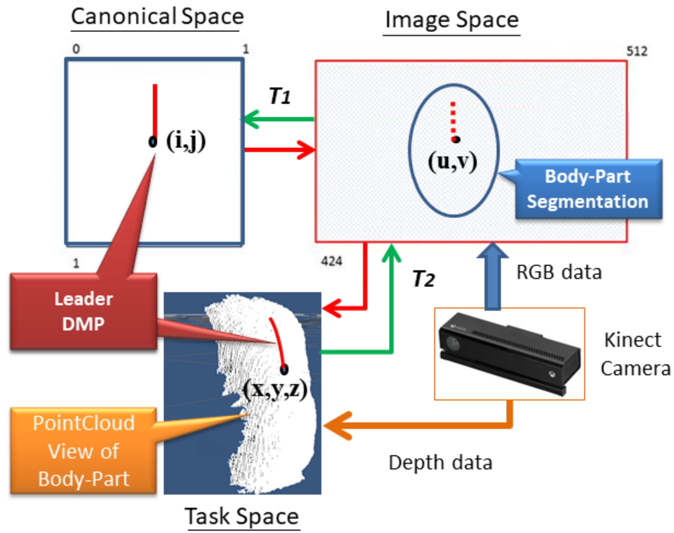

Fig. 2: Perception-based motion planning. A leader DMP point (i,j) from the Canonical space is transformed with bijective transformation T_1 to the point (u,v) of Image space and then with bijective transformation T_2 to the point (x,y,z) of the body-part. From the neighborhood of (x,y,z) we are able to calculate the reference orientation.

and the workspace of the robot at the same time. This means that the camera extrinsic parameters are known and the user is properly positioned and oriented with respect to the robot and the camera. In general, calibration of camera extrinsic parameters is a challenging task and can affect the accuracy of the algorithm, but this assumption can be fulfilled in a static set-up such as the I-Support system, Fig. 1.

Proposition 1: The transformation T_2 (which is represented by the camera projection) from the Image space (i.e. $IM = \{(u,v) : u \in [0,\ \mu], v \in [0,\ \nu]\}$, where μ, ν are the image width and height respectively, to the Task space (i.e. $V = \{(x,y,z) \in FOV\}$) at each time step is a bijection.

Proof: We provide a descriptive and intuitive proof. The basic idea results from the fact that a ray starting from the camera's optical center passes through the Image space and meets a point in the Task space. The latter is always true in an indoor environment. Therefore, $\forall (u,v) \in IM \Rightarrow \exists\,(x,y,z) \in v$. Using ray-casting technique it is easy to show that this point is unique, since the same ray cannot meet two points in the Task space at the same time. From the previous we can conclude that the transformation $T_2 : IM \rightarrow V$ is one-to-one (injective) and onto (surjective), so it is bijective. ■

In order to calculate the leader reference position and orientation, we use depth information of the pixel (u,v) and its neighboring pixels in the Image space, collecting a group of 3D points. This group of points in the Task space forms a small planar segment of the body part surface. For the reference position we calculate the percentile median point, whereas for the reference orientation we apply Principal Component Analysis to the covariance matrix of the collected points and use the eigenvectors as a local reference frame, as described in detail in [23].

One major issue of this approach is the visual occlusion of the surface, which occurs during the robot's operation. This

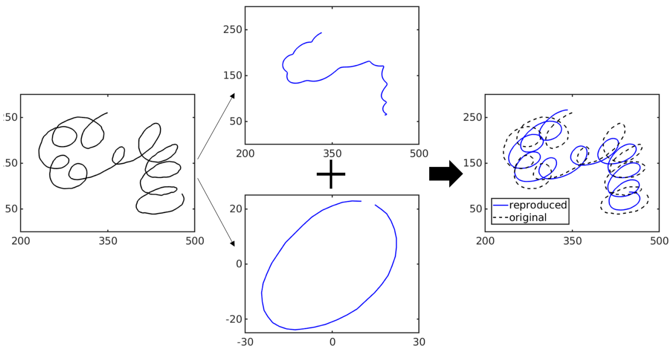

Fig. 3: The procedure of learning CC-DMP by human demonstration includes the separation of the motion into discrete and periodic part. **Left:** A demonstrated washing action **Middle:** Separation of the demonstrated motion into primitive discrete and periodic motions. **Right:** The reproduced motion by the CC-DMP method (blue) is similar to the demonstrated one (dashed).

problem is tackled by adjusting the size of the neighborhood of pixels mentioned above. The larger occlusion occurs, the larger the neighborhood should be so we can locally reconstruct the missing depth information from the surrounding pixel's depth.

