Robotics and Autonomous Systems 128 (2020) 103515

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Skill transfer learning for autonomous robots and human–robot cooperation: A survey☆

Yueyue Liu [a], Zhijun Li [b,*], Huaping Liu [c], Zhen Kan [b]

[a] College of Automation Science and Engineering, South China University of Technology, Guangzhou, 510640, China
[b] Department of Automation, University of Science and Technology of China, Hefei, 230027, China
[c] Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Article history:
Available online 3 April 2020

Keywords:
Skill transfer
Robot learning
Human demonstrations
Neurophysiological skill acquisition

ABSTRACT

Designing a robot system with reasoning and learning ability has gradually become a research focus in robotics research field. Recently, *Skill Transfer Learning* (STL), i.e., the ability of transferring human skills to robots, has become a research thrust for autonomous robots and human–robot cooperation. It provides the following benefits: (i) the skill transfer learning system with independent decision-making and learning ability enables the robot to learn and acquire manipulation skills in a complex and dynamic environment, which can overcome the shortages of conventional methods such as traditional programming, and greatly improve the adaptability of the robot to complex environments and (ii) human physiological signals allow us to extract motion control characteristics from physiological levels which create a rich sensory signal. In this survey, we provide an overview of the most important applications of STL by analyzing and categorizing existing works in autonomous robots and human–robot cooperation area. We close this survey by discussing remaining open challenges and promising research topics in future.

1. Introduction

Autonomous robots and human–robot cooperation are becoming more personalized, interactive, and engaging than ever, providing assistance ranging from daily life to manufacturing, healthcare, and transportation [1]. In these applications, robots are desired to have the capabilities of handling tasks autonomously in different environments and interact with human safely. A key enabler to these applications is to design a system with reasoning and learning ability.

Early robotic manipulation or motion behaviors are usually composed of a series of prescribed motion sequences, which cannot adapt to changing and complex environments [2]. However, for complex tasks, pre-programming a robot is not only inefficiency and tedious, but also impracticable, especially if tasks are constantly updated or changed. In addition, traditional programming approaches cannot achieve autonomous behaviors due to the overlook of human actions and external environments. For example, in a household service scenario, it is hard to pre-program the robot's tasks taking into account all potential human behaviors and indoor environment configurations. To date, few existing robots can easily perform tie shoelaces, cook, or cut hair.

For human grasping and lifting tasks, neurophysiological research reveals that human relies on the detection of discrete mechanical events that occur when grasping, lifting and replacing an object. Such events represent transitions between phases of the evolving manipulation task (e.g., object contact, lift-off, etc.), and provide critical elements required for the sequential control of the task as well as for corrections and parameterization of the task [3]. Even a simple power grasp manipulation task would engage large parts of the human brain [4], requiring sophisticated control processing. Coordinated and graceful lifting patterns observed in adults are not realized until humans are 8–10 years old [5]. Therefore, humans need nearly a decade of daily practice to master this seemingly simple sensorimotor task.

Consequently, robots are desirable to have the abilities of perception, decision-making and learning in a complex and dynamic environment [6]. It is well known that human manipulation behavior essentially relies on the constant exploration and understanding of the relationship between actions and sensory responses. Human usually preserves the skill knowledge learned in the past and utilizes it to help future learning and problem solving. Inspired by human learning and skill transfer

☆ This work was supported in part by the National Natural Science Foundation of China under Grant 61625303, and Grant 61751310, in part by the National Key Research and Development Program of China under Grant 2017YFB1302302, Grant 2018YFC2001600, in part by the Anhui Science and Technology Major Program under Grant 17030901029.

* Corresponding author.
E-mail address: zjli@ieee.org (Z. Li).

https://doi.org/10.1016/j.robot.2020.103515
0921-8890/© 2020 Elsevier B.V. All rights reserved.

Nomenclature

BCIs	Brain–computer interfaces
BMIs	brain–machine interfaces
CMPs	compliant movement primitives
CNN	convolutional neural network
DDPG	Deep deterministic policy gradient
DMPs	dynamic movement primitives
DPG	Deterministic policy gradient
DQN	Deep Q-network
DRL	deep reinforcement learning
EEG	electroencephalogram
GMMs	Gaussian mixture models
GPIRL	Gaussian process inverse reinforcement learning
GPS	Guided policy search
GQ-CNNs	Grasp Quality Convolutional Neural Networks
HMMs	hidden Markov models
IRL	inverse reinforcement learning
POMDP	partially observable Markov decision process
ProMPs	probabilistic movement primitives
RL	reinforcement learning
sEMG	surface electromyography
SNR	signal-to-noise
SOSC	scalable online sequence clustering
STL	skill transfer learning
TMR	targeted muscle reinnervation
TRPO	Trust region policy optimization

process, developing robot functionality with human level capability of perception, planning and control, has always been an essential goal. Similar to human behaviors, robots typically need to physically interact with environments or humans while performing tasks with rich and informative neurophysiological sensory signals, which are all occurring simultaneously with actions. Moreover, there is a relationship between the sensory responses and actions which could be explored to predict and interpret these behaviors. For autonomous robots and human–robot cooperation, skill transfer learning enables robots to retain or utilize the behaviors observed from human as their skills, improve them by practice, and then apply them into new task environments. The main idea of STL is to develop technical solutions by imitating and exploiting the natural models, systems, and processes.

Motivated by this idea, robots gradually gained the ability to automatically generate motion sequences to perform desired tasks according to the characteristics of the environment and the object, e.g., size and weight [7]. In addition, human neurophysiological signals have been adopted to restore human manipulation functionality by using human muscle activity or cerebral cortex to control the movements of different autonomous devices and perform different human-cooperation tasks [8–10].

Based on STL, autonomous robots and human–robot cooperation have been a key research area in advanced robotics. The STL methods have been widely applied in perception, learning and control, which integrate knowledge from neurophysiological signals [11], cognitive and executive processing [12]. These new skill acquisition mechanisms significantly facilitate the development of robotic systems with desired properties inspired from neurophysiological and human skill learning processes, such as adaptivity, robustness, versatility, and agility [13]. Because of the advantages listed above, STL has attracted great research attention, and becomes a vital tool enabling robots to deal with environment uncertainties. However, the acquisition of human autonomous skills is quite challenging.

There exist excellent surveys in the literature regarding the skill acquisition process in robot area [14–19]. However, most of them either focus on the robot learning (e.g, machine learning in robotics, RL in robot control) or pure affordances in psychology and neuroscience. Few of them reviews the topic of STL, especially for the skill acquisition via neurophysiological signals. In this paper, we summarize the status and challenges of STL systems. The remainder of this paper is organized as follows. In Section 2, we present the state of the art of STL which includes the categorizations, framework and application of STL. Section 3 introduces the robot learning of STL. Section 4 reviews the recent developments of STL via neurophysiological signals. Section 5 discusses future directions for STL in autonomous robots and human–robot cooperation.

2. Skill transfer learning

Autonomous robots and human–robot cooperation are desirable to handle objects of different size and weight in a dynamic environment. Skill learning cannot only address challenges caused by the lack of accurate object model and interaction dynamic model, but also the increasing complexity of perception and control of systems with large degrees of freedom. In addition, with the development of recent neuroscience technologies, precise nature of human neural representations can be utilized to transfer human skills to robots.

By exploiting robot learning and human–robot skills, STL can enhance robot skill acquisition and enable robust perception-guided manipulation behaviors, utilizing a large number of machine learning approaches in robotics. Fig. 1 shows the categorization of STL. Generally, STL has two benefits. First, the STL system with independent decision-making and learning ability enables the robot to learn and acquire manipulation skills in a complex and dynamic environment, which can overcome the shortages of conventional methods such as traditional programming, and greatly improve the adaptability of the robot to the environment. Second, human physiological signals allow us to extract motion control characteristics from physiological levels to create rich sensory signal. In this section, we will introduce the framework and interfaces of STL, and elaborate the benefits of STL in more detail.

2.1. Skill transfer learning framework

A robot system usually consists of a body part, and perception and control modules. The robot body interacts directly with the physical environment via actuators, sensors, and the guidance of human intension. As shown in Fig. 2, STL can be divided into two parts: robot autonomous learning and human intention acquisition transfer. Skill model is the core of STL which can be characterized by the type of feedback and the process of data generation. Skill reproduction is the other important component of STL, which produces the corresponding actions according to the learned skill model.

