On-line periodic movement and force-profile learning for adaptation to new surfaces

Andrzej Gams*, Martin Do†, Aleš Ude*, Tamim Asfour† and Rüdiger Dillmann†

*Department of Automation, Biocybernetics and Robotics, Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia
†Institute for Anthropomatics, Humanoids and Intelligence Systems Laboratories, Karlsruhe Institute of Technology, Karlsruhe, Germany

Abstract—To control the motion of a humanoid robot along a desired trajectory in contact with a rigid object, we need to take into account forces that arise from contact with the surface of the object. In this paper we propose a new method that enables the robot to adapt its motion to different surfaces. The initial trajectories are encoded by dynamic movement primitives, which can be learned from visual feedback using a two-layered imitation system. In our approach these initial trajectories are modified using regression methods. The data for learning is provided by force feedback. In this way new trajectories are learned that ensure that the robot can move along the object while maintaining contact and applying the desired force to the object. Active compliance can be used more effectively with such trajectories. We present the results for both movement imitation and force profile learning on two different surfaces. We applied the method to the ARMAR-IIIB humanoid robot, where we use the system for learning and imitating a periodic task of wiping a kitchen table.

I. INTRODUCTION

The use of a humanoid robot in a kitchen is becoming a viable scenario [1], [2]. The ability to learn new tasks in a natural way and without the need of an expert is one of the key functionalities for a truly effective and useful system. In this paper we present a two-layered system for movement imitation, which utilizes both visual and force feedback. We applied the system on a humanoid robot to achieve natural learning and effective use in a kitchen environment for the task of wiping flat or uneven surfaces.

The task of wiping a kitchen table is easy for a human, but involves several challenges for a humanoid robot. The robot has to perform periodical movement in contact with a heavy rigid object. Besides the height of the object, which can differ, depending on the height of kitchen counter or kitchen table, the location relative to the base of the robot is not fixed. Furthermore, wiping movement can be of different types. We therefore need effective methods to learn from humans and to generalize from existing knowledge.

Imitating movement with robots is a common approach [3], [4]. Different trajectory encodings have been proposed in the literature, e.g. splines [5] or dynamic movement primitives (DMP) [6]. To learn the associated parameters methods like statistical regression [7] or reinforcement learning [8] have been proposed. In this work we use dynamic movement primitives. DMPs use a set of kernel functions to reproduce trajectories. They have favorable properties in the sense of continuous and smooth trajectories in the presence of obstacles [9] and their modulation characteristics make them especially useful for learning whole families of similar movements from a single demonstration.

For the task of learning of periodic movements we need to extract the frequency and the waveform of the demonstrated movement. A two-layered movement imitation system based on adaptive frequency oscillators in a feedback loop and a DMP formulation for waveform learning allows easy, natural, and computationally inexpensive learning and replaying of demonstrated trajectories [9]. The two-layered system of movement imitation can be used for learning of periodic trajectories in as many degrees of freedom as we can effectively measure.

When imitating human movement with a humanoid robot we can either learn movements in joint or in task space. Both have their advantages. For example when measuring the movement in joint space we have direct control over separate joints (see for example [10]). Measuring in task space, on the other hand, hands over the control of separate joints to on-line inverse kinematics algorithms. Different algorithms that try to maintain human-like postures when applying inverse kinematics algorithms exist, e.g. by minimizing the sum of exerted torques [11] or by optimizing the manipulability index [12]. As the influence of external forces on the movements is specified in task space, we specify trajectories in task space.

Contact with the environment requires tactile sensing abilities, such as for example a force-torque sensor. Controlling rigid robots while in contact with the environment is a difficult task [13], and often passive compliance elements are used to reduce the complexity. Although wiping a kitchen table, with a sponge might be successful, adding active compliance increases the versatility of feasible tasks and allows faster and more precise movement along the surface. This can be further increased by learning the force-profile of a periodic task.

The paper is organized as follows. In Section II we give a review of the two-layered movement imitation system and present force-profile learning, which allows performing periodic tasks while in contact with an object of an arbitrary shape.
In Section III we present the kitchen experiment with the task of wiping an arbitrary shaped surface. We give conclusions in Section IV.