Remark 1: The size of the robotic end-effector should be related to the curvature of the surface area. For example, if the robotic arm is large and causes a large visual occlusion, the local estimation of surface's curvature would be coarse in a highly curved area.

Remark 2: The reconstruction of the missing depth data from Kinect sensor can be solved with efficient image inpainting techniques presented in [25], [26]. The implementation of these computer vision algorithms is out of the scope of this paper. In our implementation we apply a planar fit in the missing data. In addition, the problem of missing visual data is highly reduced in multi-camera systems such as the I-Support, Fig. 1.

The described bijective transformations serve as a feedback to the controller as well. For example obstacle areas in the Task space (e.g. the hands of the user, or injuries on the back region) which are visible by the camera can be transformed back to the "Canonical" space by using the inverse procedure. In more detail, the black region in Fig. 5 represents a bandage on a body-part, which covers an injured region. This region is visually perceived and is transformed back and maximizes the values of the navigation function vector field in the corresponding coordinates. This modification will affect the execution (blue path) of a demonstrated leader DMP (red path) which passes through this area, preventing the robot from washing this sensitive area. Therefore, the described approach provides augmented perception properties to the washing system, which include user motion compensation, adaptability to different body-part size together with obstacle avoidance.

B. Learning and Motion Adaptation

Instead of hard-coded trajectories, the robot can achieve more human friendly washing motion by observing human demonstrations. We choose DMP to represent washing movement primitives as mentioned before, [7]. DMP is a damped-spring system coupled with a nonlinear term: $\tau \cdot \dot{v} = K \cdot (g - y) - D \cdot v + scale \cdot f$, with the spring factor K, the damping factor D and the nonlinear force term f, which can be learned by observing the demonstration. The temporal factor is τ, and g is the goal for discrete movement or the anchor point for periodic movement. Also, v, $\dot{v}$ and y specify the current state of the motion. The scaling factor $scale$ is used for changed g or start position y_0.

However, the traditional DMP cannot handle interactive actions, such as wiping a dynamic surface. Hence, we developed a leader-follower framework called Coordinate Change Dynamic Movement Primitive (CC-DMP), [15]. The idea of CC-DMP is that we learn the follower's DMP in the leader's coordinate system. In order to get the follower's motion in the global coordinate system, we multiply both sides of DMP transformation system with a coordinate transformation R^L_G, as in (2), where the superscript G denotes the global coordinate and L denotes the local coordinate. The leader's motion can also be encoded by an another DMP, which, together with the follower's DMP, constructs a leader-follower framework realizing the adaptation of the follower's movements to the leader's behavior.

$$\begin{aligned} \tau \cdot R^L_{G,t+1} \cdot \dot{v}^G &= R^L_{G,t} \cdot (K \cdot (g^G - y^G) - D \cdot v^G + \\ & \quad scale^G \cdot f^G) \\ \tau \cdot R^L_{G,t+1} \cdot \dot{y}^G &= R^L_{G,t} \cdot v^G \end{aligned} \tag{2}$$

In order to learn a washing action and keep its capacity of generalization, we first detect and separate periodic pattern of motion from its discrete part by performing signal analysis such as Fourier transformation described in [15]. The discrete part of a washing motion encodes the action direction such as top-down, left-right or some special movement. The periodic pattern encodes the functional primitive, which can be a cyclic motion, Fig. 3. In the extreme case, a motion whose periodic pattern has zero amplitude is a simple discrete motion. By this separation and representing both parts with DMP, we can modify the motion according to the user's preference or task constraints. Fig. 3 shows one simple way to extract both parts of a washing action and reproduce it with CC-DMP. The accuracy of the reproduction is dependent on both separation and learning. Despite the accurate learning properties of DMP, signal splitting might cause information loss. Nevertheless, signal splitting can be avoided by customizing the expert's demonstration strategy.