As for the learning process, demonstrations of the task are collected and then used to extract the desired motion information which include position, velocity, force etc. In addition, the stiffness can be collected for some collaboration tasks [20,21]. After demonstration, the data sets of skills information are acquired. Specially, the motion representing, demonstration alignment, motion segmentation and generation should be considered [22]. Note that human intention acquisition transfer only

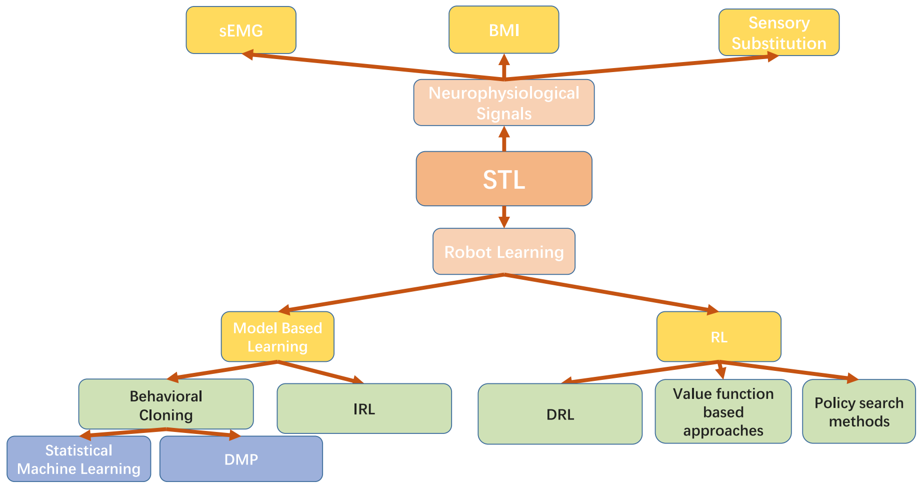


Fig. 1. Categorization of STL.

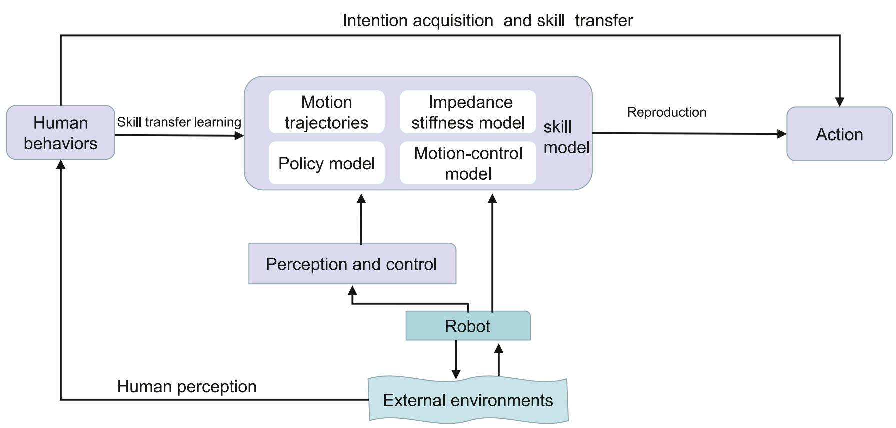


Fig. 2. Diagram of robot STL model.

needs to acquire the human decoding signals from human intention. After obtaining the skill model, the motion policies can be mapped to robot controller. The learned skill model can then be used for reproducing the robot behavior in new environment.

2.2. Interfaces for skill transfer learning

The interface for STL plays a key role in collecting and transmitting information. Generally, STL interfaces can be divided into following modes.

2.2.1. Physical interaction

Physical interaction is also called kinesthetic teaching [23,20]. In this case, the robot is under the physical guidance of humans. This approach allows users to use the robot's own capabilities to demonstrate skills in the robot's environment. A natural demonstration interface is also provided to correct skills reproduced by the robot. In [24], the skills of tactile capabilities were exploited by the kinesthetic teaching. In [25], Keyframe-based learning was proposed by human via physical guidance. Since the user directly manipulates the robot during physical interaction, the robot's movement is restricted within the workspace. Therefore the robot's kinematic are all restricted in the workspace. In addition, no extra devices (e.g., motion capture or teleoperation devices) are needed. However, physical interaction may be challenging for everyday users who have limited experience in manipulating robots with many degrees of freedom.

2.2.2. Teleoperation

In this case, a human user needs to use the robot's own sensors and effectors to perform tasks. Teleoperation can be accomplished using simple joysticks or other remote control devices (e.g., haptic devices [26]). In [27], a bilateral coupling teleoperation system was utilized to perform demonstrations, and statistical model was trained in pure follower/leader role assignment mode that alternates between human and robot. Compared with external motion tracking systems, the advantage of teleoperation is that it completely solves the correspondence problem, because the system directly records the perception and movement from the robot's configuration space. It also outperforms kinesthetic training because it allows remote training and therefore is particularly suitable for teaching navigation and locomotion modes. Since robots and teachers do not need to share same space, teleoperation is used to transmit the kinematics of motion. The kinematics of the demonstration is directly transmitted to the robot and sometimes it combines with the haptic devices to train a motion model based on perceived forces.

2.2.3. Vision and wearable devices

Vision and wearable devices capture human motion by using human body model, which can then be mapped to the robot. The angular displacement of human limbs and joints can be accurately measured by these external means. In [28], human demonstration motions were captured by Vicon, an optical motion capture system. In [29], movements were collected by using a Kinect v2 sensor. Besides mentioned visual capture system, exoskeleton [21], were also utilized to extract human motion for rehabilitation. Human can move freely when using vision and wearable devices; however, the correspondence issues is the main challenge. Since human body and robot have different kinematics and dynamics constraints, transferring human demonstrations to robot can be difficult and even impossible in certain tasks.

2.2.4. Human physiological signals

With the help of various sensors, biological signals can be collected to perceive human's motion intention and assist the estimated motion in real-time. In [30,31], sEMG was used to manipulate a series of tasks consisting of tracking/recogniting/grasping of an object. In [32], the EEG was applied to enable the robot to perform manipulation tasks guided by human operator's mind. Since physiological signals can directly encode the perception and action of human, this method has the benefits of easy accessibility, fast adaptivity, and stability, and thus is particularly suited for limb prosthesis or exoskeleton robot. However, human intension decoding is still in its infancy, which limits its potential applications.

2.3. Skill transfer learning application for autonomous robots and human-robot cooperation

Currently, most STL methods focus on either grasping or manipulation goals. Example applications of STL are listed in Fig. 3, which can be classified into four categories based on its operation environments: industrial manufacturing, service robot, surgery robot, and medical rehabilitation.

Since robot skills are transformable to various scenarios, designing a robot skill model with compactness, comprehensiveness, stability, safety, learning ability and complexity is the key for the robot to acquire, learn, and optimize these skills. Table 1 summaries four STL application domains of autonomous robots and human–robot cooperation.

2.4. Benefits of skill transfer learning

Perform task without programming: Traditional robot programming has to consider all possible situations that a robot might encounter during mission operation. Therefore, the desired task may need to be decomposed into dozens or hundreds of small steps, and the robustness of each step should be tested before applying in practice. When a failure occurs, the system has to be updated to adapt to the new situation by repeating the tedious programming process mentioned before. STL only requires the demonstration of the desired performance from the end-user. No coding as in traditional methods is needed. Even if a failure occurs, the end-user only needs to provide more demonstrations. Robots can even learn to improve their performance based on the interactions with the environment. Therefore, STL is able to transfer human skills via demonstrations to improve the robot's performance in mission operation.

Physiological signals enhance skill transfer: The study of STL is closely related to the analysis and cognition of human behavioral characteristics, where neural signals are often considered as an ideal tool to understand the relationship between neural representations and skill transfer in autonomous robots and human–robot cooperation. Human superb sensorimotor can enhance robot–robot, human–robot, and robot-environment interactions. An immediate application of STL is to decode and transmit neural signals by understanding neural representations to control external devices. Similar to the use of sEMG or EEG in prosthetic and rehabilitative robots [64,65] the skill learning captures the essence of natural behavior and has become one of the focuses in the research of human–machine interface.

3. Skill transfer based on robot learning

Robot learning can implicitly train a robot, so that human users can minimize or eliminate explicit tedious programming of tasks. Most of robot learning methods are data-driven. The data required for robot learning can be generated by interactions between the robot and the environment or provided by experts. Based on this idea, the robot learning for skill acquisition can be classified in the following types.