II. PERIODIC TRAJECTORY LEARNING

In this section we give a recap of the two-layered movement imitation system [9], which allows learning of periodic trajectories and the frequency of execution. After learning of periodic movements in task space with visual feedback we apply the system to learn the force-profile of objects by means of force feedback.

A. Learning of movements

The structure of the imitation system has two layers, as is presented in Fig. 1. The first layer – the Canonical dynamical system – extracts the fundamental frequency of the input signal. The second layer of the system is the Output dynamic system, which outputs the desired trajectory for the robot. The input signal can be any measured periodic quantity, in our case it is the visual feedback.

The Canonical dynamical system has to extract the fundamental frequency $\Omega$ and the phase signal $\Phi$ from the input signal. As the basis of the Canonical dynamical system we use a set of $M$ adaptive frequency phase oscillators [14] in a feedback structure [15], as shown in Fig. 2.

The feedback structure of $M$ adaptive frequency phase oscillators is governed by

$$\dot{\phi}_i = \omega_i - K e(t) \sin(\phi_i),$$  

$$\dot{\alpha}_i = -K e(t) \sin(\phi_i),$$  

$$e(t) = y_m(t) - \bar{y}(t),$$  

$$\bar{y}(t) = \sum_{i=1}^{M} \alpha_i \cos(\hat{\phi}_i),$$  

$$\alpha_i = \eta \cos(\hat{\phi}_i) e(t),$$

where $K$ is the coupling strength, $\phi_i, i = 1, \ldots, M$ is the phase of a separate oscillator, $e(t)$ is the input into the oscillators, $y_m$ is the input signal, $M$ is the number of oscillators, $\alpha_i$ is the amplitude associated with the $i$-th oscillator, and $\eta$ is the learning constant.

Each of the oscillators of the feedback structure receives the same input signal, which approaches zero when the weighted sum of separate frequency components $\dot{\bar{y}}$ approaches the input signal (see Fig. 2). Such a feedback structure can extract $M$ frequency components of the input signal. Since we imitate periodic human movement and are interested only in the basic frequency, we use $M = 3$ oscillators. The feedback structure is followed by a logical algorithm, which chooses the basic or fundamental frequency $\Omega$ from the three extracted frequencies.

The Canonical dynamical system can be used as an imitation system by itself, see [15], [16]. Adding the Output dynamical system allows exploiting the modulations and robustness properties of DMPs, which are important for movements in contact with rigid surfaces.

The Output dynamical system is a periodic DMP structure, where the waveform is anchored to the extracted phase. As illustrated in Fig. 1, the inputs are 1) the signal we are trying to learn and imitate, given by triplets of position, velocity and acceleration in discrete time steps; 2) the frequency $\Omega$ and 3) the phase $\Phi$. The latter two are provided by the Canonical dynamical system.

A periodic DMP structure is given as a second order system of differential equations

$$\dot{s} = \Omega \left( \alpha_s (\beta_s (g - y) - s) + \sum_{i=1}^{N} \frac{\psi_i w_i r}{\sum_{i=1}^{N} \psi_i} \right),$$  

$$\dot{\bar{y}} = \Omega s.$$  

The system above includes a nonlinear term specified by basis functions $\psi_i$. The nonlinear term provides the desired waveform of the output by multiplying the Gaussian-like kernel functions $\psi_i$ and the weights $w_i$. In (6) $\Omega$ is the frequency given by the Canonical dynamical system, and $\alpha_s = 12$ and $\beta_s = 3$ are positive constants which ensure that the system monotonically varies to the trajectory that is oscillating around an anchor point $g$. $N$ is the number of kernel functions and $r$ is the amplitude scaling factor. The kernel functions are given by

$$\psi_i = \exp \left( h (\cos (\phi - c_i) - 1) \right),$$

where $h$ determines their width, and $c_i$ which distribute them over the course of one period. $c_i$ are equally spaced between 0 and $2\pi$ in $N$ steps.

The waveform of the learned movement is updated incrementally in every cycle as is explained for the force profile.
learning in the next section. We use visual feedback to learn
the initial periodic movement, which we later modify by force
profile adaptation.

B. Force profile adaptation

Learning of a movement that brings the robot into contact
with the environment must be based on force control, otherwise
to which there can be damage to the robot or the object to
which the robot applies its force. In the context of wiping
a table or any other object of arbitrary shape constant contact
with the object is required. To teach the robot the necessary
movement, we decoupled the learning of the movement from
the learning of the shape of the object. We first apply the
described two-layered movement imitation system to learn the
desired trajectories by means of visual feedback. We then use
force-feedback to adapt the motion to the shape of the object
that the robot acts upon.

