Manipulation Learning on Humanoid Robots

Andrej Gams[1] · Tadej Petrič[1] · Bojan Nemec[1] · Aleš Ude[1,2]

Abstract

Purpose of Review The ability to autonomously manipulate the physical world is the key capability needed to fulfill the potential of cognitive robots. Humanoid robots, which offer very rich sensorimotor capabilities, have made giant leaps in their manipulation capabilities in recent years. Due to their similarity to humans, the progress can be partially attributed to the learning by demonstration paradigm. Supplemented by the autonomous learning methods to refine the demonstrated manipulation actions, humanoid robots can effectively learn new manipulation skills. In this paper we present continuous effort by our research group to advance the manipulation capabilities of humanoid robots and bring them to autonomously act in an unstructured world.

Recent Findings The paper details progress in the area of humanoid robot learning, ranging from trajectory imitation, motion adaptation in order to maintain feasibility and stability, and learning of dynamics to statistical generalization of actions, autonomous learning, and end-to-end vision-to-action learning that exploits deep neural networks.

Summary With the focus on manipulation, the presented research provides the means to overcome the complexity behind the problem of engineering manipulation skills on robots, especially humanoid robots where programming by demonstration is most effective.

Keywords Robot learning · Humanoid robots · Robot manipulation · Autonomous learning

Introduction

Achieving cognition relies on robots to provide embodiment — embodiment with rich and complex motor skills that provide means to interact with and manipulate in the physical world [1]. The ability to autonomously manipulate the physical world is the key capability needed to fulfill the potential of cognitive robots. It has an enormous potential for various applications, where autonomous robots can be deployed in all kinds of unstructured and even hazardous environments. Applications can range far beyond today's utilization of robots in factories; from helping in households, hospitals, and care facilities, to work in radioactive environments or even in space [1, 2].

As one of the key aspects of robotics, manipulation and learning of manipulation has been at the center of research for a long time [3]. The developed applications vary in complexity, venue, and robotic mechanisms [4, 5]. In the long run, research has been progressing towards evaluating approaches and results at the complete system level rather than focusing on the performance of separate components, or even component parts [6].

Manipulation on humanoid robots, which typically offer very rich sensorimotor capabilities [7, 8], represents one of the most complex settings. Physical capabilities of humanoid robots have drastically increased over the last decade [7, 9]. Research in the field of robot manipulation, also applied to humanoid robotics, has brought about rapid

✉ Aleš Ude
ales.ude@ijs.si

Andrej Gams
andrej.gams@ijs.si

Tadej Petrič
tadej.petric@ijs.si

Bojan Nemec
bojan.nemec@ijs.si

1 Department of Automatics, Biocybernetics and Robotics, Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia

2 Faculty of Electrical Engineering, University of Ljubljana, Tržaška cesta 25, 1000 Ljubljana, Slovenia

advances in the manipulation capabilities. However, general purpose and general manipulation skills on humanoid robots remain open research questions [5].

Several aspects must be considered when realizing an effective manipulation learning on humanoid robots. First of all, due to their similarity to humans, humanoid robots can learn manipulation skills by observing human performance. Learning (or programming) by demonstration has long been an important topic in humanoid robotics research [10, 11]. By observing human performance, a humanoid robot can compute an initial model of the desired manipulation skill. However, since the human and humanoid robot kinematics and dynamics are not exactly the same, such models usually provide only a rough approximation of the desired skill and need to be refined through practicing, which is a form of autonomous robot learning [1, 12]. Autonomous learning, e.g., reinforcement learning, is an essential component to compute more performant manipulation skills for humanoid robots. Other issues include (1) learning from multiple demonstrations of the desired manipulation skill, where the skill is applied under different environmental conditions; (2) learning of bimanual manipulation skills (humanoid robots have two arms in a known kinematic arrangement); and (3) preserving the stability of a humanoid robot while performing manipulation or any other tasks.

The purpose of this paper is to provide an overview of continuous efforts of our research group in these areas. The progressive evolution of research questions and capabilities, evident in the structure of the paper, illustrates the complexity behind the problem of engineering general-purpose robot manipulation skills. Examples of our work on different platforms are shown in Fig. 1.

Motor Representations and Learning Spaces

An effective movement representation is essential for a successful implementation of robot learning methodologies [16], and even more so for robot manipulation learning. Classical encoding of robotic trajectories encodes motions as a function of time, but representations without explicit time dependency, i.e., autonomous motion representations, are often advantageous when the robot need to react to unexpected events and changes in the environment [17]. Different approaches with and without explicit time dependency have been proposed in the literature. Examples include simple storing of large time-indexed vectors [18], spline fitting and via-points [10], Gaussian Mixture Models [19], function approximators such as different kinds of (deep) neural networks [20–22], nonlinear dynamic systems [23–25], and others. The reduction of the search space brought about by

Fig. 1 Examples of our work on different humanoid robotic platforms. **a** End-to-end vision-to-motion learning [13••]. **b** Learning of bimanual discrete-periodic manipulations on a humanoid robot (© [2015] IEEE. Reprinted, with permission, from [14]). **c** Arm synchronization for bimanual motion and obstacle avoidance and **d** bimanual human-robot collaboration (© [2014] IEEE. Reprinted, with permission, from [15])

a)

b)

c)

d)
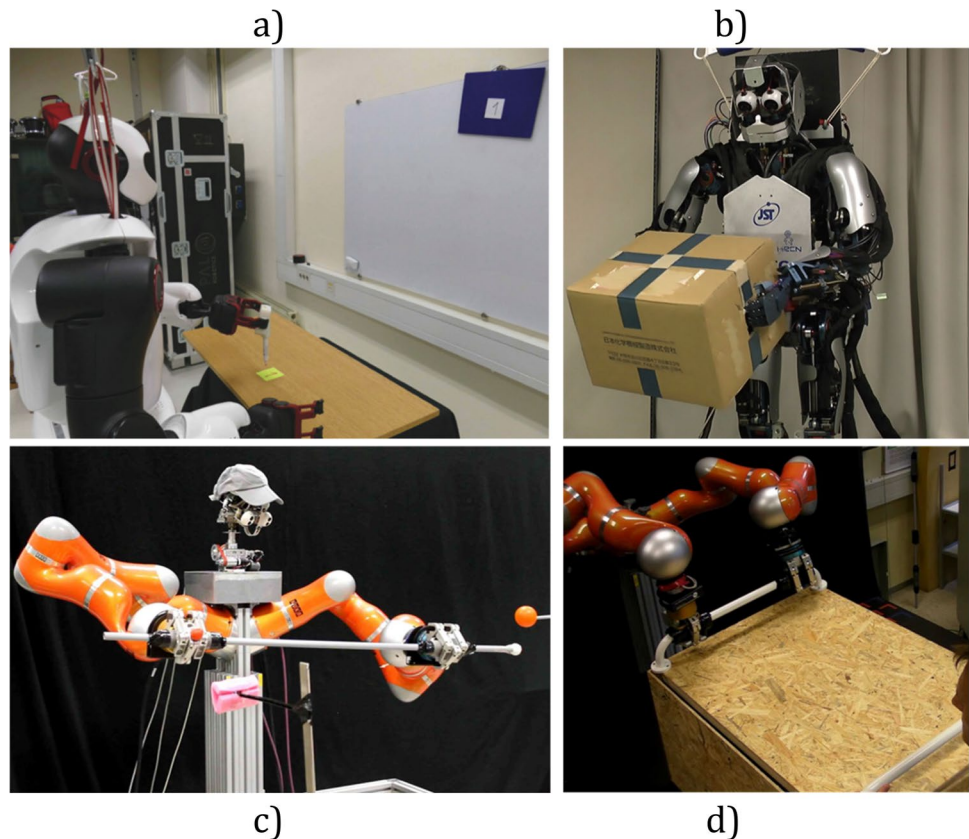
parametric representations is important for the development of effective robot manipulation learning methods [26].

Due to many favorable properties, nonlinear dynamic systems, which form a class of autonomous motion representations, have been widely applied as motion representation for robots acting in dynamic environments. The favorable properties include easy computation of free parameters to encode specific motions, ease of modulation, inclusion of coupling terms for interaction with the environment or other agents, robustness against perturbation, etc.