3.1. Model based learning

Model based learning [66] can reduce the complexity of robot search strategy, and improve the learning efficiency of robot operation skills. According to the use of demonstration data, the learning can be divided into behavioral cloning [67] and IRL [68].

Behavioral cloning is essentially a supervised learning method. It is based on the observed expert's demonstration. Demonstration data sets generically include a series of trajectories, composed states $\mathbf{s}_i$, actions $\mathbf{a}_i$ and next actions $\mathbf{s}_{i+1}$. Then the sampled state–action $\mathcal{D}$ was observed. By the methods of supervised learning, we can obtain the state–action mapping. There are two main methods for skill learning: one is to mimic the motion data using statistical machine learning, and the other is based on dynamical systems [69].

Methods of statistical machine learning. Based on statistical machine learning, STL consists of GMMs, and HMMs. In [33], a human–robot cooperative lifting task was presented by utilizing GMMs to encode and reproduce human behaviors. Human and robot collaboratively lift an object through teleoperation. GMMs was used to capture the robot's motion and interacting forces, while Gaussian mixture regression was utilized to generate the reproduction forces. In [34], the robot was endowed with cognitive ability by using HMMs. The abilities of segmentation, encoding, and clustering for collaborative behavioral primitives were demonstrated. The proposed HMMs with a primitive graph and a primitive tree were incrementally updated for behaviors reproduction [70]. In [35], a hybrid structure was proposed to simplify learning process. The feedforward control signals were generated by GMMs to learn a model, and the parameters of the impedance controller were adjusted based on the motion and force errors generated during the task. In [21], based on hierarchical control scheme, task-parameterized GMMs were presented for cooperations of an exoskeleton robot and human users. The proposed learning model consists of high-level and low-level tasks, where the high-level tasks are the cooperative impedance-based manipulation tasks while the low-level tasks were to address symmetric constraint in an admittance controller. In [71], a non-parametric SOSC algorithm was proposed for online learning and motion synthesis of high-dimensional robot manipulation tasks, which could systematically adapt the model parameters to changing situations such as position/orientation/size of the objects.

Dynamic movement primitives. DMPs is another successfully used behavioral cloning method which represents motion as policy primitives based on nonlinear dynamic systems. Second-order differential equations are often used to describe the system dynamics [72]. The advantage of expressing motion is that the dynamics of the system can automatically correct the disturbance (robustness against disturbance). Early research regarding the robot learning skills only focuses on the system acceleration. In [49], a learning approach was presented to find an acceleration-based predictive reaction for coupled agents to minimize the differences of force signals caused by obstacle avoidance or different paths to follow.

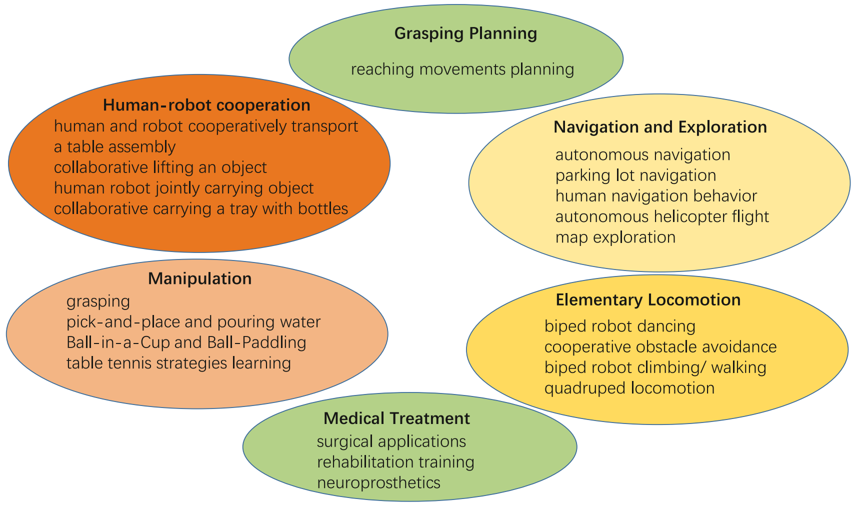


Fig. 3. Some examples of STL applications.

Table 1
STL application domains for autonomous robots and human–robot cooperation.

Application areas	Task descriptions	Examples	References
Industrial manufacture	Making the end effectors have the desired posture relationship while satisfying certain constraints.	peg-in-hole task, bending steel wire, grinding and polishing, weld seam welding, human and robot work together to carry objects, etc.	[20,33,27,34,35], etc.
Service robot	Manipulation or moving in the human environment where the arrangement of objects is usually irregular and dynamic changes often occur.	Ironing clothes and handling objects, grasping, pouring water, playing ball, autonomous navigation and obstacle avoidance, etc.	[36–48,28,49–54], etc.
Surgery robot	Surgical robots have complete difficulty to complete skills autonomously, but robots can learn local surgical skills and perform the local skill autonomously under supervision.	Semiautonomous simulated brain tumor ablation, multilateral cutting of tissue phantoms, flexible insertion and mucosal dissection of tubular instruments, etc.	[55–59], etc.
Medical rehabilitation	Obtaining certain skills to enhance human ability, realize rehabilitation training of paralyzed patients, assist people to carry out various operations and allow users to recover lost functionality by controlling a robotic device with their remaining muscle activity.	Neuroprosthetics controlling and rehabilitation training	[60,21,61–63], etc.

However, system velocities were not considered in the aforementioned learning methods. Recently, various DMPs have been proposed for the human–robot cooperation. In [50], coupled movement primitives were learned by adopting iterative learning control which exploited the force feedback generated during multiple tasks. In [73], trajectory tracking and compliant control were realized without using explicit models of system dynamics. As a variant of DMPs, CMPs encode the position trajectory and the associated torque profiles. In [74], a CMPs based approach was proposed to obtain the motion data set the generation of compliant and accurate motion. A distinguished feature of DMPs is its ability to handle disturbances and system feedback, which makes it a powerful tool for human–robot interaction, manipulation and obstacle avoidance.

DMPs consider one-shot learning and provide spatial and temporal scalability features as well as guaranteed convergence to the target. Besides system acceleration, skill acquisition can also be learned from system velocity. However, the motion dimension in DMPs is represented independently and the correlation between dimensions is lost. Statistical machine learning-based approach directly learns from spatial data which makes it easy to encode multiple demonstrations at one time. However, some characteristics of DMPs are missing, such as the spatial scaling of motion and convergence to the target position. In addition, DMPs require a large number of demonstration data. Recently, ProMPs [75] have been proven to have better reasoning ability than traditional DMPs and multiple demonstrations can be combined to achieve task-specific generalizations.

3.2. Inverse reinforcement learning

Learning a skill from scratch in a complex robotic system is often not feasible or time consuming especially in the condition of limited samples. IRL can obtain reward function based on the given limited demonstration data, thus improving the generalization performance of learning strategy. Consequently, the IRL addresses how to extract appropriate cost function from observed behaviors and reconstruct the expert's policy using RL.

IRL jointly solves the problems of what to imitate and how to imitate. Essentially, IRL is to find a reward function for the task. To determine the reward function, many approaches have been proposed. In [68], max margin principle was proposed based on the demonstration data, according to which the difference between the optimal policy learned based on the reward function and other sub-optimal policies can be minimized. In [76], maximum margin planning framework was proposed based on the max margin principle. The reward function was then transformed into a maximum margin structured prediction problem over a policy space. However, due to potential noise, IRL may not perform well for real robots. In [77], a probabilistic approach based on the principle of maximum entropy was proposed to determine the reward function. The robot control policy still shows good performance even when the demonstration data is corrupted with noise.

The aforementioned methods describe the reward function as a linear combination of hand selected features. Recently, many nonlinear functions have been adopted to design the reward function. Margin-based methods have been proposed to learn nonlinear reward functions through feature construction [78,79]. In [80], GPIRL algorithm was proposed which combined probabilistic reasoning of stochastic expert behavior with the ability to learn the reward as nonlinear functions of features. Thus, GPIRL surpasses previous approaches on tasks with nonlinear rewards and suboptimal examples.

