IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 4, 2021, accepted April 7, 2021, date of publication May 17, 2021, date of current version May 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3081449

Sensorimotor Skill Communication: A Literature Review

VAHAN BABUSHKIN[1], MUHAMMAD HASSAN JAMIL[2], WANJOO PARK[2], AND MOHAMAD EID[2], (Senior Member, IEEE)

[1]NYU Tandon School of Engineering, New York University, Brooklyn, NYC 11201, USA
[2]Division of Engineering, New York University at Abu Dhabi, Abu Dhabi, United Arab Emirates

Corresponding author: Vahan Babushkin (vahan.babushkin@nyu.edu)

This work was supported by the ADEK Award for Research Excellence (AARE) 2019 Program under Project AARE19-159.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board for Protection of Human Subjects at New York University Abu Dhabi under Application No. HRPP-2019-026.

ABSTRACT A sensorimotor skill is a sequence of motions generated in response to external stimuli and aiming to accomplish a particular task. It can be communicated to reproduce the task in a distant environment with similar settings. In this work, we conceptualize a multi-modal sensorimotor skill communication system that incorporates modeling, simulation, and evaluation of the sensorimotor skill. The proposed sensorimotor skill communication system can be applied for learning a specific style of human sensorimotor skill and teaching the skill to distant learners, which can be implemented in a variety of applications such as Tele-consultation, Tele-diagnosis, Tele-treatment, Tele-monitoring, and Tele-support. To understand the processes behind the communication of sensorimotor skill we review the representation of a human sensorimotor system from the neurobiological perspective. Then we analyze the existing literature on sensorimotor skill communication systems and propose a taxonomy of currently available methods for sensorimotor skill modeling, simulation, and evaluation. Furthermore, we propose a benchmark for evaluating the quality of the sensorimotor skill communication system. We present a case study aiming to demonstrate modeling the dental sensorimotor skill of periodontal probing. Lastly, we discuss challenges and limitations and provide perspectives for future research in developing sensorimotor skill communication systems.

INDEX TERMS Haptics and haptic interfaces, learning from demonstration, sensorimotor learning, virtual reality and interfaces.

I. INTRODUCTION

A. SENSORIMOTOR SKILL

A human skill is the ability to perform a variety of tasks using past knowledge and previous experience and can be mastered gradually through learning and practice [1]. A sensorimotor skill, sometimes referred to as a perceptual-motor or psychomotor skill, is the process of receiving information about our bodies and/or the physical environment through our sensory systems (such as vision, audition, cutaneous, and proprioception), and generating a perceptual or cognitive state that produces an appropriate motor response (movement or force) in order to complete a physical task [2]. Examples of sensorimotor skills include walking, running, handwriting, drawing, etc.

The associate editor coordinating the review of this manuscript and approving it for publication was Chaitanya U. Kshirsagar.

Intuitively, a sensorimotor skill can be considered as a mapping of stimuli to responses. From this point of view, learning a sensorimotor skill is a process of inferring this mapping [1]. An accurate model of human sensorimotor skill will not only improve the communication of the skill, thus facilitating its acquisition, but will also help to understand human behavior, and how behavioral patterns affect the proficiency in a given skill and vice versa [2]–[4]. The expertise in the skill is determined by several factors, such as the precision of movement, the latency of the gaze [5], [6] and the speed of decision-making to name a few [2]. Modeling the human's sensorimotor skill will facilitate a better understanding of the processes involved in sensorimotor skill communication and learning, leading to new developments in science and industry.

Researchers' interest in sensorimotor skill theories shows a steady growth for the past decade. The publication search in

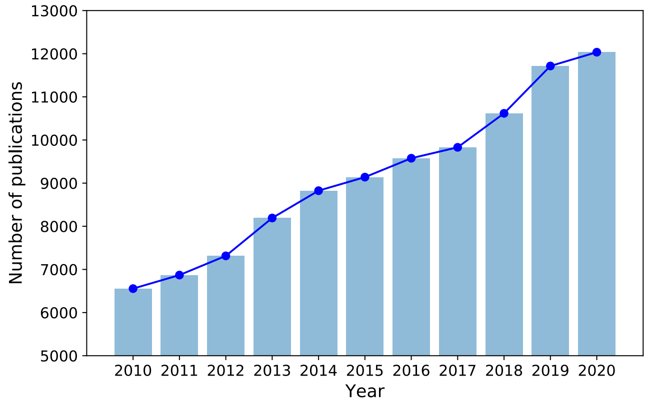

FIGURE 1. Publications in ACM, IEEE, Science Direct, and Springer on sensorimotor skill communication for the last decade. The search has been conducted by keywords with the condition "(Sensori-motor OR Sensorimotor OR Haptic) AND (Modeling OR Simulation OR Evaluation OR Communication)".

top four engineering digital libraries (ACM, IEEE, Science Direct and Springer) with the keywords ''(Sensori-motor OR Sensorimotor OR Haptic) AND (Modeling OR Simulation OR Evaluation OR Communication)'' reveals a strongly growing trend in the number of publications from 2010 to 2020 (see Figure 1). These results highlight the importance of surveying the current situation in the growing field of sensorimotor skill communication.