One of the most widely used nonlinear dynamic systems for movement representation are the Dynamic Movement Primitives (DMPs) [23]. DMPs describe a control policy by a set of nonlinear differential equations with well-defined attractor dynamics for either point-to-point [27], periodic motions [28, 29], and combined discrete-periodic motions [14, 30]. The approach has been expanded over the years to represent orientation trajectories with quaternions [31, 32], enable speed adaptation [33] and other modulation and adaptation features [15], encode variations of movements by adding probabilistic distributions to DMPs [34], and to arc-length dynamic movement primitives (AL-DMPs) [35], where spatial and temporal aspects of motion are well separated. The latter are especially well-suited for action recognition and learning from multiple demonstrations.

The advent of robots with joint torque sensors led to the development of Compliant Movement Primitives (CMPs) [36], which are used to describe both kinematic and dynamic aspects of motion. To specify a CMP, the teacher first demonstrates the desired motion which is recorded by a DMP. Next the robot executes the recorded DMP with high control gains, which ensures accurate motion tracking. The torques arising during this execution are recorded, encoded with radial basis functions, and used to generate feedforward torques during subsequent execution of the desired motion. By providing the feedforward torques, the robot can remain compliant while ensuring accurate motion tracking.

Representation of Bimanual Motion

Unlike standard robot manipulators, humanoid robots can use two arms in a known kinematic arrangement to carry out the desired manipulation tasks. While dual-arm tasks can be described by independently specifying the motion of the two arms arranged in two separate kinematic chains with the common base, it is often better to separate dual-arm motion in absolute and relative coordinates [37]. This way it becomes possible to directly control the relative motion between the two arms, which is often the key to a successful implementation of dual-arm (bimanual) manipulation skills. Just like Cartesian space manipulator arm movements, movements in relative and absolute frames can be specified with standard Cartesian space DMPs [31]. We exploited the

separation of dual-arm humanoid robot motion into relative and absolute coordinates to implement several manipulation skills, e.g., bimanual peg-in-hole [38], bimanual human-robot cooperation for object transportation [39•], and compliant bimanual operations [40].

Humanoid Robot Imitation Learning

Imitation learning, also referred to as learning by demonstration and programming by demonstration (PbD), offers the means to quickly transfer skills from the demonstrator to the learning agent, in this case the robot [41]. This was demonstrated in many robotic applications, which include not only waving in the air, but interaction and manipulation [42–44]. See Fig. 2 for several different examples.

For fixed-base robots, transferring the demonstrated motion in joint- or task-space has been addressed in numerous publications [42]. The demonstrated motion needs to be collected through a proper interface, be it kinesthetically [50–52], visually [53], or through some other sensory system [54], and then transferred to the robot. The data collection does not need to be limited to the kinematic aspects of the task, but can include also the arising forces and torques [55–57]. Depending on the task, the collected motion can be encoded, for example, in DMPs, and then transferred in joint-space or in task-space, for one arm or two, in relative or absolute space [39•], etc. Even for such robots, the difference in embodiment between the demonstrator and the robot might distort the learned motion so that the imitated task is not properly performed. This correspondence problem [58] is much more evident in floating-base, potentially only dynamically stable robots, such as bipedal humanoid robots. Some methods for whole-body motion retargeting proposed in recent years include [59–61].

Besides demonstrator–imitator correspondence, imitation learning cannot be used for direct transfer of motion also because the state of the environment or of the manipulated object(s) is never exactly as in the demonstrated motion. In this sense, learning by demonstration is only useful if it allows for subsequent adaptation of the transferred motion [51]. Thus, the main advantage of imitation learning is in narrowing down the search space for subsequent learning, be it for manipulation learning or any other task. The adaptation is based on acquired sensory information, which can be visual feedback, force feedback, tactile feedback, etc. This was demonstrated for a fixed robot in [62], where a reactive impedance controller was added to the demonstrated trajectory at the acceleration level of the DMP. On a statically stable ARMAR-3 robot, the demonstrated planar wiping motion was adapted in one dimension through an admittance-based force control that combined iterative learning
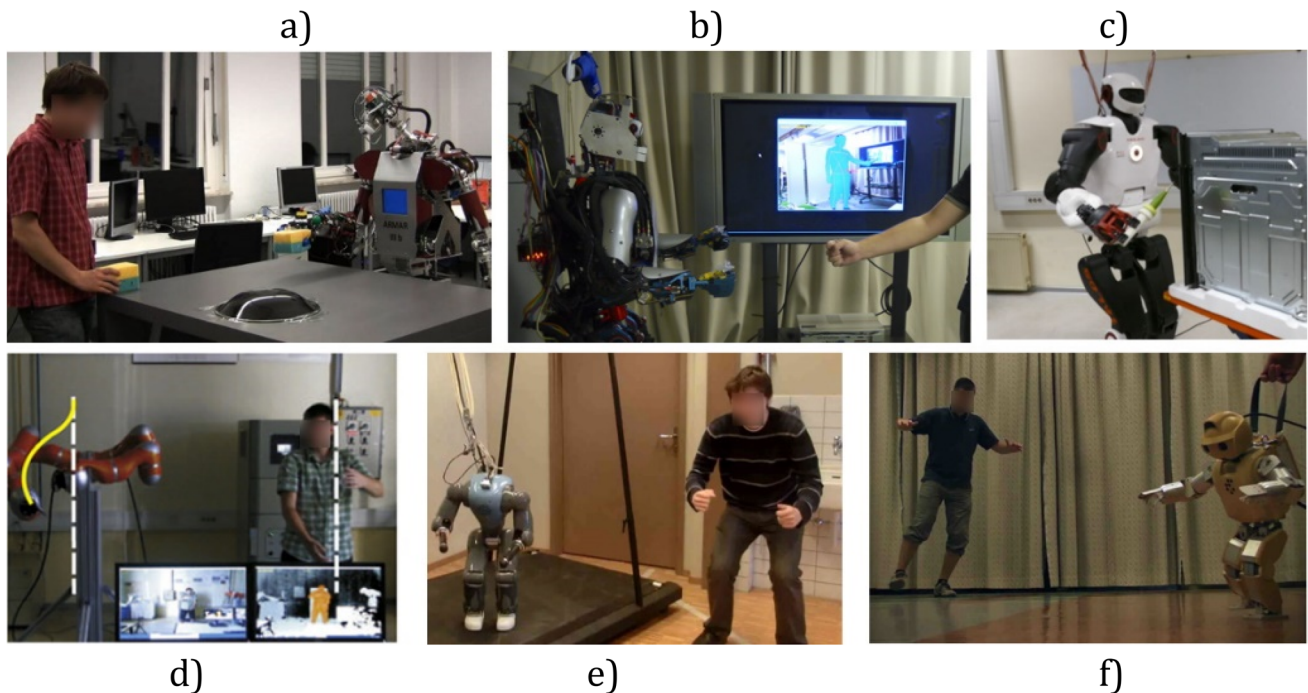
Fig. 2 Examples of adaptation of motion acquired through LbD on different platforms. **a** ARMAR-3 learning to wipe with visual feedback [45]. **b** CbI learning arm gesture motions (reprinted from [46], with permission from Elsevier). **c** TALOS manipulating a measuring tool. **d** Modification of bimanual motion on a bimanual KUKA LWR-4 platform (reprinted from [47], with permission from Elsevier). **e** COMAN humanoid robot full-body motion imitation (reprinted from [48], with permission from Cambridge University Press). **f** HOAP-3 performing full-body motion imitation — walking (© [2013] IEEE. Reprinted, with permission, from [49])

[45, 46]. Both these examples demonstrate that adaptation and additional learning was necessary.

Preserving Postural Stability in Imitation Learning

Let us first address fixed robots and statically stable humanoid robots. An example of the latter is the aforementioned ARMAR-3 humanoid robot, which has a humanoid head and upper body, but a wheeled platform [63]. Here, dynamics of the motion are not a problem, as the stability is ensured by the fixation, or by the platform. Even with statically stable robots, the embodiment might ask for modification of demonstrated motion, for example, to avoid self-collision [47], where the authors implemented an effective methodology that only modifies the motion if necessary, which was implemented through blending of primary and secondary tasks.

On dynamically stable robots, differences in embodiment are even more emphasized, and just transfer of the motion to the robot will result not only in poor execution of the demonstrated desired task, but also in the robot tipping over, at the least. Therefore, adaptation is required in the very core of imitation. In [64], a similar approach as in [47] was used, exploiting the blending of primary and secondary tasks. The demonstrated task was directly transferred in joint-space, unless the projection of the center of mass was approaching

the edge of the stability polygon. Then, the primary task of maintaining stability would take over, and the demonstrated task would only be executed in its null-space. This was also the basis for off-line adaptation of demonstrated motion [49], where the robot would record the demonstrated task and then optimize the whole-body motion in order to maintain stability and approach the likeness of the demonstrated task as much as possible. Similar was also applied on the COMAN humanoid robot [48].