Hence, a complete washing system based on CC-DMP has multiple leaders and followers. In the high level, the user's movement is the leader and a periodic motion is the follower. However, since the user's movement is not predictable in the general case, we need sensor feedback to perceive the change of the surface instead of learning user's movement with a DMP. In the low level, the discrete part of the motion is leader and the periodic pattern is the follower, both of which can be learned by demonstration with DMPs. However, in our previous work [15], we have not included the perception of

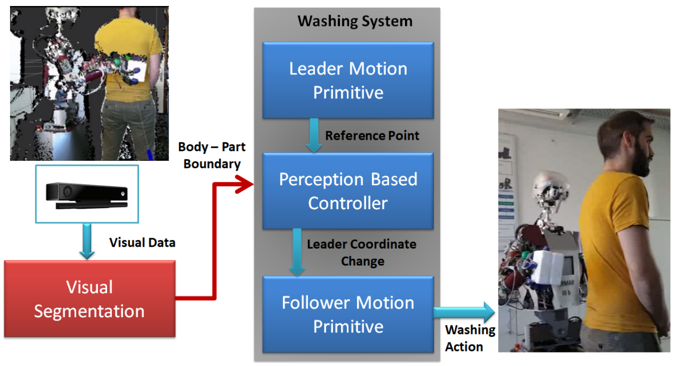

Fig. 4: Perception-based washing system. The output of the body-part visual segmentation and the depth data provided by the camera is the input of this system, while the output is the target washing action of the robot's end-effector.

the environment. All the experiments mentioned in that paper are conducted in the simulator, in which the movement of the user is known by the system.

IV. Perception-Based Washing System

If the desired washing movement is simple and predefined, a vision-based controller described in Sec. III-A can successfully adjust trajectory points one-by-one on a dynamical surface to generate a desired action. In the meantime, if the surface has known structure and does not change significantly during the motion evolution, CC-DMP described in Sec. III-B can flexibly generate complex trajectories learned by demonstration and adapt the movement to the surface's already modeled dynamic behavior.

However, in a washing case study, the size of each body-part differs among users and the body shape may change during the washing procedure, thus, we cannot generate an appropriate motion by pure imitation learning which cannot generalize for a relatively large change in the environment. On the other end, a pure perception-based controller cannot generate human-demonstrated washing trajectories. Furthermore, the preferences of each user may differ or change during the washing procedure, which requires the online modifications of motion parameters (e.g. amplitude, velocity). Therefore, this online motion modification requires the properties of a dynamical system such as CC-DMP.

Hence, we create a hierarchical washing system, shown in Fig. 4, by merging the described approaches, to achieve more robust behavior and to increase the capabilities of the system. In this system, we consider the discrete part of a washing action as the leader and the periodic pattern as the follower. The learned leader's motion primitive outputs a reference point in the "Canonical" space, which is followed by the Navigation Function controller. The output is adjusted in the body-part extends and then transformed on its surface, Fig. 2. The leader global pose is calculated by the analysis of camera's depth data in a small neighborhood of the visually segmented target area as described in Sec. III-A. In the latter step of this workflow, the follower movement primitive calculates the next point in the leader's local coordinate system, then transforms it to the global or robot's coordinate system. The final step is to use inverse kinematics to calculate the next required joint configurations of the robot and its low-level controller to drive the robot's end-effector to the next desired pose.

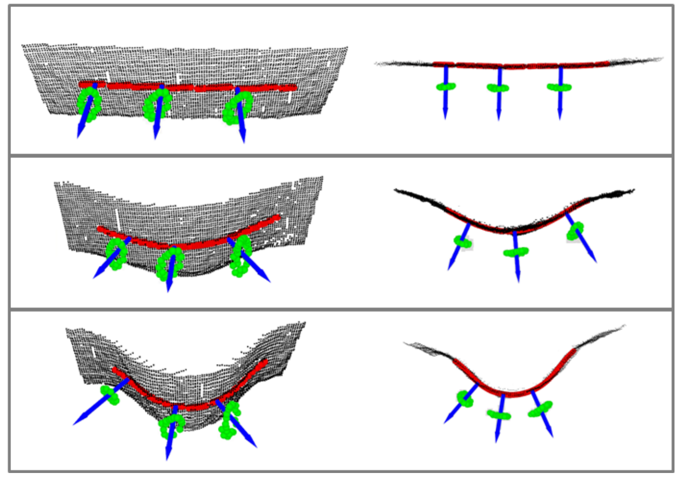

Fig. 6: Adaptation of a linear leader DMP (red) on a deformable surface (PointCloud view). The normal vector (blue) and the application of a follower periodic washing action is demonstrated on several segments of the path. Perspective and top views of a surface are depicted, subject to several unknown levels of deformation. **Top:** No deformation. **Middle:** Medium deformation. **Bottom:** High deformation.