These above traditional methods for obtaining reward function were artificial designed, however, it is still difficult to design an effective reward function algorithm from demonstration for two reasons. First, IRL is fundamentally poorly defined because different cost functions may lead to the same behavior [77]. Second, many standard IRL algorithms solve the forward problem in the inner loop of an iterative cost optimization. Since accurate dynamics models are not always available and the robots are often complex systems with high dimensions, the solution of forward problem is in general difficult to obtain. To address these problems, deep neural network has been applied to represent the reward function [81]. In [82], guided cost learning was proposed which combined the cost function with policy optimization to generate the reward function close to the expert demonstration trajectory. In [83], a generative adversarial network was adopted to optimize the reward function, which obtained significant performance gains over traditional learning methods in large, high-dimensional environments.

3.3. Reinforcement learning

One of the most widely used learning approaches is RL, which enables robots to discover optimal behavior through trial and error with the environment. Instead of providing explicit guidance, the overall performance of the RL is based on a scalar objective function. In contrast to IRL, the goal of RL is to find a policy that optimizes reward function [14]. Different from other approaches that learn from demonstrations, RL allows the robot to discover new control policies by freely exploring the state–action space. According to the type of reward function, RL can be divided into value-function-based approaches and policy search methods.

Value-function-based approaches aim to identify (possibly approximate) solutions to reward function. Many value-function-based RL approaches have been applied in robotic domain. In [84], local linear system was adopted to estimate the reward function in the manipulation of a two-link robot. In [85], a highly-tailored convolutional network architecture with domain-specific optimizations was constructed for predictive models in the manipulation of granular media.

When applying RL to robotic systems, policy search methods are often preferred over value function-based approaches. There are two reasons: (1) policy search methods are easy to incorporate expert experience, which can improve the convergence rate of policy optimization. (2) policy search methods have fewer learning parameters than value-function-based approaches, which is more efficient. In [86], a policy search based reinforcement learning method was applied for an anthropomorphic robot arm to learn to hit a baseball. In [87], a model-free reinforcement learning algorithm was presented to learn motion primitive goals. The proposed method could perform grasping and the pick-and-place task subject to position uncertainties, which improves the robustness of object manipulation skills.

3.4. Deep reinforcement learning

With the development of artificial intelligence and neuroscience, more machine cognitive methodologies have been applied in robot learning. Deep learning [88] and deep reinforcement learning (DRL) [89] have been adopted to help robots learn complex skills. In [89], the end-to-end skills that take raw camera images to compute corresponding motor torques (visual servoing) have been demonstrated impressively. The work of [90] presented an approach to hand-eye coordination for robotic grasping from continuous servoing mechanism based on CNN, which can effectively grasp a wide range of different objects, including objects that have never been trained before. To address the problems of universal picking under the inherent uncertainty in sensing, control, and contact physics, the POMDP framework for ambidextrous robot grasping was proposed which based on robust wrench resistance as a common reward function [91].

Compared to conventional RL methods, DRL algorithms use deep neural networks to express the value functions and policy search methods, which not only avoid the use of artificial features, but also easily integrate perception information into the environment. Many popular DRL algorithms have been proposed such as DQN [92], DPG [93], DDPG [94], TRPO [95], GPS [96] and GQ-CNNs [91]. These algorithms have been successfully used in electronic games and robot control in virtual environment.

Since the deep learning focuses on fitting the value functions and policy search, DRL usually utilizes the gradient descent algorithm to update cost function. As the number of dimensions grows, more data and computation are needed to cover the complete state–action space, leading to low learning efficiency. Evaluating the states quickly becomes infeasible with growing system dimension, even for discrete states.

4. Skill transfer learning via neurophysiological signals

Since human neurophysiology contains rich and useful environmental information, they can facilitate robot skill transfer. Meaningful tactile feedbacks or perceptions have been studied in a large number of neurophysiological and behavioral researches. Human is able to understand the internal mechanism of actions and then utilize it to realize appropriate behavior [12]. Inspired by these sensorimotor abilities, detecting environment through a neuromorphic interface and initiating an automated reflex in the external devices (e.g., prosthesis, robot manipulators, wheel chairs) has increasingly become a promising research area in robotics.

4.1. sEMG based skill acquisition

It has been suggested that human sensorimotor systems can control the impedance of the neuromuscular system [97,98]. Consequently, researchers plans to leverage human adaptation skills to achieve dexterous motion control of robots, which could greatly benefit the physical human–robot interactions and autonomous manipulation in prosthetic devices. A large number of research results have shown that the sEMG signals are closely related to the muscle activations, muscular strength, and limb configuration, which contains a wealth of information about human intentions. The study of [99] aimed to decode shoulder, elbow and wrist dynamic movements continuously and simultaneously based on multi-channel surface electromyography signals, useful for electromyography controlled exoskeleton robots for upper-limb rehabilitation. Based on the sEMG stiffness modulation mechanism in human body, it has been widely used in compliant manipulation or human–robot cooperation to achieve safe interactions with the environment. In [60], a computational model utilizing the sampled sEMGs to calculate the human arm endpoint stiffness was presented. When controlling a two-arm exoskeleton robot, the adaptation skill of human body movement was achieved by adjusting sEMG stiffness.

Moreover, sEMG can also be used as an effective input to control powered prostheses. This control method, also called sEMG control, has been gradually applied to amputees or patients with congenital lack of upper limbs. In an early work of [100], two bipolar EMG electrodes were implanted in the flexor and extensor muscles of the residual limb to control the velocity of closing and opening of prosthesis proportionally. This kind of two-channel sEMG control could only control one DOF at a time, acquiring few skills from human. To address the limitations of this method, researches have extracted more sEMG signals from a large number of channels, while various machine-learning techniques have been proposed to decoding human intension [101].

Machine learning algorithms for sEMG skill transfer can be classified into two categories, i.e., classification and regression. For classification in sEMG, many machine learning techniques can only activate individual functions in sequence [102,32,31], resulting in low classification accuracy. Moreover, this method cannot control the velocity of activated DOFs simultaneously. While regression-based approach is proposed to address this problem, which is available for simultaneous and proportional control [103,104]. Regression does not map discrete motion, but estimates the proportional activation of each DOF. In [105], a deep neural network based method was proposed to derive sEMG-force regression model for force prediction at eight different force levels.

This supports independent, synchronous and proportional control of all DOFs. Specifically, in [63], a proposed regression-based control enables simultaneous and proportional control of two DOFs. The method was robust to the change of arm position.

4.2. Brain-machine interfaces

Machine-driven interface that bridges machine and human nervous system has attracted growing public and research interest, as it provides a means to restore body functions dominated by nerve, spinal cord or brain. Specially, BCIs and BMIs can acquire brain signals and utilize them to control external devices (e.g., computers and machines).

Although human EEGs can vary among individuals, it is believed that the user's motivation and cognitive arousal play an important role in skill acquisition and final task execution [106]. BCIs and BMIs are based on the premise that sensory, motor, and cognitive information can be represented by networks of brain neurons, while the algorithms are designed to explain their complex electrical activities. The produced nerve signals by neurons (e.g., those related to expected actions) are then decoded to control a BMI [107,108].

Skill transfer via EEG has been developed in previous studies and successfully applied for online control of a virtual object [109], wheelchair [110], mobile robot [111] and quadcopter [112]. One of the ultimate goals for BCIs is to enable anthropomorphic movement of highly flexible prosthesis, robot manipulator or exoskeleton. In [64], by deciphering his brain activity, a tetraplegia patient could rapidly learn to grasp or reach an object using the prosthetic limb. However, most of current BMIs were constrained to discrete exploration in one dimension or a plane, without exploring the full possibilities in a 3D space [113–115]. Controlling a robotic arm for complex tasks such as reach-and-grasp in a 3D environment using BMIs is extremely difficult, especially for noninvasive interfaces. Recently, in [61], subjects were able to effectively control the extension of the robotic arm by adjusting their brain rhythm after a few trainings, and maintained this control ability for several months.

4.3. Sensory substitution

Human dexterous manipulation is a complex process where the motor commands, actions and sensory feedback are all coupled. The rapid development of peripheral neural interfaces has promoted the neurophysiology based STL [116]. Invasive intracortically implanted electrode arrays were used to measure the activity neurons in movement-related cortical areas. These arrays are featured of small size and high SNR, and thus can facilitate the control of robot to assist amputees via human neuromorphic interface [117]. Sensory substitution is proposed to provide sensory feedback for exploiting accurate muscle function. Various methods can be used to achieve this goal including implanted electrodes and surface nerve stimulation [118]. In [119], a dexterous hand prosthesis can be manipulated through stimulating the median and ulnar nerve fascicles using transversal multichannel intrafascicular electrodes. Researches have demonstrated that motor neurons can physiologically innervate the muscles that are responsible for the movement of the missing limb and reinnervate other muscle tissues by TMR [120,121], which is useful for the motor sensory skill acquisition for TMR patients.