Statistical Learning From Multiple Demonstrations

Learning by demonstration can provide several examples of the desired manipulation action, but it is unlikely that any of the demonstrated actions will be appropriate for the current state of a dynamically changing environment — both in term of the required motion and the associated items involved in the action. However, it is possible to demonstrate more than one task and then use these to generate a new, previously not demonstrated instance of the task. If every trajectory is associated with parameters that describe the characteristics of the task, typically the goal or other conditions of the task [36, 51], then these parameters can serve as query points into the example database.

As explained in [51], the inspiration for such generation comes from motor-tape theories, in which example movement trajectories are stored directly in memory [65, 66]. Generalization from a database of recorded demonstrations was demonstrated on different robotic platforms and tasks. In their seminal work, Ude et al. [51] have shown how generalization from a set of trajectories can be used to generate accurate reaching, grasping, and throwing actions represented by DMPs. The approach combined locally weighted regression [67] and Gaussian process regression [68] to generate all DMP parameters. Later, the complete approach was demonstrated using GPR on reaching with different classes of actions [69]. The approach has been applied to dynamic movement primitives [e.g., 51, 69]. Generalization has been widely adopted in generation of motion also with variations of DMPs and other movement representations [70, 71]. An important alternative is to build variability in the representation itself, such as with nonlinear dynamic systems [24] and probabilistic movement primitives [72]. Figure 3 illustrates different aspects of generalization.

Generalization can also be used to tackle the dynamic aspects of motion. This is necessary if the dynamic models of the robot and the task are not known as it is often the case in imitation learning. Such models cannot be learned by imitation. Thus the challenge is to obtain the correct dynamic models for each task variation. The aforementioned CMPs can be used to describe single instances of tasks. Just as with kinematic data, multiple instances of dynamic data can be used to generate new dynamic motions [36]. In this work, the kinematic and the corresponding dynamic components of CMPs were generalized separately.

Autonomous Augmentation of Trajectory Databases

Generalization can only be accurate if a sufficient amount of training data is made available. If not, the user must demonstrate additional example executions of the desired task. However, this is time consuming and often requires a significant effort from the user. It is therefore advantageous if the robot could augment the available database without
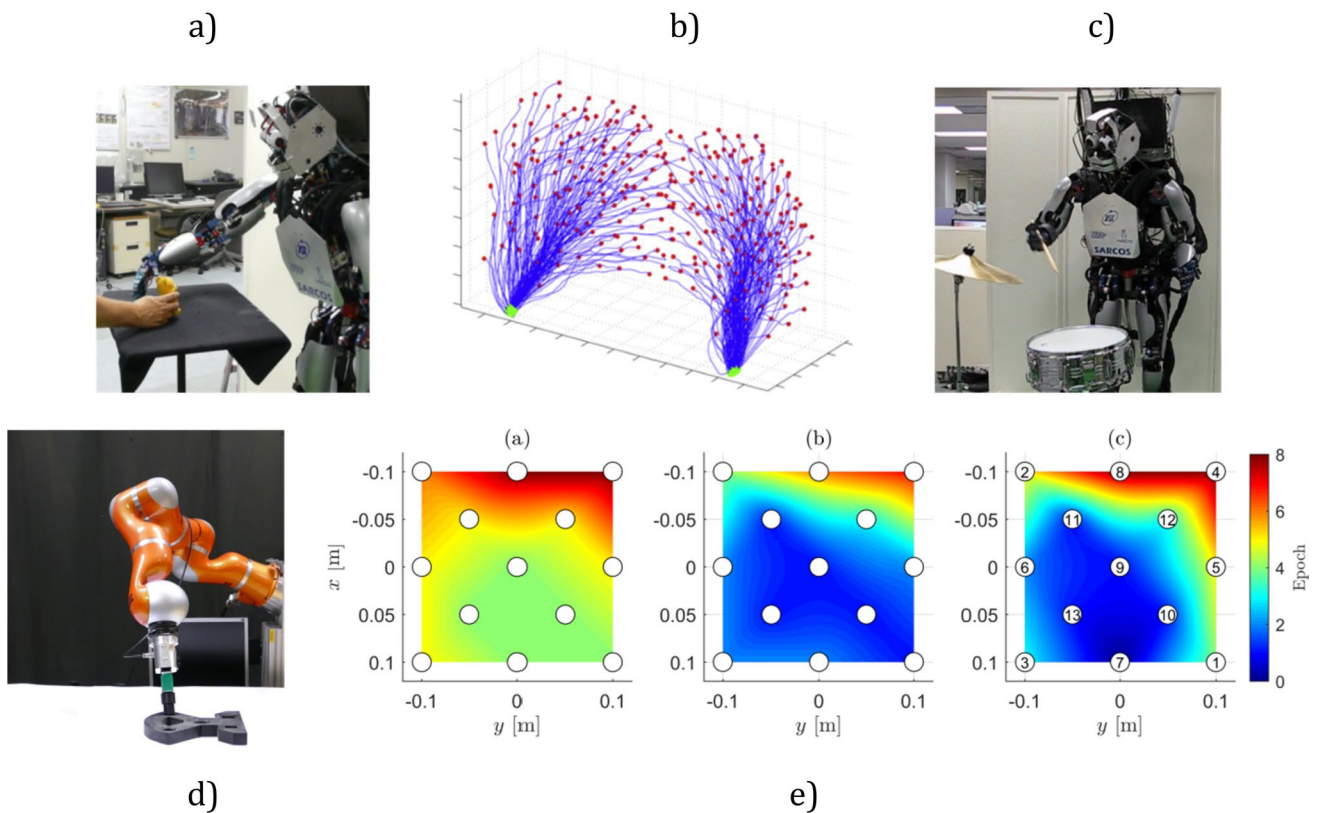

Fig. 3 **a** Generalization for grasping (© [2010] IEEE. Reprinted, with permission, from [51]). **b** Database for generalization of reaching with both arms (reprinted from [69], with permission from Elsevier). **c** Generalization for periodic actions — drumming (© [2010] IEEE. Reprinted, with permission, from [53]). **d** Learning of CMPs and **e** database expansion on the KUKA LWR robot (© [2018] IEEE. Reprinted, with permission, from [73•]). Within **e**: (a) number of learning epochs without database — five on average. (b) Number of learning epochs with leave-one-out cross-validation — two on average. (c) Number of learning epochs through incremental database expansion — two to three on average. The numbers in the circles denote the order of learning and thus the order of database expansion (© [2018] IEEE. Reprinted, with permission, from [73•])

requiring additional user demonstrations. The available database can provide structure to bootstrap the autonomous acquisition of additional task executions and speed up the data gathering process [74].

In our approach, statistical generalization is used to produce good initial approximations for the new variants of the task. If the performed action (represented by DMPs, CMPs, or any other representation) satisfies the given criterion function, e.g., hitting the target for ballistic movements or trajectory tracking accuracy for compliant movements, the new data are immediately added to the database. If not, additional autonomous learning can be applied, starting from the initial movement provided by statistical generalization. Methods such as iterative learning control or reinforcement learning (see the Autonomous Learning and Adaptation of Manipulation Actions" section) can be used for this purpose. A complete system for autonomous extension of the database was proposed in [73•], where the new compliant motion trajectories were generated by statistical generalization. The approach was recently extended also for the periodic repetition of CMPs, which includes the ability for frequency modulation in [75].

Coaching: Learning With Human-in-the-Loop

Besides demonstrating the desired task executions to the robot, a human teacher (coach) can support the learning process by direct interaction with the robot. Such interactions put the human directly in the learning loop. This teaching process is also called coaching [46].

The interactions can directly influence the manipulation policy and can be specified in different ways. Gruebler et al. [76] used voice commands as a reward function in their learning algorithm. Verbal instructions were used to modify movements obtained by human demonstration in [77]. Direct physical contact of the user with the robot is also useful to indicate how the robot should alter its motion. For example, Lee and Ott [52] used kinesthetic teaching with iterative updates to

modify the behavior of a humanoid robot. Coaching based on gestures and obstacle avoidance algorithms was applied to DMPs in [46] and combined with passivity in [78]. Coaching that involved changing of stiffness in path operational space, defined by a Frenet–Serret frame, was proposed in [39•]. Gams et al. [79] evaluated different interfaces and concluded that there is a clear advantage in using force-based coaching methods and that all coaching methods are applicable for rough approximations while accurate tracking is not viable. The coaching process should thus be enhanced by an additional autonomous adaptation method that allows fine tuning of the desired motion.