The time and spatial adaptation, together with the decomposition of the learned washing actions allows for planning of a large repertoire of motions and adaptation on deformable surfaces as well. This repertoire includes different combinations of discrete and periodic actions, which may vary according to the washing sequence (e.g. pouring water, scrubbing, soaping etc.) decided by the user or the healthcare expert. It also includes the capacity of the perception-based system for on-line adaptation on large and a-priori *unknown surface deformations* of the target part. An indicative example is presented in Fig. 6, in which a linear discrete motion (red) is adapted on a surface formed by a plane paper of unknown, but visually perceived curvature and deformation. The estimated local vector (blue) and the execution of a cyclic periodic pattern (green) are also demonstrated in several segments of the leader's path. The local curvature estimation at each time step not only compensates the surface's motion and deformation, but also permits the regulation of the perpendicular distance of the robot's end-effector to the surface. This regulation enables the execution of actions that involve physical contact (e.g. scrubbing), thus also indirectly involving the application of forces without any additional feedback, as well as actions that involve no contact with the surface (e.g. pouring water).

V. Experimental Study

An experimental setup, suitable for validating the performance of the proposed approach with similar configuration to the I-Support system, is employed that includes a single Kinect-v2 camera providing depth data for the back region of the subject, with accuracy analyzed in [27]. The segmentation of the washing surface is implemented, for the purposes of the following experiments with a color filter to the pixels of