Sensorimotor skills are typically classified as gross and fine skills. The gross sensorimotor skills incorporate general movements required for performing the task. For humans, the gross sensorimotor skills do not differ between individuals. In contrast, fine sensorimotor skills are individualized and manifest the level of proficiency in performing the task. In other words, gross motor skills define the generic aspects of the sensorimotor skill and fine motor skills describe specific expert's skill. For instance, in the periodontal task, the gross motor skills might refer to the postural control needed for positioning the hand, while the fine motor skills determine the performance in pocket probing, calculus detection, and calculus removal.

B. THE EMERGENCE OF TACTILE INTERNET

Since the commencement of the ARPANET project in 1969, humanity witnessed drastic progress in communication and information exchange technologies. The open architecture networking concept delineated the future evolution of the Internet into a worldwide decentralized infrastructure of interconnected computers and local networks. Further development of the technology facilitated the appearance of the Mobile Internet, connecting smartphones, tablets, laptops, and other electronic devices. Advancements in smart sensors, Internet protocols, and machine-to-machine (M2M) technologies led to the emergence of the Internet of Things (IoT) as the next generation of Mobile Internet [7].

The further advancements in 5G technologies spurred the emergence of highly reliable, low latency networks, enabling the development of real-time interactive systems for remote controlling and communication of tactile experiences [8]. It gave rise to the idea of Tactile Internet, as the next generation of Internet of Things [9], [10], to enable physical interaction between humans and/or machines over distance. The concept of Tactile Internet will transform the current model of the Internet to the ''Internet of Skills''; a network for haptic-audio-visual communication with ultra-low latency (1 ms end-to-end delay), extremely high availability, reliability, and security infrastructure [11].

The development of Tactile Internet and 5G networks enables the possibility of human sensorimotor skill communication over distance. The idea of transferring human skill over the Internet, synchronously in real-time or asynchronously by recording and playing back the skill, will reshape the future of education, training, tele-operation, and inter-personal communication. Imagine, for example, a skilled dentist in New York teaching periodontal procedures to students across the globe, not only through visual demonstration but also by physically guiding the student's hand movements via haptic communication, thus communicating their perceptual and tactile experience to the student. Another example would be to communicate a highly tactile artistic skill, such as calligraphy, to students anywhere in the world. In the long run, the evolution of Tactile Internet will drastically change the traditional ways of teaching, acquiring, and communicating sensorimotor skills as well as interacting with the remote or virtual environment.

C. APPLICATION SCENARIOS

In the context of Tactile Internet, sensorimotor skill communication is manifested through three application scenarios: (1) interpersonal communication in joint actions (competitive or collaborative), (2) training and learning of sensorimotor skill at a distance (synchronous or asynchronous), and (3) Tele-operation where a skilled human operator performs a sensorimotor task remotely by controlling a tele-operated robotic system.

Figure 2 summarizes the application scenarios of sensorimotor skill communication systems. The modeled and simulated sensorimotor skills from a human expert are communicated over a computer network. These skills can be used either to train another user or to control a tele-operated robot, thus enhancing the tele-presence in a remote environment. The sensorimotor skill recorded from an expert can also be implemented in a virtual reality simulation and later used for offline training, i.e. without immediate guidance by the expert. For instance, the expert's skills can be recorded and displayed in dental simulators, that use haptic guidance for training students to perform the periodontal procedures [12]. The experts can also hone their skills and work over the details of the task simulated in a virtual environment, before moving to the actual task. For example, neurosurgeons can practice locating and removal of epileptogenic brain areas in a virtual simulation of patient's brain, before performing the actual surgery.

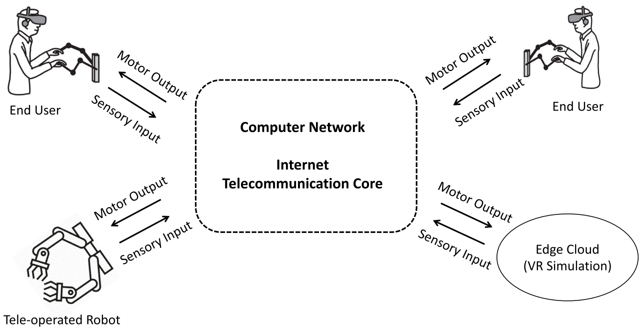

FIGURE 2. Application scenarios for the sensorimotor skill communication system.