One of the challenges in autonomous adaptation methods is the design of reward functions, which is a complex problem even for domain experts [80]. Furthermore, acting in the real world and receiving feedback through sensors implies that the true state may not be completely observable and/or noise-free [81]. Besides the robot's on-board sensors, additional external sensors are often applied. Recently it has been shown that learning systems can be effective even if the precise reward function is replaced by natural user feedback. For example, instead of precisely measuring how far the robot has thrown the ball, the user can only specify if the ball has hit the target or the throw was shorter or longer than required. Although such feedback is noisy and not optimal for teaching [82], it was nevertheless applied to successfully learn the ball-in-cup skill [80] and for robotic throwing [83].

Some examples of human-in-the-loop interaction are shown in Fig. 4.

Autonomous Learning and Adaptation of Manipulation Actions

As explained in the previous section, the robot should in most cases autonomously refine the human-demonstrated movements to achieve the required performance of its own
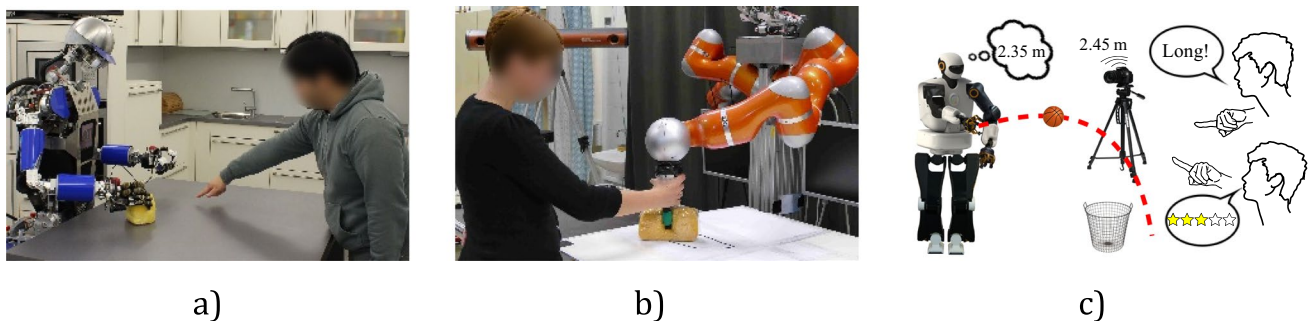

a)   b)   c)

Fig. 4 Three instances of human-in-the-loop intervention. **a** Coaching through gestures (reprinted from [46], with permission from Elsevier). **b** Coaching through physical interaction (© [2016] IEEE. Reprinted, with permission, from [79]). **c** Schematics showing quan-titative sensory, and qualitative human feedback, which acts as reward for reinforcement learning (© [2018] IEEE. Reprinted, with permission, from [83])

task execution. In the absence of the teacher, autonomous robots should also be able to find the appropriate control policies to perform the desired task either by starting from the task performed in a similar situation or even from scratch.

Traditional robot control methods assume that exact a priori models are available. Although remarkable results can be obtained in this way, model-based control can be very sensitive to inaccurately modeled system dynamics [84]. This problem is especially critical for robots operating in human environments, where compliant (low gain) control is usually required to assure the safety of humans, the environment, and the robot itself.

Most of the autonomous learning methods rely on a user-defined cost function. Reinforcement learning [81] is a method of choice for general cost functions that do not provide any additional information besides the evaluation of the motor command executed in a particular state. While general-purpose reinforcement learning can be applied to such cost functions, this type of learning is usually very slow, requiring many repetitions. More effective learning methods can be applied if the cost functions also provide some information how to change the parameters of the desired skill.

Since practicing, i.e., repeating the desired motion with real robots, especially walking robots, is extremely time consuming and also dangerous for the robots, a lot of recent research in robot learning has taken place in simulation [85]. However, many manipulation tasks cannot be learned without interacting with the real world due to the limitations of robot simulation systems. While previous sections explore both aspects of mobile manipulation, in the following sections the focus is on the approaches for manipulation learning on fixed upper-body humanoid robots. This way we avoid the problems with stability while keeping the possibility to experiment with real humanoid robots.

Iterative Learning Control

For many practical problems we can define a cost function that can be evaluated along the motion trajectory of the desired robot motion. Iterative learning control (ILC) [86] can be applied to optimize the robot motion if this cost function allows us to compute how to change the parameters of the manipulation skill; for example, it provides information about the sign of parameter change. The key idea of ILC is to use repetitive system dynamics to compensate for the errors. Although ILC is intrinsically robust to the variation of learning parameters, careful parameter tuning is still required.

We have successfully applied the ILC framework to improve the robot assembly operations acquired by human demonstrations. In automated robot assembly, the unavoidable positioning errors and tight tolerances between the objects involved require compliance and on-line adaptation of the desired trajectories. The resulting forces and torques describe the underlying assembly processes well enough to be taken as a reference for adaptation when transferring the demonstrated assembly policies to new locations [55]. At the new locations, the robot can apply ILC to autonomously improve the demonstrated policies by minimizing the discrepancies between the contact forces and torques arising during the initial human demonstration and the robot execution of the assembly task. Usually only a few epochs of learning are needed to adapt and improve the policy. Another example of the successful application of ILC for policy adaptation and refinement is bi-manual assembly of long poles [38], as shown in Fig. 5a.

Figure 5b shows that ILC framework can be successfully applied also to improve physical human-robot cooperation. In this task, the human and the humanoid robot cooperate to place a table cloth on the table. ILC can be used to transfer the cooperative table cloth placing from one location to another [87]. In this task, the bimanual robot adapts its motion in the absolute coordinates (see the "Representation of Bimanual Motion" section). The results of adaptations are shown in Fig. 5b where only three to four adaptation cycles are needed to reduce the error in relative coordinates substantially.

Hybrid Reinforcement Learning and Adaptive ILC

The application of ILC can be problematic because it is sometimes difficult to set the free parameters (gains) of ILC in such a way that the learning system remains stable. The main problem is that as ILC start converging to the optimal solution, the ILC cost function signal starts oscillating, which prevents the parameters from converging to the optimal solution. This is especially true in cases when the environment dynamics is not known. In order to overcome this problem, various adaptive ILC (AILC) algorithms were proposed in the literature [89, 90] to adapt the gains during learning. Roughly, they can be divided into two sub-classes: (a) ILC with adaptation of the feedback in the current iteration loop and (b) ILC with the adaptation of the learning mechanism (also referred to as adaptation of the previous cycle learning). In order to assure the learning and closed-loop stability of AILC, several issues have to be considered. Unfortunately, some of these issues cannot always be resolved in practice. Consequently, there are only a few examples of successful application of adaptive ILC algorithms in robotic application.
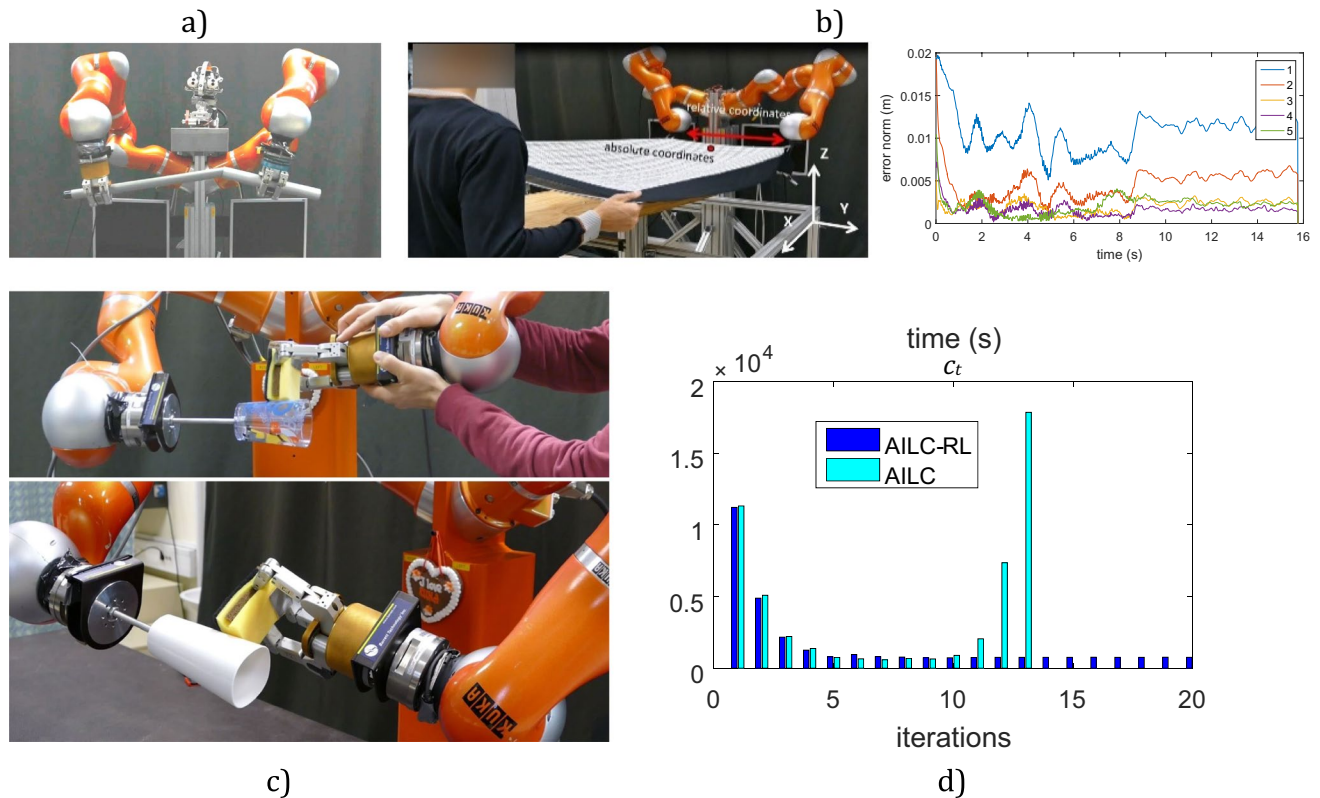
Fig. 5 **a** Humanoid robot torso in bi-manual assembly of long poles interaction (© [2015] IEEE. Reprinted, with permission, from [38]). **b** Humanoid robot during human robot cooperation to place the table cloth. The norm of the relative error in subsequent learning cycles is shown in the graph (reprinted by permission from Springer Nature from [87]). **c** The initial demonstration of wiping policy on a cylindrically shaped glass (top) and the humanoid robot while practicing the glass wiping policy on the oval-shaped glass (bottom) and **d** comparison of cost function evolution for AILC and hybrid AILC-RL scheme in a bar chart (© [2017] IEEE. Reprinted, with permission, from [88])