Periodic movements can be of any shape, yet wiping can
be effective with simple one dimensional left-right movement,
or circular movement. Once we are satisfied with the learned
movement, we can reduce the frequency of the movement by
or the initial periodic movement, which we later modify by force
vector \( F = \sqrt{F_x^2 + F_y^2 + F_z^2} \) for the feedback instead of
\( F_e \) in (11). This way the robot can move upwards every time
it hits something, for example the side of a sink. No contact
should be made from above, as this will make the robot press
up harder and harder.

The learning of the force profile is done by modifying the
weights \( w_i \) for the selected degree of freedom \( y_j \) in every
time-step by incremental locally weighted regression [7]. The
input in this case is \( Y(t) \) from (10). The target for learning is
determined by
\[
\Omega^{-1} y_j = \alpha (b_j - y_j) - \frac{1}{\Omega} \tilde{y}_j,
\]
which is obtained by matching \( y \) from (6–7) to \( y_j \), \( s \) to \( \tilde{s} \), and \( \tilde{s} \) to \( \tilde{y}_j \). Given the target data \( f_{\text{arg}}(t) \) and \( r(t) \), with the frequency \( \Omega \)
and the phase \( \Phi \) the values set for the playback, \( w_i \), \( i = 1, \ldots, N \)
are updated in recursion by
\[
w_i(t + 1) = w_i(t) + \Psi P_i(t + 1) r_i(t) \varepsilon_i(t),
\]

\[
P_i(t + 1) = \frac{1}{\lambda} \left( P_i(t) - \frac{P_i(t) r_i(t)^2}{\varepsilon_i(t) P_i(t + 1) r_i(t)^2} \right),
\]

\[
\varepsilon_i(t) = f_{\text{arg}}(t) - w_i(t) r_i(t).
\]
The recursion is started with \( w_i = 0 \) and \( P_i = 1 \), where \( i = 1, \ldots, N \). \( P \) is in general the inverse covariance matrix and \( \lambda = 0.99 \) is the forgetting factor.

The \( K_F \) matrix controls the behaviour of the movement. The
correcting movement has to be fast enough to move away
from the object if the robot hand encounters sufficient force,
and at the same time not too fast so that it does not produce
instabilities due to the discrete-time sampling when in contact
with an object. A dead-zone of response has to be included,
for example \( |F| < 1 \) N, to take into account the noise. We
empirically set \( k_F = 20 \), and limited the force feedback to
allow maximum linear velocity of 120 mm/s.

Feedback from a force-torque sensor is often noisy due
to the sensor itself and mainly due to vibrations of the
robot. A noisy signal is not the best solution for the learning
algorithm because we also need time-discrete first and second
derivatives. The described active compliance algorithm uses
the position of the end-effector as input, which is the integrated
desired velocity and therefore has no difficulties with the noisy
measured signal.

Having adapted the trajectory to the new surface enables
very fast movement with a constant force profile at the contact
of the robot/sponge and the object, without any time-sampling
and instability problems that may arise when using compliance
control only. Furthermore, we can still use the compliant control
once we have learned the shape of the object. Active compliance,
combined with a passive compliance of a sponge, and the
modulation and perturbation properties of DMPs, such as slow-down feedback [9], allow fast and safe execution of
periodic movement while maintaining a sliding contact with
the environment.

III. EXPERIMENTAL EVALUATION

In this section we describe the experimental setup and
present the results on the ARMAR-IIIB humanoid robot.
A. Experimental setup

1) ARMAR-IIIb humanoid robot: The humanoid robot ARMAR-IIIb, which serves as the experimental platform in this work, is a copy of the humanoid robot ARMAR-IIIa [1]. From the kinematics point of view, the robot consists of seven subsystems: head, left arm, right arm, left hand, right hand, torso, and a mobile platform. The head has seven DOF and is equipped with two eyes, which have a common tilt and can pan independently. Each eye is equipped with two digital color cameras, one with a wide-angle lens for peripheral vision and one with a narrow-angle lens for foveal vision. The upper body of the robot provides 33 DOF: 2·7 DOF for the arms and three DOF for the torso. The arms are designed in an anthropomorphic way: three DOF for each shoulder, two DOF in each elbow and two DOF in each wrist. Each arm is equipped with a five-fingered hand with eight DOF. The locomotion of the robot is realized using a wheel-based holonomic platform.