The study of sensorimotor forms of communication to improve inter-personal interaction has a long history [13]. Indeed, sensorimotor communication is an agile and convenient way to prompt information transfer in both cooperative and competitive online interactions, especially in the settings in which the effectiveness of other forms of communications is compromised [14]. Signaling is a form of sensorimotor communication when individuals deliberately alter the kinematics of their actions to convey information and to clarify their goals [14], [15]. The ability to encode observed actions (motor coding) plays an important role in coordinating joint actions [16]. Simulating these actions in a sensorimotor system makes it possible to predict, monitor, and adapt to the behavior of others [17]. Sensorimotor communication also appears in leader-follower interactions, i.e. in tasks with asymmetric information, when one of the partners (leader) possesses the knowledge, which the other partner (follower) cannot acquire independently [15], [17]–[19]. Communication of sensorimotor skill is also vital for Human-Robot Interaction [20], [21].

Advancements in haptic technologies and Virtual Reality have aroused the interest in teaching novel sensorimotor skills over a computer network. A classical example involves an expert using multimodal sensing devices to capture the haptic, auditory, and visual cues necessary to reproduce a sensorimotor skill, communicate these cues over a computer network, and utilize multimodal actuators to display the corresponding sensorimotor skill to a learner. Indeed, leader-follower signaling can be implemented by amplifying those components of sensorimotor skill that are difficult to comprehend [14]. For instance, some specific details of fine sensorimotor skills associated with tele-surgery [22] or handwriting [23], [24] can be communicated implicitly from a teacher to a student via real-time (online) haptic interaction.

Sensorimotor skill communication is greatly manifested through tele-operated robotic systems involved in physical interaction including locomotion, object manipulation, carrying loads, etc. Providing the tele-operator with haptic, audio, and visual feedback, significantly improves the quality of tele-operation [25]. Examples of applications in tele-operation include tele-medicine and tele-surgery [26], tele-rehabilitation [27], [28], tele-operation for space exploration [29], micromanipulation and microassembly [30], and tele-operation of underwater [31] or aerial [29] robots.

To our best knowledge, currently, there are not so many studies available that address the communication of sensorimotor skills over a computer network. Our research employs a pragmatic methodology to provide a comprehensive literature review on sensorimotor skill communication studies in academic journals and conference proceedings. Our main contributions are in the following areas:

- We dived deeper into neuroscience literature to examine the neural representation of the human sensorimotor skill. A visual representation of brain areas involved in sensorimotor skill development/execution as well as functional connectivity/relationships between these regions are shown in Figure 3.
- We conceptualized sensorimotor skill communication systems as a composition of three functional elements: skill modeling, skill simulation, and skill evaluation. These elements must be able to support multimodal interaction, perform either online or offline and are tunable for gross and fine motor skills. Furthermore, we propose a taxonomy, based on the presented conceptualization, of methods that are applicable to sensorimotor skill communication. We also present a case study to

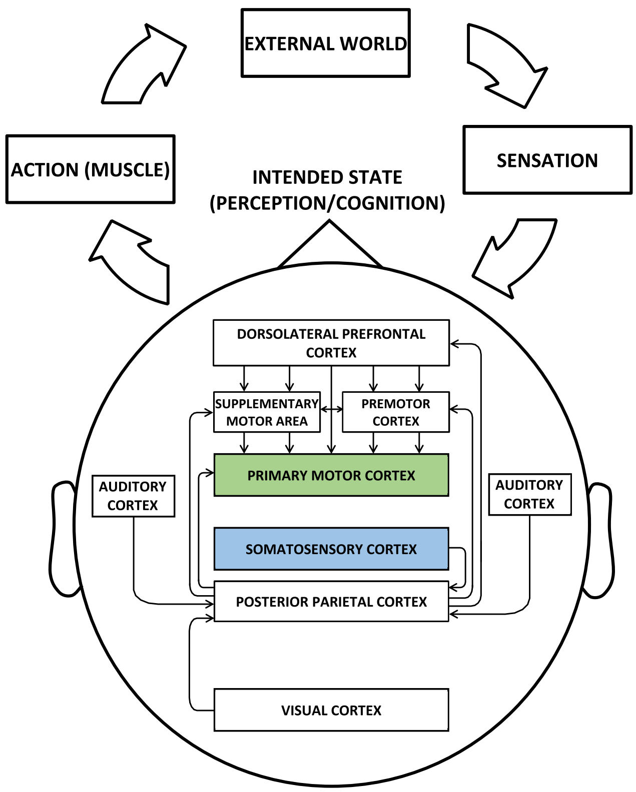

FIGURE 3. Sensorimotor skill communication system from neurobiological perspective.