Reinforcement Learning (RL) enables a robot to autonomously find an optimal policy by direct trial-and-error exploration within its environment [81]. It is often used in robotics to solve problems where models are not available. The main issue with general-purpose reinforcement learning is the high dimensionality of the parameter space arising in motor skill learning. Without any additional information, the robot must estimate the gradient of the cost function to compute the parameter updates, which is an expensive operation.

We have designed a hybrid system that combines the strengths of adaptive ILC and reinforcement learning. In the proposed system, reinforcement learning acts as a supervisor to compute the optimal skill parameters and ILC gains after every learning cycle. Since the general direction of adaptation is provided by ILC, this hybrid system converges much faster than standard reinforcement learning. On the other hand, the reinforcement learning selects the optimal set of parameters from the previous and the current learning cycle, which ensures stable operation of AILC even when the task

dynamics is not known. We used PI^2 reinforcement learning algorithm to implement the proposed hybrid scheme [88].

An example of successful application of hybrid AILC-RL learning framework is bi-manual glass wiping (see Fig. 5c). The initial control policy for cylindrically shaped glasses was obtained by human demonstration using kinesthetic guidance. After that, we replaced the glass with an oval, cone-shaped glass. Instead of demonstrating a new control policy for this glass, we applied AILC-RL to adapt the demonstrated control policy to the new shape. The aim was to achieve the same force-torque profile as applied to the cylindrically shaped glass. During adaptation it was necessary to consider that we handle fragile objects, so the forces and torques during adaptation were limited to small values. AILC-RL ensures convergence even when the feedback gains are low, as it is necessary to prevent high forces and torques from arising. The comparison of the evolution of both cost functions shows that AILC-RL preserves both the stability of reinforcement learning and the adaptation speed of AILC.

Reinforcement Learning in Physically Constrained Environments

Many robotic tasks are performed in contact with an environment that restricts movement to only one degree of freedom. Examples of such tasks are opening and closing doors, drawers, cabinets, sliding doors, latches, etc. Learning such tasks is easier because the space of parameters is one dimensional. However, we do not know the limitations of space in advance. Similar to the previous section, where we used AILC as the search algorithm, this time we use an intelligent compliant controller for this purpose. The underlying controller, which acts as a policy search agent, generates movements along the admissible directions defined by the physical constraints of the task. We employ variable compliance to assure that the robot is stiff in the tangential direction of motion and compliant in the orthogonal directions. This is accomplished by attaching a Frenet-Serret frame (coordinate frame constructed from tangential, normal, and bi-normal vector) to the motion trajectory and defining stiffness along the axes of this frame [39•]. Experimental results show that only a few learning cycles are required for a robot to learn such tasks completely autonomously, without any prior demonstration [91].

Deep Neural Networks for Perception-Action Coupling

The statistical learning approaches described in the "Statistical Learning From Multiple Demonstrations" section require that the programmer defines a query point, which is used to index into the database of example trajectories to compute the appropriate motion for the current state of the external world. The query points usually relate to the desired task, e.g., the desired final pose for reaching movements or the target position for ballistic skills. While this can be highly effective when the goal of the task can be easily described with a few parameters, this is not always possible. In some cases it is better to specify query points by images or even videos.

End-to-End Generation of Manipulation Policies

In our work we focused on how images and image sequences can be transformed into manipulation primitives represented by dynamic movement primitives. Our starting point was the universal approximation theorem for deep neural networks, which indicates that neural networks have a sufficient representational power to learn highly nonlinear mappings that link high-dimensional inputs such as raw images to DMPs. We first tackled the issue of handwriting, that is, how to translate between visual representations of digits perceived from humanoid robot's visual system and the action representations needed to control the humanoid robot's motion trajectories required for handwriting. We addressed this issue by proposing a fully connected image-to-motion encoder-decoder neural network architecture (IMEDNet) [92], which took inspiration from autoencoder neural networks [93]. While the original IMEDNet network was useful for converting images of digits into motion trajectories, certain difficulties became apparent when considering real-world scenarios in which, for example, a robot is shown a digit on a piece of paper held in front of its camera or written in free-form on a whiteboard, and must generate the corresponding handwriting motion. In such cases, the position and orientation of the digit in the acquired camera image is not known, which we overcame by including the spatial transformer into the proposed architecture [94]. Moreover, to reduce the number of parameters in the neural network and take into account the nature of input data, i.e., camera images, we included convolutional layers [95] into the proposed architecture. Finally, to improve the accuracy of the learned neural network models, we developed an optimal criterion function to train the proposed neural network and showed how to compute its gradients for backpropagation [13••]. Its distinguishing feature is that it measures the real distance between handwriting trajectories as opposed to the distance between DMP parameters, which have no physical meaning.

Finally, in many tasks, especially in the context of human-robot interaction, it is insufficient to use single images as input to generate the appropriate robot responses. Full videos should often be used instead, but variable-length videos cannot be processed by feedforward neural networks. We therefore included the LSTM components into the feedforward neural network architectures described above. The resulting recurrent neural network architecture has been shown to be effective for prediction of human intentions [96].

Reduction of Search Space

One of the possible applications of deep neural networks is the reduction of dimensionality [93]. While DMPs provide a relatively low-dimensional representation of the action space, the dimensionality of the DMP parameter space is still rather high [97]. It has been shown that deep autoencoder (AE) neural networks, where the data is pushed through the layer with the smallest number of neurons — the latent space, are superior to standard dimensionality reduction approaches such as Principal Component Analysis (PCA) [93]. In our work we showed that faster convergence of autonomous learning methods, e.g., reinforcement learning, can be achieved when latent space representations computed by deep autoencoder neural networks are used to generate a low-dimensional representation of robot manipulation skill

[98]. In addition, we have demonstrated that generalization methods can be used to generate data for autoencoder neural network training in simulation [99].

Conclusions

Significant progress has been achieved in the area of manipulation learning on humanoid robots over the last 25 years. The main contributions of our group include statistical methods for learning movement primitives from multiple demonstrations, new learning methodologies and representations that combine kinematic and dynamic aspects of manipulation tasks, manipulation learning of bimanual tasks, new algorithms for autonomous learning of manipulation tasks by combining adaptive ILC and reinforcement learning, and the development of new neural network architectures to directly translate sensory signals into manipulation primitives. Together these methods contribute the building blocks to develop behaviors and learning methodologies at the sensorimotor level of the humanoid robot's overall cognitive architecture.