To date, most of the applications on skill transfer via neurophysiological signals focus on prosthetic and rehabilitative robots, while some of them have been used in remote or dangerous environments [122,123]. Human–robot cooperation motion behaviors rely on the contemplated stimulation of spinal cord, muscles or nerves. Therefore, by decoding low and high level neural information, sophisticated human-to-device interfaces could significantly enhance the human STL capabilities.

5. Discussion and open questions

As for human skill acquisition, to be able to acquire manipulation skills from human, robots should have the ability to learn behaviors from autonomous perception and control. STL can address uncertain models of manipulated objects and robot dynamics, as well as complex robotic systems with a large number of DOFs. Moreover, STL could utilize a user's motivation and cognitive arousal to achieve skill transfer directly. Many of the proposed STL approaches to skill acquisition for robot either depend on learning methodologies or human cognitive and neural signals in place of conventional control-theoretic approaches.

Based on the existing literature on robot STL, the application of STL for autonomous robots and human–robot cooperation mainly focus on the following three points:

- How to generalize and improve the skill acquisition performance by robot learning.
- How to acquire new operational skills with less training data and lower training cost.
- How to decode human's movement intention from neural activity while transferring them to autonomous devices for improved human–robot cooperation.

To this end, the open issues are discussed and future directions are highlighted in this section.

5.1. Robot learning algorithm

The core of robot learning is to model robot perceptions and actions, and thus endows robot with independent decision-making and learning capabilities. STL through robot learning from human demonstration and RL can help us understand the self-organization properties of robot actions. Perception and action are the two essential elements in robot learning, which suggests that the robot motor behavior is based on rich perceptual variables, and vice versa. In addition, the perception can be exploited to produce novel actions.

There are some limitations in endowing robots with new skills by robot learning only. First, learning process requires a large number of data to train the model, e.g., the DRL, and prohibitively long interaction with the system to learn complex skills — typically weeks or months of real-time execution. Over the tedious course of training, the controller may exhibit sudden and chaotic behavior, leading to logistical complications and safety concerns. Though training process are often tested in a simulation environment, refine the training model from simulation environment to a real robot system is still a challenging problem, as it is hard to design an accurate simulator corresponding to the real robot. Alternatively, robot can learn skills from an expert by demonstration, however, sometimes it is hard to find an expert for demonstration, especially for dangerous tasks. In addition, since most robot learning methods are end-to-end, function approximation may not be ensured due to potential instability and divergence when implementing the off-policy.

5.2. Human movement intention acquisition

Skill acquisition through human neurophysiological signals bridges the human intension and external devices. Although human skill can be transferred to robot via human intention, the mechanism of human neurophysiological control is not fully understood. In addition, the invasive approaches face the potential issues of post-surgery complications and infections, and the challenge of maintaining stable chronic recordings, which might limit its application in medical care. Since no surgery is needed and the electrodes can be conveniently placed, noninvasive approaches are more preferred. However, noninvasive systems could not achieve proficient multi-dimensional control of a robotic manipulation in three-dimensional (3D) space. This is because noninvasive approaches do not use proprioceptive feedback, resulting in low accuracy and efficiency of STL. Currently, most applications on neurophysiological signals STL are limited in the sense that only one DOF can be controlled at the same time. A co-contraction or other heuristics algorithms can be adopted to control of multiple DOFs. However, such approaches are counterintuitive and time consuming. Moreover, the simultaneous control of multiple DOFs and the velocity is challenging.

6. Conclusion

In conclusion, an overview on the current state of the art of STL research has been presented. In addition to discussing STL, various other robot learning algorithms are discussed and compared. We also include neurophysiological skill acquisition using STL. In the end, we present the challenges and open questions of applying STL to autonomous robot and human-cooperation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Schaal, The new robotics—towards human centered machines, HFSP J. 1 (2) (2007) 115–126.

[2] S. Mitsi, K.D. Bouzakis, G. Mansour, D. Sagris, G. Maliaris, Off-line programming of an industrial robot for manufacturing, Int. J. Adv. Manuf. Technol. 26 (3) (2005) 262–267.

[3] R.S. Johansson, J. Flanagan, Tactile sensory control of object manipulation in human, in: Handbook of the Senses: Vol. 5-Somatosensation, Elsevier, 2007.

[4] H.H. Ehrsson, A. Fagergren, T. Jonsson, G. Westling, R.S. Johansson, H. Forssberg, Cortical activity in precision- versus power-grip tasks: An fMRI Study, J. Neurophysiol. 83 (1) 528–536.

[5] A.C. Eliasson, H. Forssberg, K. Ikuta, I. Apel, G. Westling, R. Johansson, Development of human precision grip, Exp. Brain Res., 106 (3) 425–433.

[6] S. Schaal, Is imitation learning the route to humanoid robots, Trends Cogn. Sci. 3 (6) (1999) 233–242.

[7] B. Dufay, J. Latombe, An approach to automatic robot programming based on inductive learning, Int. J. Robot. Res. 3 (4) (1984) 3–20.

[8] Y. Yuan, W. Su, Z. Li, G. Shi, Brain-computer interface-based stochastic navigation and control of a semiautonomous mobile robot in indoor environments, IEEE Trans. Cogn. Dev. Syst. 11 (1) (2019) 129–141.

[9] Z. Li, J. Li, S. Zhao, Y. Yuan, Y. Kang, C.L.P. Chen, Adaptive neural control of a kinematically redundant exoskeleton robot using brain-machine interfaces, IEEE Trans. Neural Netw. (2019) 1–14.

[10] J.E. Downey, L. Brane, R.A. Gaunt, E.C. Tylerkabara, M.L. Boninger, J.L. Collinger, Motor cortical activity changes during neuroprosthetic-controlled object interaction, Sci. Rep. 7 (1) (2017) 16947.

[11] K.Z. Zhuang, N. Sommer, V. Mendez, S. Aryan, E. Formento, E. Danna, F. Artoni, F.M. Petrini, G. Granata, G. Cannaviello, et al., Shared human–robot proportional control of a dexterous myoelectric prosthesis, Nat. Mach. Intell. 1 (9) (2019) 400–411.

[12] G. Rizzolatti, L. Fogassi, V. Gallese, Neurophysiological mechanisms underlying the understanding and imitation of action, Nat. Rev. Neurosci. 2 (9) (2001) 661–670.

[13] R. Pfeifer, M. Lungarella, F. Iida, Self-organization, embodiment, and biologically inspired robotics, Science 318 (5853) (2007) 1088–1093.

[14] J. Kober, J. Bagnell, J. Peters, Reinforcement learning in robotics: A survey, Int. J. Robot. Res. 32 (11) (2013) 1238–1274.

[15] B. Kiumarsi, K.G. Vamvoudakis, H. Modares, F.L. Lewis, Optimal and autonomous control using reinforcement learning: A survey, IEEE Trans. Neural Netw. 29 (6) (2018) 2042–2062.

[16] K. Chatzilygeroudis, V. Vassiliades, F. Stulp, S. Calinon, J. Mouret, A survey on policy search algorithms for learning robot controllers in a handful of trials, 2018, arXiv:Robotics.

[17] J. Li, Z. Li, F. Chen, A. Bicchi, Y. Sun, T. Fukuda, Combined sensing, cognition, learning, and control for developing future neuro-robotics systems: A survey, IEEE Trans. Cogn. Dev. Syst. 11 (2) (2019) 148–161.

[18] L. Jamone, E. Ugur, A. Cangelosi, L. Fadiga, A. Bernardino, J. Piater, J. Santosvictor, Affordances in psychology, neuroscience, and robotics: A survey, IEEE Trans. Cogn. Dev. Syst. 10 (1) (2018) 4–25.

[19] W. Burgard, A. Valada, N. Radwan, T. Naseer, J. Zhang, J. Vertens, O. Mees, A. Eitel, G. Oliveira, Perspectives on deep multimodel robot learning, in: Robotics Research, Springer, 2020, pp. 17–24.

[20] L. Rozo, S. Calinon, D.G. Caldwell, P. Jimenez, C. Torras, Learning Physical collaborative robot behaviors from human demonstrations, IEEE Trans. Robot. 32 (3) 513–527.