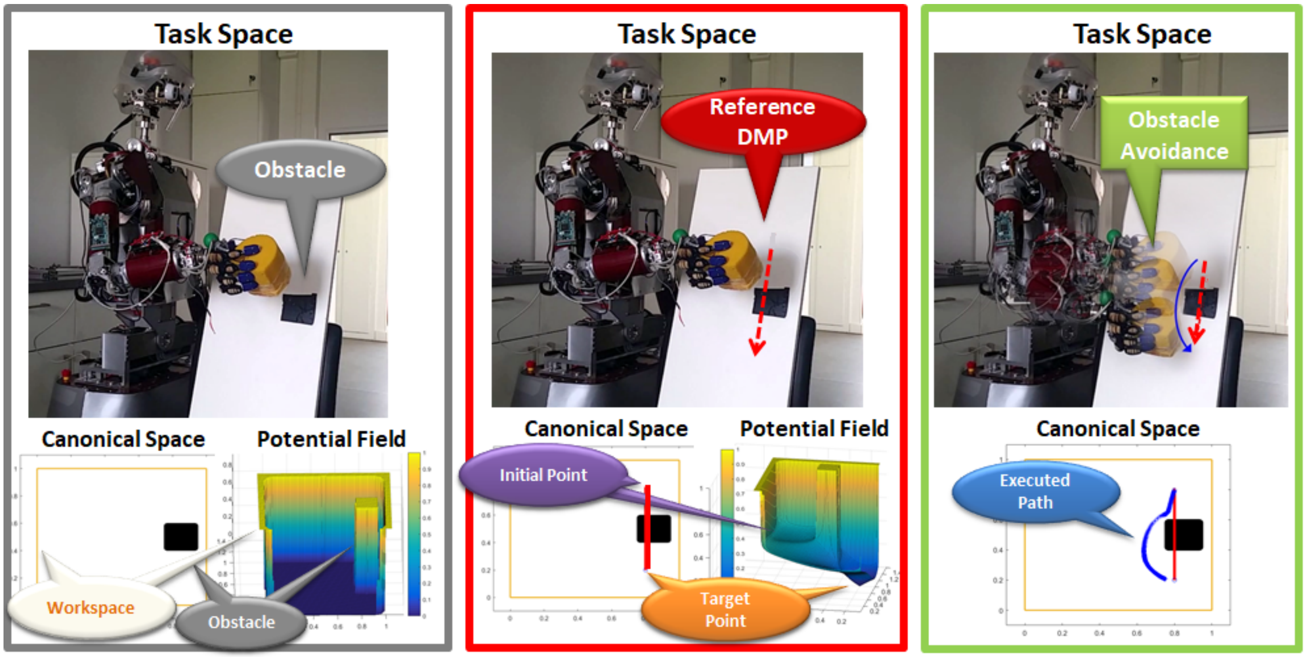

Fig. 5: **Scenario I:** ARMAR-III wipes a static whiteboard. **Left:** An obstacle area (e.g. injury depicted with black patch) is detected in the Task space and transformed back to the Canonical space. The Navigation Function potential field is maximized in the corresponding area and the boundary of the workspace. **Middle:** The leader DMP path (red) is defined and an attractive vector field leads to the target point. **Right:** The controllers output is the blue path and is executed by the robot, avoiding the sensitive injured area. After the obstacle avoidance the end-effector's motion converges again the indicated linear motion primitive.

the image obtained from Kinect-v2 camera. The setup also includes an ARMAR-III robot.

A. Experimental Scenarios

In order to test our methods, we consider the following experiments:

- **Scenario I:** ARMAR-III wipes a static whiteboard. The discrete part of this washing action is a vertical top-down movement. Obstacle avoidance is demonstrated in this experiment, Fig. 5.
- **Scenario II:** ARMAR-III wipes a dynamic whiteboard, which is held by a person and rotated/translated from time to time, Fig. 7.
- **Scenario III:** ARMAR-III wipes a male subject's back. He is moving his back during the experiment to demonstrate the adaptation of the robot motion to the subject's movement. For safety reasons, ARMAR-III has no real contact with the person, Fig. 10- 11.

We choose a whiteboard for this experimental procedure in order to bypass difficulties imposed by image segmentation, which is out of the scope of this paper, and as a reference surface for validation purposes.

B. Results & Discussion

In Figures 5 – 11, the results of all experimental scenarios are presented. ARMAR-III uses all 7 DOFs of its right arm and its hip yaw joint to generate functional washing actions. A washing task with obstacle avoidance (Scenario I) is demonstrated in Fig. 5, in which an obstacle area indicated with a black patch (e.g. an injury on the back region) intersects with the motion of the robot if a leader top-down path (red) is directly executed. In particular, the injured area is visually perceived and is transformed back to the Canonical space using the inverse T_1 and T_2 transformations. This information is inserted into the Navigation function as in (1), maximizing the values of the potential field in the corresponding coordinates. As soon as the leader DMP path (red) is defined, an attractive vector field is formed and leads to the target point. The robot executes a modified path (blue) and avoids to wash the obstacle area.

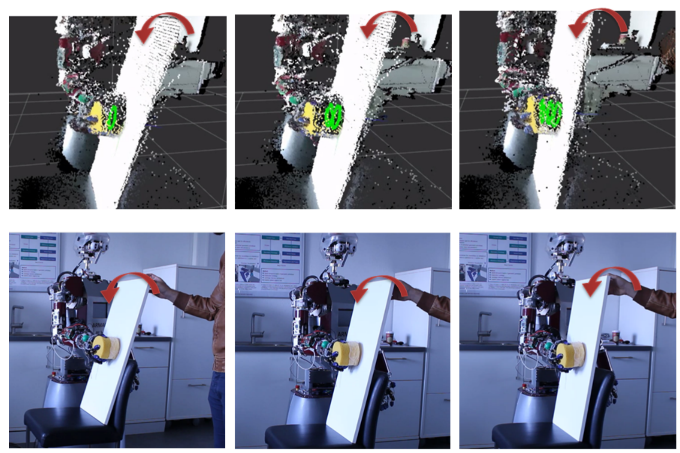

Fig. 7: **Scenario II:** ARMAR-III is wiping a dynamic whiteboard and a person moves the board. ARMAR-III is holding a yellow sponge, which keeps contact with the surface. The wiping movement is adapted to the surface's motion. **Top:** PointCloud view of the whiteboard and the robot's end-effector showing instances of the adaptation of the wiping motion with the green trajectory. **Bottom:** Side camera view of the wiping action indicating with the red arrow the motion of the whiteboard implied by the human.