Manipulation learning on humanoid robots remains an open research area, eagerly awaiting progress in humanoid robots' capabilities. On the hardware side, soft robotics and compliant actuator designs can make a significant contribution. Together with new AI approaches based on the availability of vast quantities of data, increased computational power, and deep neural networks, we expect significant progress in the near future. The important problems that remain to be resolved include transferability of results from simulation to the real world and between different robots, which is problematic especially for dynamic tasks.

Declarations

References

Papers of particular interest, published recently, have been highlighted as:
● Of importance
●● Of major importance

1. Kroemer O, Niekum S, Konidaris G. A review of robot learning for manipulation: challenges, representations, and algorithms. J Mach Learn Res. 2021;22(30):1–82.
2. Wonsick M, Long P, Önol AÖ, Wang M, Padir T. A holistic approach to human-supervised humanoid robot operations in extreme environments. Front Robot AI. 2021;8:148. https://doi.org/10.3389/frobt.2021.550644.
3. Kemp CC, Edsinger A, Torres-Jara E. Challenges for robot manipulation in human environments [grand challenges of robotics]. IEEE Robot Autom Mag. 2007;14(1):20–9. https://doi.org/10.1109/MRA.2007.339604.
4. Mason MT. Toward robotic manipulation. Annu Rev Control Robot Autonom Syst. 2018;1(1):1–28. https://doi.org/10.1146/annurev-control-060117-104848.
5. Cui J, Trinkle J. Toward next-generation learned robot manipulation. Sci Robot. 2021;6(54):9461. https://doi.org/10.1126/scirobotics.abd9461.
6. Sun Y, Falco J, Roa M, Calli B. Research challenges and progress in robotic grasping and manipulation competitions. IEEE Robot Autom Lett. 2022;7(2):874–81. https://doi.org/10.1109/LRA.2021.3129134.
7. Stasse O, Flayols T, Budhiraja R, Giraud-Esclasse K, Carpentier J, Mirabel J, Del Prete A, Souères P, Mansard N, Lamiraux F, Laumond J-P, Marchionni L, Tome H, Ferro F. Talos: A new humanoid research platform targeted for industrial applications. In: IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids); 2017. p. 689–95. https://doi.org/10.1109/HUMANOIDS.2017.8246947.
8. Rojas-Quintero JA, Rodríguez-Liñán MC. A literature review of sensor heads for humanoid robots. Robot Auton Syst. 2021;143:103834. https://doi.org/10.1016/j.robot.2021.103834.
9. Pratt J, Ragusila V, Sanchez S, Shrewsbury B, Thornton S, Landry D, Morfey S, Yu E, Cowen S, Farina J, et al. Towards extreme mobility humanoid resupply robots. In: Unmanned Systems Technology XXIII, vol. 11758; 2021. p. 1175804.
10. Schaal S. Is imitation learning the route to humanoid robots? Trends Cogn Sci. 1999;3(6):233–42. https://doi.org/10.1016/S1364-6613(99)01327-3.
11. Ude A, Atkeson CG, Riley M. Programming full-body movements for humanoid robots by observation. Robot Auton Syst. 2004;47:93–108. https://doi.org/10.1016/j.robot.2004.03.004.
12. Ibarz J, Tan J, Finn C, Kalakrishnan M, Pastor P, Levine S. How to train your robot with deep reinforcement learning: lessons we have learned. Int J Robot Res. 2021;40(4-5):698–721. https://doi.org/10.1177/0278364920987859.
13.●● Pahič R, Ridge B, Gams A, Morimoto J, Ude A. Robot skill learning in latent space of a deep autoencoder neural network. Neural Netw. 2020;127:121–31. https://doi.org/10.1016/j.neunet.2020.04.010 **The findings of this paper show how complex sensorimotor skills can be generated using an end-to-end approach that directly maps visual information to robotic skills. The approach is based on deep neural networks with sufficient expressive power to learn such highly nonlinear mappings.**
14. Gams A, Ude A, Morimoto J. Accelerating synchronization of movement primitives: dual-arm discrete-periodic motion of a humanoid robot. In: IEEE/RSJ International Conference

on Intelligent Robots and Systems (IROS); 2015. p. 2754–60. https://doi.org/10.1109/IROS.2015.7353755.

15. Gams A, Nemec B, Ijspeert AJ, Ude A. Coupling movement primitives: interaction with the environment and bimanual tasks. IEEE Trans Robot. 2014;30(4):816–30. https://doi.org/10.1109/TRO.2014.2304775.

16. Calinon S, Guenter F, Billard A. On learning, representing, and generalizing a task in a humanoid robot. IEEE Trans Syst Man Cybern B (Cybernetics). 2007;37(2):286–98. https://doi.org/10.1109/TSMCB.2006.886952.

17. Schaal S, Mohajerian P, Ijspeert AJ. Dynamics systems vs. optimal control — a unifying view. Prog Brain Res. 2007;165:425–45. https://doi.org/10.1016/S0079-6123(06)65027-9.

18. Kawamura S, Fukao N. A time-scale interpolation for input torque patterns obtained through learning control on constrained robot motions. In: IEEE International Conference on Robotics and Automation (ICRA); 1995. p. 2156–61. https://doi.org/10.1109/ROBOT.1995.525579.

19. Calinon S. Mixture models for the analysis, edition, and synthesis of continuous time series. In: Bouguila N, Fan W, editors. Mixture models and applications. Cham: Springer; 2019. p. 39–57. https://doi.org/10.1007/978-3-030-23876-63.

20. Levine S, Finn C, Darrell T, Abbeel P. End-to-end training of deep visuomotor policies. J Mach Learn Res. 2016;17(1):1334–73.

21. Bousmalis K, Irpan A, Wohlhart P, Bai Y, Kelcey M, Kalakrishnan M, Downs L, Ibarz J, Pastor P, Konolige K, Levine S, Vanhoucke V. Using simulation and domain adaptation to improve efficiency of deep robotic grasping. In: 2018 IEEE International Conference on Robotics and Automation (ICRA); 2018. p. 4243–50. https://doi.org/10.1109/ICRA.2018.8460875.

22. Nguyen H, La H. Review of deep reinforcement learning for robot manipulation. In: Third IEEE International Conference on Robotic Computing (IRC); 2019. p. 590–5. https://doi.org/10.1109/IRC.2019.00120.

23. Ijspeert AJ, Nakanishi J, Hoffmann H, Pastor P, Schaal S. Dynamical movement primitives: learning attractor models for motor behaviors. Neural Comput. 2013;25(2):328–73. https://doi.org/10.1162/NECO_a_00393.

24. Khansari-Zadeh SM, Billard A. Learning stable nonlinear dynamical systems with Gaussian mixture models. IEEE Trans Robot. 2011;27(5):943–57. https://doi.org/10.1109/TRO.2011.2159412.

25. Salehian SSM, Figueroa N, Billard A. A unified framework for coordinated multi-arm motion planning. Int J Robot Res. 2018;37(10):1205–32. https://doi.org/10.1177/0278364918765952.

26. Stulp F, Sigaud O. Robot skill learning: from reinforcement learning to evolution strategies. Paladyn J Behav Robot. 2013;4(1):49–61. https://doi.org/10.2478/pjbr-2013-0003.

27. Ijspeert AJ, Nakanishi J, Schaal S. Movement imitation with nonlinear dynamical systems in humanoid robots. In: IEEE International Conference on Robotics and Automation (ICRA). Washington, DC; 2002. p. 1398–403. https://doi.org/10.1109/ROBOT.2002.1014739.

28. Ijspeert AJ, Nakanishi J, Schaal S. Learning rhythmic movements by demonstration using nonlinear oscillators. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Lausanne, Switzerland; 2002. p. 958–63. https://doi.org/10.1109/IRDS.2002.1041514.

29. Gams A, Ijspeert AJ, Schaal S, Lenarčič J. On-line learning and modulation of periodic movements with nonlinear dynamical systems. Auton Robot. 2009;27:3–23. https://doi.org/10.1007/s10514-009-9118-y.

30. Ernesti J, Righetti L, Do M, Asfour T, Schaal S. Encoding of periodic and their transient motions by a single dynamic movement primitive. In: 2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012); 2012. p. 57–64. https://doi.org/10.1109/HUMANOIDS.2012.6651499.

31. Ude A, Nemec B, Petrič T, Morimoto J. Orientation in Cartesian space dynamic movement primitives. In: IEEE International Conference on Robotics and Automation (ICRA). Hong Kong; 2014. p. 2997–3004. https://doi.org/10.1109/ICRA.2014.6907291.