[21] D. Mingdi, L. Zhijun, K. Yu, C.C.L. Philip, C. Xiaoli, A learning-based hierarchical control scheme for an exoskeleton robot in human-robot cooperative manipulation, IEEE Trans. Cybern. (2018).

[22] S. Calinon, F. Guenter, A. Billard, On learning, representing, and generalizing a task in a humanoid robot, IEEE Trans. Syst. Man Cybern. B 37 (2) (2007) 286–298.

[23] G. Maeda, M. Ewerton, G. Neumann, R. Lioutikov, J. Peters, Phase estimation for fast action recognition and trajectory generation in human–robot collaboration, Int. J. Robot. Res. 36 (2017) 1579–1594.

[24] E.L. Sauser, B.D. Argall, G. Metta, A. Billard, Iterative learning of grasp adaptation through human corrections, Robot. Auton. Syst. 60 (1) (2012) 55–71.

[25] B. Akgun, M. Cakmak, K. Jiang, A.L. Thomaz, Keyframe-based learning from demonstration method and evaluation, Int. J. Soc. Robot. 4 (4) (2012) 343–355.

[26] C. Yang, X. Wang, Z. Li, Y. Li, C. Su, Teleoperation control based on combination of wave variable and neural networks, IEEE Trans. Syst. Man Cybern. A 47 (8) (2017) 2125–2136.

[27] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, A. Kheddar, Teaching physical collaborative tasks: Object-lifting case study with a humanoid, in: 2009 9th IEEE-RAS International Conference on Humanoid Robots, IEEE, 2009, pp. 399–404.

[28] S. Nakaoka, A. Nakazawa, F. Kanehiro, K. Kaneko, M. Morisawa, H. Hirukawa, K. Ikeuchi, learning from observation paradigm: Leg task models for enabling a biped humanoid robot to imitate human dances, Int. J. Robot. Res., 26 (8) 829–844.

[29] C. Li, C. Yang, Z. Ju, A.S. Annamalai, An enhanced teaching interface for a robot using DMP and GMR, Int. J. Intell. Robot. Appl. 2 (1) (2018) 110–121.

[30] G. Lin, Z. Li, L. Liu, H. Su, W. Ye, Development of multi-fingered dexterous hand for grasping manipulation, Sci. China F 57 (12) (2014) 1–10.

[31] Y. Hu, Z. Li, G. Li, P. Yuan, C. Yang, R. Song, Development of sensory-motor fusion-based manipulation and grasping control for a robotic hand-eye system, IEEE Trans. Syst. Man Cybern. A 47 (7) (2016) 1169–1180.

[32] Z. Li, B. Wang, F. Sun, C. Yang, Q. Xie, W. Zhang, SEMG-based joint force control for an upper-limb power-assist exoskeleton robot, IEEE J. Biomed. Health Inf. 18 (3) (2014) 1043–1050.

[33] S. Calinon, P. Evrard, E. Gribovskaya, A. Billard, A. Kheddar, Learning collaborative manipulation tasks by demonstration using a haptic interface, in: 2009 International Conference on Advanced Robotics, IEEE, 2009, pp. 1–6.

[34] J.R. Medina, M. Lawitzky, A. Mörtl, D. Lee, S. Hirche, An experience-driven robotic assistant acquiring human knowledge to improve haptic cooperation, in: 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2011, pp. 2416–2422.

[35] E. Gribovskaya, A. Kheddar, A. Billard, Motion learning and adaptive impedance for robot control during physical interaction with humans, in: Proceedings - IEEE International Conference on Robotics and Automation, 2011.

[36] S. Levine, N. Wagener, P. Abbeel, Learning contact-rich manipulation skills with guided policy search, 2015, arXiv preprint arXiv:1501.05611.

[37] A. Billard, M.J. Mataric, Learning human arm movements by imitation:: Evaluation of a biologically inspired connectionist architecture, Robot. Auton. Syst. 37 (2) (2001) 145–160.

[38] L.P. Kaelbling, T. Lozano-Pérez, Unifying perception, estimation and action for mobile manipulation via belief space planning, in: 2012 IEEE International Conference on Robotics and Automation, IEEE, 2012, pp. 2952–2959.

[39] M.R. Dogar, M.C. Koval, A. Tallavajhula, S.S. Srinivasa, Object search by manipulation, in: International Conference on Robotics and Automation, vol. 36, (1) 2013, pp. 153–167.

[40] R. Platt, L. Kaelbling, T. Lozano-Perez, R. Tedrake, Efficient planning in non-Gaussian belief spaces and its application to robot grasping, in: Robotics Research, Springer, 2017, pp. 253–269.

[41] Z. Li, T. Zhao, F. Chen, Y. Hu, C. Su, T. Fukuda, Reinforcement learning of manipulation and grasping using dynamical movement primitives for a humanoidlike mobile manipulator, IEEE/ASME Trans. Mechatronics 23 (1) (2017) 121–131.

[42] T. Zhao, M. Deng, Z. Li, Y. Hu, Cooperative manipulation for a mobile dual-arm robot using sequences of dynamic movement primitives, IEEE Trans. Cogn. Dev. Syst. (2018) 1.

[43] O. Kroemer, R. Detry, J. Piater, J. Peters, Combining active learning and reactive control for robot grasping, Robot. Auton. Syst. 58 (9) (2010) 1105–1116.

[44] Z. Xue, S.W. Ruehl, A. Hermann, T. Kerscher, R. Dillmann, Autonomous grasp and manipulation planning using a ToF camera, Robot. Auton. Syst. 60 (3) (2012) 387–395.

[45] R. Jonschkowski, O. Brock, Learning state representations with robotic priors, Auton. Robots 39 (3) (2015) 407–428.

[46] D. Silver, J.A. Bagnell, A. Stentz, Learning from demonstration for autonomous navigation in complex unstructured terrain, Int. J. Robot. Res. 29 (12) (2010) 1565–1592.

[47] P. Abbeel, A. Coates, A.Y. Ng, Autonomous helicopter aerobatics through apprenticeship learning, Int. J. Robot. Res. 29 (13) (2010) 1608–1639.

[48] T. Kollar, N. Roy, Trajectory optimization using reinforcement learning for map exploration, Int. J. Robot. Res. 27 (2) (2008) 175–196.

[49] T. Kulvicius, M. Biehl, M.J. Aein, M. Tamosiunaite, F. Worgotter, Interaction learning for dynamic movement primitives used in cooperative robotic tasks, Robot. Auton. Syst. 61 (12) (2013) 1450–1459.

[50] A. Gams, B. Nemec, A.J. Ijspeert, A. Ude, Coupling movement primitives: Interaction with the environment and bimanual tasks, IEEE Trans. Robot. 30 (4) (2014) 816–830.

[51] J. Kolter, A.Y. Ng, Policy search via the signed derivative., in: Robotics: Science and Systems, 2009, p. 34.

[52] T. Matsubara, J. Morimoto, J. Nakanishi, M. Sato, K. Doya, Learning CPG-based biped locomotion with a policy gradient method, Robot. Auton. Syst. 54 (11) (2006) 911–920.

[53] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, M. Hutter, Learning agile and dynamic motor skills for legged robots, Sci. Robot. 4 (26) (2019).

[54] E. Theodorou, J. Buchli, S. Schaal, Reinforcement learning of motor skills in high dimensions: A path integral approach, in: 2010 IEEE International Conference on Robotics and Automation, IEEE, 2010, pp. 2397–2403.

[55] S. Calinon, D. Bruno, M.S. Malekzadeh, T. Nanayakkara, D.G. Caldwell, Human–robot skills transfer interfaces for a flexible surgical robot, Comput. Methods Programs Biomed. 116 (2) (2014) 81–96.

[56] M.S. Malekzadeh, S. Calinon, D. Bruno, D.G. Caldwell, Learning by imitation with the STIFF-FLOP surgical robot: a biomimetic approach inspired by octopus movements, Robot. Biomim. 1 (1) (2014) 13.

[57] D. Hu, Y. Gong, B. Hannaford, E.J. Seibel, Semi-autonomous simulated brain tumor ablation with ravenii surgical robot using behavior tree, in: 2015 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2015, pp. 3868–3875.

[58] W. Xu, J. Chen, H.Y. Lau, H. Ren, Automate surgical tasks for a flexible serpentine manipulator via learning actuation space trajectory from demonstration, in: 2016 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2016, pp. 4406–4413.