2) Vision and force feedback: In order to obtain reliable motion data of a human wiping demonstration through observation by the robot, we exploited the color features of the sponge to track its motion. Using the stereo camera setup of the robot, the implemented blob tracking algorithm based on color segmentation and a particle filter framework provides a robust location estimation of the sponge in 3D. The resulting trajectories were captured with a frame rate of 30 Hz.

For learning of movements we first define the area of demonstration by measuring the lower-left and the upper-right position within a given time frame. These coordinates make a rectangular area (marked with dashed lines) where the robot tracks the demonstrated movements.

B. Evaluation

1) Learning of periodic movement: The two-layered movement imitation system extracts the frequency and learns one period of the waveform. As we can see from the results in Fig. 4, the system extracts three frequency components, of which one is the offset (0 Hz). As we expect relatively simple waveform for wiping movements we only use \( M = 3 \) oscillators and always choose the middle of the three frequencies for the fundamental frequency. This way the chosen frequency is never associated with the offset or the higher frequency components present. We simply reset the oscillators if more than one frequency component converges to the 0 Hz frequency (see also [9]).

The learning of the waveform happens simultaneously with the learning of the frequency. Once we are satisfied with the performed movement, we stop the learning. After learning we can modulate the frequency value \( \Omega \).

Fig. 5 shows the adaptation of the waveform during learning for both dimensions in the \( xy \) plane. The input signal is transformed from task-space coordinates to the normalized deviation from the center of the visual learning frame (see Fig. 3). As we can see the waveform adapts quickly. Some delay can be observed, which can be attributed to the slow frame-rate of the visual feedback at 30 Hz, and the learning algorithm itself. Once we are satisfied with the waveform we stop the learning, and can then arbitrarily scale and offset the learned movement.

2) Learning of force profile: Once we stop the movement learning we start learning the force-profile, with which we learn the trajectory required to maintain constant contact between the sponge and the robot. A video showing the experiment can be found at [http://www.ijs.si/~aude/ForceLearning.mov](http://www.ijs.si/~aude/ForceLearning.mov).

Fig. 6 shows the results of learning the force-profile for a flat surface. As the robot grasps the sponge, its orientation and location are unknown to the robot, and the tool center point (TCP) changes. Should the robot simply perform a planar trajectory it would not ensure constant contact with the table.
As we can see from the results, the hand initially moves down until it makes contact with the surface. The force profile later changes the desired height by approx. 5 cm within one period. After the learning (stopped manually, marked with a vertical dashed line) the robot maintains such a profile. A manual increase in frequency was introduced to demonstrate the ability to perform the task at an arbitrary frequency. The bottom plot shows the measured length of the force vector $|F|$. As we can see the force vector keeps the same force profile, even though the frequency is increased. No increase in the force profile proves that the robot has learned the required trajectory.

Fig. 7 shows the results for a bowl-shaped object. As we can see from the results the height of the movement changes for more than 6 cm within a period. The learned shape (after the vertical dashed lined) maintains the shape of a bowl, but has an added local minimum. This is the result of the dead-zone within the active compliance, which comes into effect when going up one side, and down the other side of the bowl. No significant change in the force profile can be observed in the bottom plot after a manual increase in frequency. Some drift, as the consequence of an error of the sensor and of wrist control on the robot, can be observed.

IV. CONCLUSION

We have shown the use of a two-layered imitation system with decoupled periodic movement learning and force profile learning. The main contribution of the paper is a new algorithm for force profile learning. We have demonstrated the usefulness of our approach for the implementation of a wiping behaviour. Learning of the force-profile (or of the shape of the object), in
combination with both active and passive compliance, allows faster and more robust execution of tasks. Our approach thus implements an important behaviour for the future.

ACKNOWLEDGMENT

The work described in this paper was partially conducted within the EU Cognitive Systems projects PACO-PLUS (FP6-027657) and GRASP (FP7-215821) funded by the European Commission. It was partially funded by the Slovenian Research Agency grant Z2-3672.

REFERENCES