32. Koutras L, Doulgeri Z. Exponential stability of trajectory tracking control in the orientation space utilizing unit quaternions. In: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2021. p. 8151–8. https://doi.org/10.1109/IROS51168.2021.9636171.

33. Vuga R, Nemec B, Ude A. Speed adaptation for self-improvement of skills learned from user demonstrations. Robotica. 2016;34(12):2806–22. https://doi.org/10.1017/S0263574715000405.

34. Zhou Y, Gao J, Asfour T. Movement primitive learning and generalization: using mixture density networks. IEEE Robot Autom Mag. 2020;27(2):2–12. https://doi.org/10.1109/MRA.2020.2980591.

35. Gašpar T, Nemec B, Morimoto J, Ude A. Skill learning and action recognition by arc-length dynamic movement primitives. Robot Auton Syst. 2018;100:225–35. https://doi.org/10.1016/j.robot.2017.11.012.

36. Deniša M, Gams A, Ude A, Petrič T. Learning compliant movement primitives through demonstration and statistical generalization. IEEE/ASME Trans Mechatron. 2016;21(5):2581–94. https://doi.org/10.1109/TMECH.2015.2510165.

37. Chiacchio P, Chiaverini S, Siciliano B. Direct and inverse kinematics for coordinated motion tasks of a two-manipulator system. J Dyn Syst Meas Control. 1996;118(4):691–7. https://doi.org/10.1115/1.2802344.

38. Likar N, Nemec B, Žlajpah L, Ando S, Ude A. Adaptation of bimanual assembly tasks using iterative learning framework. In: IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids); 2015. p. 771–6. https://doi.org/10.1109/HUMANOIDS.2015.7363457.

39.● Nemec B, Likar N, Gams A, Ude A. Human robot cooperation with compliance adaptation along the motion trajectory. Auton Robot. 2018;42(5):1023–35. https://doi.org/10.1007/s10514-017-9676-3 **This paper describes a new methodology for intuitive and natural physical human-robot interaction in cooperative tasks. From an initial behavior obtained by imitation learning, the robot behavior naturally evolves into a cooperative task.**

40. Batinica A, Nemec B, Ude A, Raković M, Gams A. Compliant movement primitives in a bimanual setting. In: IEEE-RAS 17th International Conference on Humanoid Robots (Humanoids); 2017. p. 365–71. https://doi.org/10.1109/HUMANOIDS.2017.8246899.

41. Billard AG, Calinon S, Dillmann R. In: Siciliano B, Khatib O, editors. Learning from humans. Cham: Springer; 2016. p. 1995–2014. 74. https://doi.org/10.1007/978-3-319-32552-1.

42. Argall BD, Chernova S, Veloso M, Browning B. A survey of robot learning from demonstration. Robot Auton Syst. 2009;57(5):469–83. https://doi.org/10.1016/j.robot.2008.10.024.

43. Zhu Z, Hu H. Robot learning from demonstration in robotic assembly: a survey. Robotics. 2018;7(2). https://doi.org/10.3390/robotics7020017.

44. Fang B, Jia S, Guo D, Xu M, Wen S, Sun F. Survey of imitation learning for robotic manipulation. Int J Intell Robot Appl. 2019;3(4):362–9.

45. Gams A, Do M, Ude A, Asfour T, Dillmann R. On-line periodic movement and force-profile learning for adaptation to new surfaces. In: 10th IEEE-RAS International Conference on Humanoid Robots; 2010. p. 560–5. https://doi.org/10.1109/ICHR.2010.5686306.

46. Gams A, Petrič T, Do M, Nemec B, Morimoto J, Asfour T, Ude A. Adaptation and coaching of periodic motion primitives through physical and visual interaction. Robot Auton Syst. 2016;75:340–51. https://doi.org/10.1016/j.robot.2015.09.011.

47. Petrič T, Žlajpah L. Smooth continuous transition between tasks on a kinematic control level: obstacle avoidance as a control problem. Robot Auton Syst. 2013;61(9):948–59. https://doi.org/10.1016/j.robot.2013.04.019.

48. Gams A, van den Kieboom J, Dzeladini F, Ude A, Ijspeert AJ. Real-time full body motion imitation on the coman humanoid robot. Robotica. 2015;33(5):1049–61. https://doi.org/10.1017/S0263574714001477.

49. Vuga R, Ogrinc M, Gams A, Petrič T, Sugimoto N, Ude A, Morimoto J. Motion capture and reinforcement learning of dynamically stable humanoid movement primitives. In: IEEE International Conference on Robotics and Automation (ICRA), Karlsruhe, Germany; 2013. p. 5284–90. https://doi.org/10.1109/ICRA.2013.6631333.

50. Hersch M, Guenter F, Calinon S, Billard A. Dynamical system modulation for robot learning via kinesthetic demonstrations. IEEE Trans Robot. 2008;24(6):1463–7. https://doi.org/10.1109/TRO.2008.2006703.

51. Ude A, Gams A, Asfour T, Morimoto J. Task-specific generalization of discrete and periodic dynamic movement primitives. IEEE Trans Robot. 2010;26(5):800–15. https://doi.org/10.1109/TRO.2010.2065430.

52. Lee D, Ott C. Incremental kinesthetic teaching of motion primitives using the motion refinement tube. Auton Robot. 2011;31:115–31. https://doi.org/10.1007/s10514-011-9234-3.

53. Kulić D, Ott C, Lee D, Ishikawa J, Nakamura Y. Incremental learning of full body motion primitives and their sequencing through human motion observation. Int J Robot Res. 2012;31(3):330–45. https://doi.org/10.1177/0278364911426178.

54. Ott C, Lee D, Nakamura Y. Motion capture based human motion recognition and imitation by direct marker control. In: 8th IEEE-RAS International Conference on Humanoid Robots (Humanoids); 2008. p. 399–405. https://doi.org/10.1109/ICHR.2008.4755984.

55. Abu-Dakka FJ, Nemec B, Jorgensen JA, Savarimuthu TR, Krüger N, Ude A. Adaptation of manipulation skills in physical contact with the environment to reference force profiles. Auton Robot. 2015;39(2):199–217. https://doi.org/10.1007/s10514-015-9435-2.

56. Koropouli V, Hirche S, Lee D. Generalization of force control policies from demonstrations for constrained robotic motion tasks. J Intell Robot Syst. 2015;80(1):133–48. https://doi.org/10.1007/s10846-015-0218-y.

57. Kramberger A, Gams A, Nemec B, Chrysostomou D, Madsen O, Ude A. Generalization of orientation trajectories and force-torque profiles for robotic assembly. Robot Auton Syst. 2017;98:333–46. https://doi.org/10.1016/j.robot.2017.09.019.

58. Alissandrakis A, Nehaniv CL, Dautenhahn K. Correspondence mapping induced state and action metrics for robotic imitation. IEEE Trans Syst Man Cybern B (Cybernetics). 2007;37(2):299–307. https://doi.org/10.1109/TSMCB.2006.886947.

59. Penco L, Clement B, Tsagarakis NG, Mouret J-B, Ivaldi S, et al. Robust real-time whole-body motion retargeting from human to humanoid, IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids); 2018. p. 425–32. https://doi.org/10.1109/HUMANOIDS.2018.8624943.

60. Rouxel Q, Yuan K, Wen R, Li Z. Multicontact motion retargeting using whole-body optimization of full kinematics and sequential force equilibrium. IEEE/ASME Trans Mechatron. 2022. https://doi.org/10.1109/TMECH.2022.3152844.

61. Lannan N, Zhou L, Fan G. Human motion enhancement via Tobit Kalman filter-assisted autoencoder. IEEE Access. 2022;10:29233–51. https://doi.org/10.1109/ACCESS.2022.3157605.

62. Pastor P, Righetti L, Kalakrishnan M, Schaal S. Online movement adaptation based on previous sensor experiences. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2011. p. 365–71. https://doi.org/10.1109/IROS.2011.6095059.

63. Asfour T, Regenstein K, Azad P, Schroder J, Bierbaum A, Vahrenkamp N, Dillmann R. Armar-III: An integrated humanoid platform for sensory-motor control. In: 6th IEEE-RAS International Conference on Humanoid Robots (Humanoids); 2006. p. 169–75. https://doi.org/10.1109/ICHR.2006.321380.

64. Petrič T, Gams A, Babič J, Žlajpah L. Reflexive stability control framework for humanoid robots. Auton Robot. 2013;34(4):347–61. https://doi.org/10.1007/s10514-013-9329-0.