In addition, the results of Scenario II are intuitively vi-

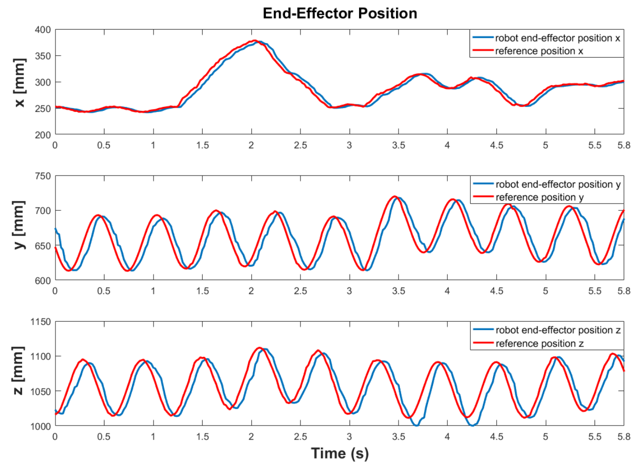

Fig. 8: Evolution of the reference position for the wiping motion and the executed robot end-effector position, during **Scenario II** that involves a moving planar surface.

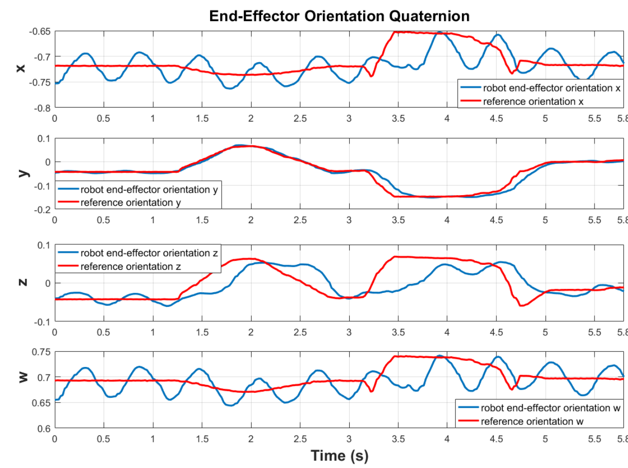

Fig. 9: Evolution the reference orientation for the wiping motion and the executed robot end-effector orientation, during **Scenario II** that involves a moving planar surface

sualized in the top of Fig. 7, using the Point cloud view provided by the Kinect camera and a sequence of green points showing the evolution of the circular motion and its adaptation to the movement of the whiteboard. In order to validate the performance of the robot during the periodic action, we compared the reference path computed by the washing system with the end-effector path calculated from the robot's forward kinematics. More specifically, Fig. 8, 9 depict the time evolution of the reference pose (position & orientation quaternion) in red color and the executed robot end-effector pose in blue color during the wiping motion of the planar surface (**Scenario II**). It is apparent, that the robot manages to follow the surfaces pose compensating with it's motion and simultaneously execute the wiping action. Furthermore, the executed path quickly converges to reference with bounded error both in position and orientation.

In Scenario III, the robot executes the washing trajectory over the back region of a male subject. In more detail, in Fig. 10 the subject moves to the right, while in Fig. 11 he moves backwards. In both cases the robot with the aid of the proposed motion planning approach manages to compensate

Fig. 10: **Scenario III:** ARMAR-III is wiping a male subject's back region. The subject is moving to the right and the robot follows the motion. **Left:** Side camera view of the subject performing a translation to the right. **Right:** Zoomed Point cloud view of the experiment highlighting the adaptation of the wiping motion (green trajectory) to the movement of the subject. Safe distance is kept between the back and ARMAR-III's end-effector.