[59] A. Murali, S. Sen, B. Kehoe, A. Garg, S. McFarland, S. Patil, W.D. Boyd, S. Lim, P. Abbeel, K. Goldberg, Learning by observation for surgical subtasks: Multilateral cutting of 3d viscoelastic and 2d orthotropic tissue phantoms, in: 2015 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2015, pp. 1202–1209.

[60] Z. Li, C. Xu, Q. Wei, C. Shi, C. Su, Human-inspired control of dual-arm exoskeleton robots with force and impedance adaptation, IEEE Trans. Syst. Man Cybern. (2018) 1–10.

[61] J. Meng, S. Zhang, A. Bekyo, J. Olsoe, B. Baxter, B. He, Noninvasive electroencephalogram based control of a robotic arm for reach and grasp tasks, Sci. Rep. 6 (1) (2016) 38565.

[62] D. Borton, S. Micera, J.d.R. Millan, G. Courtine, Personalized neuroprosthetics, Sci. Transl. Med. 5 (210) 210rv2–210rv2.

[63] J.M. Hahne, S.M. A., K. Mario, F. Dario, Simultaneous control of multiple functions of bionic hand prostheses: Performance and robustness in end users, Sci. Robot. 3 (19) eaat3630.

[64] J.L. Collinger, B. Wodlinger, J.E. Downey, W. Wang, E.C. Tylerkabara, D.J. Weber, A.J.C. Mcmorland, M. Velliste, M.L. Boninger, A.B. Schwartz, High-performance neuroprosthetic control by an individual with tetraplegia, Lancet 381 (9866) (2013) 557–564.

[65] O. Fukuda, T. Tsuji, M. Kaneko, A. Otsuka, A human-assisting manipulator teleoperated by EMG signals and arm motions, in: International Conference on Robotics and Automation, vol. 19, (2) 2003, pp. 210–222.

[66] D. Nguyentuong, J. Peters, Model learning for robot control: a survey, Cogn. Process. 12 (4) (2011) 319–340.

[67] S. Ross, G. Gordon, D. Bagnell, A reduction of imitation learning and structured prediction to no-regret online learning, in: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, 2011, pp. 627–635.

[68] P. Abbeel, A.Y. Ng, Apprenticeship learning via inverse reinforcement learning, in: Proceedings of the Twenty-First International Conference on Machine Learning, 2004, pp. 1.

[69] A.J. Ijspeert, J. Nakanishi, S. Schaal, Movement imitation with nonlinear dynamical systems in humanoid robots, in: Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292), Vol. 2, IEEE, 2002, pp. 1398–1403.

[70] D. Kulic, W. Takano, Y. Nakamura, Incremental learning, clustering and hierarchy formation of whole body motion patterns using adaptive hidden Markov chains, Int. J. Robot. Res. 27 (7) (2008) 761–784.

[71] A.K. Tanwani, S. Calinon, Small-variance asymptotics for non-parametric online robot learning, Int. J. Robot. Res. 38 (1) (2019) 3–22.

[72] P. Pastor, H. Hoffmann, T. Asfour, S. Schaal, Learning and generalization of motor skills by learning from demonstration, in: 2009 IEEE International Conference on Robotics and Automation, IEEE, 2009, pp. 763–768.

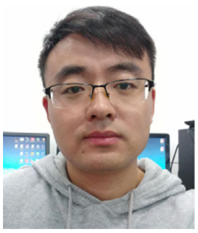
[73] M. Denisa, A. Gams, A. Ude, T. Petric, Learning compliant movement primitives through demonstration and statistical generalization, IEEE/ASME Trans. Mechatronics 21 (5) (2016) 2581–2594.

[74] T. Petric, A. Gams, L. Colasanto, A.J. Ijspeert, A. Ude, Accelerated sensorimotor learning of compliant movement primitives, IEEE Trans. Robot. 34 (6) (2018) 1636–1642.

[75] E. Rueckert, J. Mundo, A. Paraschos, J. Peters, G. Neumann, Extracting low-dimensional control variables for movement primitives, in: 2015 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2015, pp. 1511–1518.

[76] N.D. Ratliff, J. Bagnell, M.A. Zinkevich, Maximum margin planning, in: Proceedings of the 23rd International Conference on Machine Learning, 2006, pp. 729–736.