65. Poggio T, Bizzi E. Generalization in vision and motor control. Nature. 2004;431(7010):768–74. https://doi.org/10.1038/nature03014.

66. Atkeson CG, Hale JG, Pollick F, Riley M, Kotosaka S, Schaul S, Shibata T, Tevatia G, Ude A, Vijayakumar S, Kawato E, Kawato M. Using humanoid robots to study human behavior. IEEE Intell Syst Appl. 2000;15(4):46–56. https://doi.org/10.1109/5254.867912.

67. Atkeson CG, Moore AW, Schaal S. Locally weighted learning. Artif Intell Rev. 1997;11(1-5):11–73. https://doi.org/10.1023/A:1006559212014.

68. Rasmussen CE, Williams CKI. Gaussian processes for machine learning. Cambridge: The MIT Press; 2005.

69. Forte D, Gams A, Morimoto J, Ude A. On-line motion synthesis and adaptation using a trajectory database. Robot Auton Syst. 2012;60(10):1327–39. https://doi.org/10.1016/j.robot.2012.05.004.

70. Matsubara T, Hyon S-H, Morimoto J. Learning parametric dynamic movement primitives from multiple demonstrations. Neural Netw. 2011;24(5):493–500. https://doi.org/10.1016/j.neunet.2011.02.004.

71. Calinon S. A tutorial on task-parameterized movement learning and retrieval. Intell Serv Robot. 2016;9(1):1–29. https://doi.org/10.1007/s11370-015-0187-9.

72. Paraschos A, Daniel C, Peters J, Neumann G. Using probabilistic movement primitives in robotics. Auton Robot. 2018;42:529–51. https://doi.org/10.1007/s10514-017-9648-7.

73.● Petrič T, Gams A, Colasanto L, Ijspeert AJ, Ude A. Accelerated sensorimotor learning of compliant movement primitives. IEEE Trans Robot. 2018:1–7. https://doi.org/10.1109/TRO.2018.2861921 **This paper explains how to exploit existing sensorimotor knowledge to autonomously improve and expand the robot's skill knowledge. It combines various aspects of learning, imitation, and generalization, finally demonstrating the applicability of the proposed methodology for effective learning of dynamic skills.**

74. Wörgötter F, Geib C, Tamosiunaite M, Aksoy EE, Piater J, Xiong H, Ude A, Nemec B, Kraft D, Krüger N, Wächter M, Asfour T. Structural bootstrapping — a novel, generative mechanism for faster and more efficient acquisition of action-knowledge. 2015;7(2):140–54. https://doi.org/10.1109/TAMD.2015.2427233.

75. Petrič T. Phase-synchronized learning of periodic compliant movement primitives (P-CMPs). Front Neurorobot. 2020;14:1–12. https://doi.org/10.3389/fnbot.2020.599889.

76. Gruebler A, Berenz V, Suzuki K (2011) Coaching robot behavior using continuous physiological affective feedback. In: 2011 11th IEEE-RAS International Conference on Humanoid Robots, pp. 466–471. https://doi.org/10.1109/Humanoids.2011.6100888

77. Riley M, Ude A, Atkeson C, Cheng G. Coaching: an approach to efficiently and intuitively create humanoid robot behaviors. In: 6th IEEE-RAS International Conference on Humanoid Robots (Humanoids); 2006. p. 567–74. https://doi.org/10.1109/ICHR.2006.321330.

78. Papageorgiou D, Kastritsi T, Doulgeri Z. A passive robot controller aiding human coaching for kinematic behavior modifications. Robot Comput Integr Manuf. 2020;61:101824. https://doi.org/10.1016/j.rcim.2019.101824.

79. Gams A, Ude A. On-line coaching of robots through visual and physical interaction: analysis of effectiveness of human-robot interaction strategies. In: 2016 IEEE International Conference on Robotics and Automation (ICRA); 2016. p. 3028–34. https://doi.org/10.1109/ICRA.2016.7487467.

80. Vollmer A-L, Hemion NJ. A user study on robot skill learning without a cost function: optimization of dynamic movement primitives via naive user feedback. Front Robot AI. 2018;5:77. https://doi.org/10.3389/frobt.2018.00077.

81. Kober J, Bagnell JA, Peters J. Reinforcement learning in robotics: a survey. Int J Robot Res. 2013;32(11):1238–74. https://doi.org/10.1177/0278364913495721.

82. Weng P, Busa-Fekete R, Hüllermeier E. Interactive Q-learning with ordinal rewards and unreliable tutor. In: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD); 2013.

83. Pahič R, Lončarević Z, Ude A, Nemec B, Gams A. User feedback in latent space robotic skill learning. In: IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids); 2018. p. 270–6. https://doi.org/10.1109/HUMANOIDS.2018.8624972.

84. Peters J, Mülling K, Kober J, Nguyen-Tuong D, Krömer O. In: Pradalier C, Siegwart R, Hirzinger G, editors. Towards motor skill learning for robotics. Berlin: Springer; 2011. p. 469–82. https://doi.org/10.1007/s10339-011-0404-1.

85. Hwangbo J, Lee J, Dosovitskiy A, Bellicoso D, Tsounis V, Koltun V, Hutter M. Learning agile and dynamic motor skills for legged robots. Science. Robotics. 2019;4(26):eaau5872. https://doi.org/10.1126/scirobotics.aau5872.

86. Norrlöf M, Gunnarsson S. Experimental comparison of some classical iterative learning control algorithms. IEEE Trans Robot Autom. 2002;18(4):636–41. https://doi.org/10.1109/TRA.2002.802210.

87. Nemec B, Likar N, Gams A, Ude A. Adaptive human robot cooperation scheme for bimanual robots. In: Lenarcic J, Merlet JP, editors. Advances in Robot Kinematics. Cham: Springer; 2018. p. 385–93. https://doi.org/10.1007/978-3-319-56802-7_39.

88. Nemec B, Simonič M, Likar N, Ude A. Enhancing the performance of adaptive iterative learning control with reinforcement learning. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2017. p. 2192–9. https://doi.org/10.1109/IROS.2017.8206038.

89. Tayebi A. Adaptive iterative learning control for robot manipulators. Automatica. 2004;40(7):1195–203. https://doi.org/10.1109/ACC.2003.1240553.

90. Ouyang PR, Petz BA, Xi FF. Iterative learning control with switching gain feedback for nonlinear systems. J Comput Nonlinear Dyn. 2011;6(1). https://doi.org/10.1109/TIC-STH.2009.5444376.

91. Nemec B, Žlajpah L, Ude A. Door opening by joining reinforcement learning and intelligent control. In: 2017 18th International Conference on Advanced Robotics (ICAR); 2017. p. 222–8. https://doi.org/10.1109/ICAR.2017.8023522.

92. Pahič R, Gams A, Ude A, Morimoto J. Deep encoder-decoder networks for mapping raw images to dynamic movement primitives. In: IEEE International Conference on Robotics and Automation (ICRA). Brisbane, Australia; 2018. p. 5863–8. https://doi.org/10.1109/ICRA.2018.8460954.

93. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science. 2006;313(5786):504–7. https://doi.org/10.1126/science.1127647.

94. Ridge B, Pahič R, Ude A, Morimoto J. Learning to write anywhere with spatial transformer image-to-motion encoder-decoder networks. In: IEEE International Conference on Robotics and Automation (ICRA). Montreal, Canada; 2019. p. 2111–7. https://doi.org/10.1109/ICRA.2019.8794253.

95. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to handwritten zip code recognition. Neural Comput. 1989;1(4):541–51. https://doi.org/10.1162/neco.1989.1.4.541.

96. Mavsar M, Deniša M, Nemec B, Ude A. Intention recognition with recurrent neural networks for dynamic human-robot collaboration. In: International Conference on Advanced Robotics (ICAR); 2021. p. 208–15. https://doi.org/10.1109/ICAR53236.2021.9659473.

97. Nemec B, Vuga R, Ude A. Efficient sensorimotor learning from multiple demonstrations. Adv Robot. 2013;27(13):1023–31. https://doi.org/10.1080/01691864.2013.814211.

98. Pahič R, Lončarević Z, Gams A, Ude A. Robot skill learning in latent space of a deep autoencoder neural network. Robot Auton Syst. 2021;135:103690. https://doi.org/10.1016/j.robot.2020.103690.

99. Lončarević Z, Pahič R, Ude A, Gams A. Generalization-based acquisition of training data for motor primitive learning by neural networks. Appl Sci. 2021;11:1013. https://doi.org/10.3390/app11031013.