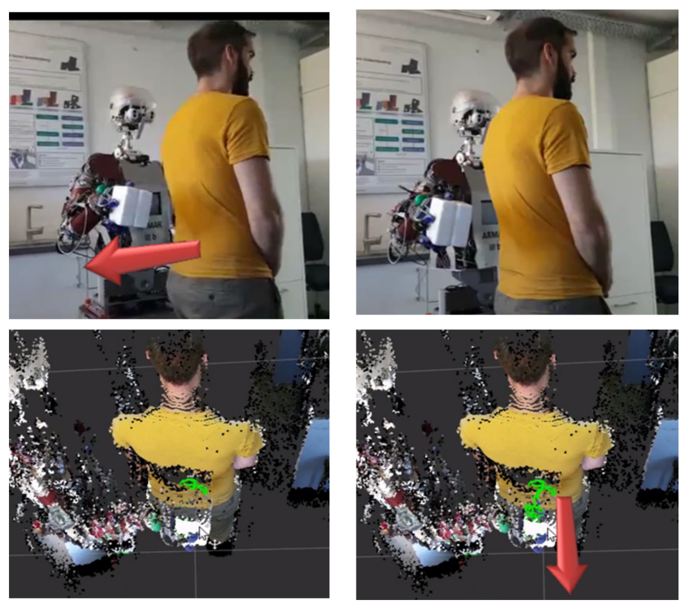

Fig. 11: **Scenario III:** ARMAR-III is wiping a male subject's back region. The subject is moving backwards and the robot follows the motion. **Top:** Side camera view of the subject performing a translation backwards. **Bottom:** Point cloud view of the experiment highlighting the adaptation of the wiping motion (green trajectory) to the movement of the subject. Safe distance is kept between the back and ARMAR-III's end-effector.

with the movement of the subject, without interrupting the washing task indicating the applicability of this approach to real life scenarios.

VI. Conclusion and Future Work

This paper presents a vision-based washing system, which is capable of adapting the motion of a robotic end-effector to a moving and non-rigid surface, such as the back region of a person. This goal is achieved by merging two methods, a leader-follower motion primitive framework (CC-DMP) with a visual perception based controller. This fusion carries out human-friendly washing tasks, by incorporating the expertise of health-care experts with imitation learning techniques, while enhancing on-line adaptation to dynamic moving and deformable objects (in our case, body parts). We conducted several experiments with a humanoid robot ARMAR-III, which applies washing actions on a planar surface (either static or moving) and over the back region of a subject.

For further research, we intent to ameliorate this system by expanding the repertoire of learned washing actions and smoothly integrate them to the technical requirements of a multi-camera robotic system, which will reduce the occlusion problem. We can also make the system more interactive by applying shared control techniques and by letting the user adjust on the fly the parameters of the robotic motion and the contact forces according to his/her feeling. In the future, we aim towards smoother and more robust contact with the human body, by incorporating force based control approaches [28]–[30] to the current method, fusing information from both vision and tactile sensors.

The system will be further integrated and experimentally validated using the soft arm shower robot developed in the frames of the I-Support project.

REFERENCES

[1] D. D. Dunlop, S. L. Hughes, and L. M. Manheim, "Disability in activities of daily living: patterns of change and a hierarchy of disability," *American Journal of Public Health*, vol. 87, pp. 378–383, March 1997.

[2] (2016) Oasis seated shower system. [Online]. Available: www.adl-solutions.com/Oasis%2DSeated%2DShower%2DSystem%2Dfor%2DOptimal%2DAccessible%2DBathing%2DTempe%2DAZ.html

[3] (2016) Robotics care poseidon. [Online]. Available: http://www.roboticscare.com/robotics-care-poseidon/

[4] T. Hirose, S. Fujioka, O. Mizuno, and T. Nakamura, "Development of hair-washing robot equipped with scrubbing fingers," in *IEEE International Conference on Robotics and Automation (ICRA)*, May 2012, pp. 1970–1975.

[5] B. Driessen, H. Evers, and J. v Woerden, "Manus—a wheelchair-mounted rehabilitation robot," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 215, pp. 285–290, March 2001.

[6] S. Katz, A. Ford, R. Moskowitz, B. Jackson, and M. Jaffe, "Studies of illness in the aged: The index of adl: a standardized measure of biological and psychosocial function," *JAMA*, vol. 185, pp. 914–919, September 1963.

[7] S. Schaal, "Dynamic movement primitives - a framework for motor control in humans and humanoid robots," in *The International Symposium on Adaptive Motion of Animals and Machines*, March 2006, pp. 261–280.

[8] S. Calinon, F. Guenter, and A. Billard, "On learning, representing and generalizing a task in a humanoid robot," *IEEE Transactions on Systems, Man and Cybernetics, Part B. Special issue on robot learning by observation, demonstration and imitation*, vol. 37, pp. 286–298, March 2007.