[77] B.D. Ziebart, A.L. Maas, J.A. Bagnell, A.K. Dey, Maximum entropy inverse reinforcement learning., in: AAAI, Vol. 8, Chicago, IL, USA 2008, pp. 1433–1438.
[78] S. Levine, Z. Popovic, V. Koltun, Feature construction for inverse reinforcement learning, in: Advances in Neural Information Processing Systems, 2010, pp. 1342–1350.
[79] J. Bagnell, J. Chestnutt, D.M. Bradley, N.D. Ratliff, Boosting structured prediction for imitation learning, in: Advances in Neural Information Processing Systems, 2007, pp. 1153–1160.
[80] S. Levine, Z. Popovic, V. Koltun, Nonlinear inverse reinforcement learning with gaussian processes, in: Advances in Neural Information Processing Systems, 2011, pp. 19–27.
[81] M. Wulfmeier, P. Ondruska, I. Posner, Maximum entropy deep inverse reinforcement learning, 2015, arXiv preprint arXiv:1507.04888.
[82] C. Finn, S. Levine, P. Abbeel, Guided cost learning: Deep inverse optimal control via policy optimization, in: International Conference on Machine Learning, 2016, pp. 49–58.
[83] J. Ho, S. Ermon, Generative adversarial imitation learning, in: Advances in Neural Information Processing Systems, 2016, pp. 4565–4573.
[84] R. Lioutikov, A. Paraschos, J. Peters, G. Neumann, Sample-based informationl-theoretic stochastic optimal control, in: 2014 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2014, pp. 3896–3902.
[85] C. Schenck, J. Tompson, S. Levine, D. Fox, Learning robotic manipulation of granular media, in: S. Levine, V. Vanhoucke, K. Goldberg (Eds.), Proceedings of the 1st Annual Conference on Robot Learning, in: Proceedings of Machine Learning Research, vol. 78, PMLR, 2017, pp. 239–248, URL http://proceedings.mlr.press/v78/schenck17a.html.
[86] J. Peters, S. Schaal, 2008 special issue: Reinforcement learning of motor skills with policy gradients, Neural Netw. 21 (4) (2008) 682–697.
[87] F. Stulp, E.A. Theodorou, S. Schaal, Reinforcement learning with sequences of motion primitives for robust manipulation, IEEE Trans. Robot. 28 (6) (2012) 1360–1370.
[88] G.E. Hinton, R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (5786) (2006) 504–507.
[89] S. Levine, C. Finn, T. Darrell, P. Abbeel, End-to-end training of deep visuomotor policies, J. Mach. Learn. Res. 17 (1) (2016) 1334–1373.
[90] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, D. Quillen, Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection, Int. J. Robot. Res. 37 (2018) 421–436.
[91] J. Mahler, M. Matl, V. Satish, M. Danielczuk, B. DeRose, S. McKinley, K. Goldberg, Learning ambidextrous robot grasping policies, Sci. Robot. 4 (26) (2019) eaau4984.
[92] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, et al., Human-level control through deep reinforcement learning, Nature 518 (7540) (2015) 529–533.
[93] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, M. Riedmiller, Deterministic policy gradient algorithms, in: International Conference on Machine Learning, 2014, pp. 387–395.
[94] T.P. Lillicrap, J.J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, D. Wierstra, Continuous control with deep reinforcement learning, 2015, arXiv preprint arXiv:1509.02971.
[95] J. Schulman, S. Levine, P. Abbeel, M. Jordan, P. Moritz, Trust region policy optimization, in: International Conference on Machine Learning, 2015, pp. 1889–1897.
[96] S. Levine, V. Koltun, Guided policy search, in: International Conference on Machine Learning, 2013, pp. 1–9.
[97] N. Hogan, Adaptive control of mechanical impedance by coactivation of antagonist muscles, IEEE Trans. Automat. Control 29 (8) (1984) 681–690.
[98] T.S. Buchanan, D.G. Lloyd, K. Manal, T.F. Besier, Neuromusculoskeletal modeling: Estimation of muscle forces and joint moments and movements from measurements of neural command, J. Appl. Biomech. 20 (4) (2004) 367–395.
[99] J. Liu, Y. Ren, D. Xu, S.H. Kang, L.-Q. Zhang, EMG-based real-time linear-nonlinear cascade regression decoding of shoulder, elbow and wrist movements in able-bodied persons and stroke survivors, IEEE Trans. Biomed. Eng. (2019).
[100] E.D. Sherman, A. Lippay, A Russian bioelectric-controlled prosthesis, CMAJ 92 (5) (1965) 243.
[101] K. Englehart, B. Hudgins, A robust, real-time control scheme for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 50 (7) (2003) 848–854.
[102] A.J. Young, L.H. Smith, E.J. Rouse, L.J. Hargrove, Classification of simultaneous movements using surface EMG pattern recognition, IEEE Trans. Biomed. Eng. 60 (5) (2012) 1250–1258.
[103] A. Ameri, E.N. Kamavuako, E.J. Scheme, K.B. Englehart, P.A. Parker, Real-time, simultaneous myoelectric control using visual target-based training paradigm, Biomed. Signal Process. Control 13 (2014) 8–14.
[104] J.M. Hahne, F. Biessmann, N. Jiang, H. Rehbaum, D. Farina, F.C. Meinecke, K.-R. Müller, L.C. Parra, Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control, IEEE Trans. Neural Syst. Rehabil. Eng. 22 (2) (2014) 269–279.
[105] C. Li, J. Ren, H. Huang, B. Wang, Y. Zhu, H. Hu, PCA and deep learning based myoelectric grasping control of a prosthetic hand, Biomed. Eng. Online 17 (1) (2018) 107.
[106] S. Musallam, B. Corneil, B. Greger, H. Scherberger, R.A. Andersen, Cognitive control signals for neural prosthetics, Science 305 (5681) (2004) 258–262.
[107] A.B. Schwartz, Cortical neural prosthetics, Annu. Rev. Neurosci. 27 (1) (2004) 487–507.
[108] M.A. Lebedev, M.A. Nicolelis, Brain-machine interfaces: past, present and future, rends Neurosci. 29 (9) (2006) 536–546.
[109] B. Xia, O. Maysam, S. Veser, L. Cao, J. Li, J. Jia, H. Xie, N. Birbaumer, A combination strategy based brain–computer interface for two-dimensional movement control, J. Neural Eng. 12 (4) (2015) 046021.
[110] T.J. Bradberry, R.J. Gentili, J.L. Contrerasvidal, Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals, J. Neurosci. 30 (9) (2010) 3432–3437.
[111] Y. Liu, W. Su, Z. Li, G. Shi, X. Chu, Y. Kang, W. Shang, Motor-imagery-based teleoperation of a dual-arm robot performing manipulation tasks, IEEE Trans. Cogn. Dev. Syst. 11 (3) (2019) 414–424.
[112] K. LaFleur, K. Cassady, A. Doud, K. Shades, E. Rogin, B. He, Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface, J. Neural Eng. 10 (4) (2013) 046003.
[113] E. Hortal, D. Planelles, A. Costa, E. Ianez, A. Ubeda, J.M. Azorin, E. Fernandez, SVM-based brain-machine interface for controlling a robot arm through four mental tasks, Neurocomputing 151 (2015) 116–121.
[114] T. Sakurada, T. Kawase, K. Takano, T. Komatsu, K. Kansaku, A BMI-based occupational therapy assist suit: asynchronous control by SSVEP, Front. Neurosci. 7 (2013) 172.
[115] P. Horki, T. Solis-Escalante, C. Neuper, G. Müller-Putz, Combined motor imagery and SSVEP based BCI control of a 2 DoF artificial upper limb, Med. Biol. Eng. Comput. 49 (5) (2011) 567–577.
[116] X. Navarro, T.B. Krueger, N. Lago, S. Micera, T. Stieglitz, P. Dario, A critical review of interfaces with the peripheral nervous system for the control of neuroprostheses and hybrid bionic systems, J. Peripher. Nerv. Syst. 10 (3) (2005) 229–258.
[117] L.R. Hochberg, D. Bacher, B. Jarosiewicz, N.Y. Masse, J.D. Simeral, J. Vogel, S. Haddadin, J. Liu, S.S. Cash, P.V. Der Smagt, et al., Reach and grasp by people with tetraplegia using a neurally controlled robotic arm, Nature 485 (7398) (2012) 372–375.
[118] E. Danna, G. Valle, A. Mazzoni, I. Strauss, F. Iberite, J. Patton, F.M. Petrini, S. Raspopovic, G. Granata, R.D. Lorio, et al., A closed-loop hand prosthesis with simultaneous intraneural tactile and position feedback, bioRxiv, 2018, 262741.
[119] S. Raspopovic, M. Capogrosso, F.M. Petrini, M. Bonizzato, J. Rigosa, G.D. Pino, J. Carpaneto, M. Controzzi, T. Boretius, E. Fernandez, et al., Restoring natural sensory feedback in real-time bidirectional hand prostheses, Sci. Transl. Med. 6 (222) (2014).
[120] D. Farina, I. Vujaklija, M. Sartori, T. Kapelner, F. Negro, N. Jiang, K.D. Bergmeister, A. Andalib, J.C. Principe, O.C. Aszmann, Man/machine interface based on the discharge timings of spinal motor neurons after targeted muscle reinnervation, Nat. Biomed. Eng. 1 (2) (2017) 0025.
[121] T.A. Kuiken, G. Li, B.A. Lock, R.D. Lipschutz, L.A. Miller, K.A. Stubblefield, K. Englehart, Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms., JAMA 301 (6) (2009) 619–628.
[122] P. Artemiadis, K.J. Kyriakopoulos, EMG-based control of a robot arm using low-dimensional embeddings, IEEE Trans. Robot. 26 (2) (2010) 393–398.
[123] J. Vogel, C. Castellini, P. van der Smagt, EMG-based teleoperation and manipulation with the DLR LWR-III, in: 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2011, pp. 672–678.



Yueyue Liu received the M. S. degree in control engineering from the South China University of Technology, Guangzhou, China, in 2017. Now he is currently pursuing the Ph.D. degree with the College of Automation Science and Engineering. His research interests include mobile manipulation and autonomous robot.

Zhijun Li received the Ph.D. degree in mechatronics, Shanghai Jiao Tong University, P. R. China, in 2002. From 2003 to 2005, he was a postdoctoral fellow in Department of Mechanical Engineering and Intelligent systems, The University of Electro-Communications, Tokyo, Japan. From 2005 to 2006, he was a research fellow in the Department of Electrical and Computer Engineering, National University of Singapore, and Nanyang Technological University, Singapore. From 2017, he is a Professor in Department of Automation, University of Science and Technology, Hefei, China. From 2019, he is the Vice Dean of School of Information Science and Technology, University of Science and Technology of China, China.

From 2016, he has been the Co-Chairs of IEEE SMC Technical Committee on Bio-mechatronics and Bio-robotics Systems (B2S), and IEEE RAS Technical Committee on Neuro-Robotics Systems. He is serving as an Editor at-large of Journal of Intelligent & Robotic Systems, and Associate Editors of several IEEE Transactions. Dr. Li's current research interests include wearable robotics, tele-operation systems, nonlinear control, neural network optimization, etc.

Huaping Liu received the Ph.D. degree from Tsinghua University, Beijing, China, in 2004. He is currently an Associate Professor with the Department of Computer Science and Technology, Tsinghua University. His research interests include robot perception and learning. Dr. Liu served as a Program Committee Member for RSS2016 and IJCAI2016. He serves as an Associate Editor for several journals including IEEE ROBOTICS AND AUTOMATION LETTERS, Neurocomputing, Cognitive Computation, and some conferences including the International Conference on Robotics and Automation and the International Conference on Intelligent Robots and Systems.

Zhen Kan received the Ph.D. degree in mechanical and aerospace engineering from the University of Florida, Gainesville, FL, USA, in 2011. He is a Professor with the Department of Automation, University of Science and Technology of China, Hefei, China. He was a Postdoctoral Research Fellow with the Air Force Research Laboratory, Eglin AFB, FL, USA, and the University of Florida REEF, Shalimar, FL, USA, from 2012 to 2016, and an Assistant Professor with the Department of Mechanical Engineering, University of Iowa, Iowa City, IA, USA. His current research interests include networked robotic systems, Lyapunov-based nonlinear control, graph theory, complex networks, and human-assisted estimation, planning, and decision making. Prof. Kan currently serves as an Associate Editor on the Conference Editorial Board of the IEEE Control Systems Society and Technical Committee for several internationally recognized scientific and engineering conferences.