[9] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, "Probabilistic movement primitives," in *Advances in Neural Information Processing Systems 26*. Curran Associates, Inc., 2013, pp. 2616–2624.

[10] M. Mühlig, M. Gienger, and J. J. Steil, "Interactive imitation learning of object movement skills," *Autonomous Robots*, vol. 32, pp. 97–114, Feburary 2012.

[11] T. Kulvicius, M. Biehl, M. J. Aein, M. Tamosiunaite, and F. F. Wörgötter, "Interaction learning for dynamic movement primitives used in cooperative robotic tasks," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1450–1459, December 2013.

[12] A. Gams, B. Nemec, A. J. Ijspeert, and A. Ude, "Coupling movement primitives: Interaction with the environment and bimanual tasks," *IEEE Transactions on Robotics*, vol. 30, no. 4, pp. 816–830, 2014.

[13] H. B. Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters, "Interaction primitives for human-robot cooperation tasks," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, September 2014, pp. 2831–2837.

[14] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: Learning attractor models for motor behaviors," *Neural Computation*, vol. 25, no. 2, pp. 328–373, February 2013.

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

[16] C. Mandery, O. Terlemez, M. Do, N. Vahrenkamp, and T. Asfour, "Unifying representations and large-scale whole-body motion databases for studying human motion," *IEEE Transactions on Robotics*, vol. 32, no. 4, pp. 796–809, August 2016.

[17] D. C. Conner, A. Greenfield, P. N. Atkar, A. A. Rizzi, and H. Choset, "Paint deposition modeling for trajectory planning on automotive surfaces," *IEEE Transactions on Automation Science and Engineering*, vol. 2, no. 4, pp. 381–392, October 2005.

[18] P. Fiorini and E. Prassler, "Cleaning and household robots: A technology survey," *Autonomous robots*, vol. 9, no. 3, pp. 227–235, December 2000.

[19] R. Ginhoux, J. Gangloff, M. De Mathelin, L. Soler, M. M. A. Sanchez, and J. Marescaux, "Beating heart tracking in robotic surgery using 500 hz visual servoing, model predictive control and an adaptive observer," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2004, pp. 274–279.

[20] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *arXiv preprint arXiv:1606.00915*, June 2016.

[21] L. C. Chen, Y. Yang, J. Wang, W. Xu, and A. L. Yuille, "Attention to scale: Scale-aware semantic image segmentation," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[22] S. E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 4724–4732.

[23] A. C. Dometios, X. S. Papageorgiou, A. Arvanitakis, C. S. Tzafestas, and P. Maragos, "Real-time end-effector motion behavior planning approach using on-line point-cloud data towards a user adaptive assistive bath robot," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, September 2017.

[24] T. Asfour, K. Regenstein, P. Azad, J. Schröder, N. Vahrenkamp, and R. Dillmann, "ARMAR-III: An Integrated Humanoid Platform for Sensory-Motor Control," in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Genova, Italy, December 2006, pp. 169–175.

[25] M. Stommel, M. Beetz, and W. Xu, "Inpainting of missing values in the kinect sensor's depth maps based on background estimates," *IEEE Sensors Journal*, vol. 14, no. 4, pp. 1107–1116, 2014.

[26] L. Petrescu, A. Morar, F. Moldoveanu, and A. Moldoveanu, "Kinect depth inpainting in real time," in *International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, June 2016, pp. 697–700.

[27] L. Yang, L. Zhang, H. Dong, A. Alelaiwi, and A. El Saddik, "Evaluating and improving the depth accuracy of kinect for windows v2," *IEEE Sensors Journal*, vol. 15, no. 8, pp. 4275–4285, 2015.

[28] A. Montebelli, F. Steinmetz, and V. Kyrki, "On handing down our tools to robots: Single-phase kinesthetic teaching for dynamic in-contract tasks," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015.

[29] M. Racca, J. Pajarinen, A. Montebelli, and V. Kyrki, "Learning in-contact control strategies from demonstration," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, October 2016.

[30] Y. Zhou, M. Do, and T. Asfour, "Learning and force adaptation for interactive actions," in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*. IEEE, 2016, pp. 1129–1134.