TABLE 1. **Summary of the main findings: evaluation of approaches for SSC systems.**

Type	Approach	Evaluation metrics						References
		Generalizability	Explainability	Multimodality	Data Hunger	Complexity	Quality	
Modeling	Imitation-based	Low	Low	Medium	Low	Low	Low	[129] [158] [159] [160] [94]
	Activity-based	Medium	Medium	Medium	Medium	Medium	Medium	[69] [161] [1] [85] [73]
	Goal-based	Medium	Low	Low	Low	High	Low	[162] [138] [163] [70]
Simulation	Imitation-based	Low	Medium	Medium	Low	Low	Medium	[164] [100] [165]
	Activity-based	Medium	Low	Medium	High	Low	Medium	[166] [104] [24] [22]
	Goal-based	Low	Low	Medium	High	Low	Medium	[12] [119] [120] [121]
Evaluation	Metrics-based	High	Medium	High	Low	Medium	Medium	[142] [143] [144] [145].
	Algorithm-based	Medium	Low	High	High	High	Medium	[105] [106] [149]
	Implicit score-based	Low	High	Low	Low	Low	Medium	[103] [112] [125]

demonstrate the modeling of a dental sensorimotor skill, namely, periodontal probing.

- Finally, we proposed evaluation metrics to measure the quality of a sensorimotor skill communication system. The evaluation metrics comprise 6 criteria: Generalization, Explainability, Multimodality, Data hunger, Complexity, and Quality of communication. An evaluation of existing literature against the evaluation metrics is shown in Table 1.

II. SENSORIMOTOR SKILL COMMUNICATION SYSTEM

A. HUMAN SENSORIMOTOR SKILL COMMUNICATION SYSTEM

Prior to developing a sensorimotor skill communication system, it is essential to understand the neurobiological foundations of the processes governing the human sensorimotor interaction. In general, a stimulus from the external environment is perceived via human senses, e.g. vision, audition, cutaneous, proprioception, etc., and then processed in the sensorimotor system (see Figure 3) which is a complex component of the human motor control system, incorporating the sensory, motor, and central integration and processing components involved in bodily movements [32]. Based on the perceived information, the sensorimotor system generates the intent to perform a movement. This movement intent is transferred through the central nervous system to activate the selected muscles. The motion is generated as a result of coordinated contractions and relaxations of the activated muscles.

A detailed description of the human sensorimotor system can be found in [33]. At the top of the hierarchy there is the sensorimotor association cortex, which consists of posterior parietal and dorsolateral prefrontal cortices (see Figure 3), that have quite a complex structure and responsible for different functions [34], [35]. The posterior parietal association cortex receives inputs from sensory systems responsible for localizing body parts and external objects (the visual, auditory, and the somatosensory systems) [36]. Neural populations in the posterior parietal cortex group into small clusters, specializing on specific muscular activity in eyes, head, hands, and arms [37], [38]. Before initiating the movement, the posterior parietal cortex integrates the original positions of the body parts that are to be moved and the positions of those surrounding objects. The posterior parietal is also responsible for directing behavior with spatial information and directing attention [33], [39]–[41].

The sensory inputs processed in the posterior parietal association cortex are redirected to the motor cortex, dorsolateral prefrontal association cortex, and some parts of the secondary motor cortex. In its turn, the dorsolateral prefrontal association cortex transmits the processed information to the areas in the secondary motor cortex, to the frontal eye field, and to the primary motor cortex (Figure 3). The latter is responsible for controlling the movements and has a somatopic organization, i.e. each part of the body is represented in the primary motor cortex, most of which is dedicated to the body parts, capable of generating complex movements, e.g. arms and mouth [33]. The neurons of the primary motor cortex greatly contribute to initiating movements [33]. At the same time, it is suggested that the decision to initiate movement originates in the dorsolateral prefrontal cortex [42], [43], depending also on the interaction with posterior parietal cortex and other areas of frontal cortex [33], [44].

The information from posterior parietal and dorsolateral prefrontal cortices is sent to the secondary motor cortex, which redirects most of it to the primary motor cortex. Before the initiation of voluntary movement, the activation of the neurons in the secondary motor cortex is recorded and this activity is sustained during the movement. The experimental evidence suggests that the secondary motor cortex is responsible for programming the movement after receiving instructions from the dorsolateral prefrontal cortex [33], [45]–[47].

B. ARTIFICIAL SENSORIMOTOR SKILL COMMUNICATION SYSTEM

A sensorimotor skill communication system performs modeling, simulation, and evaluation of a given sensorimotor skill (Figure 4). Modeling aims to understand the skill by developing reliable policies and defining a corresponding response for a particular sensory input. Generally speaking, the model of a sensorimotor skill is a system that accepts sensory inputs and corresponding sensory outputs and extracts some knowledge that defines how the motor commands are generated.

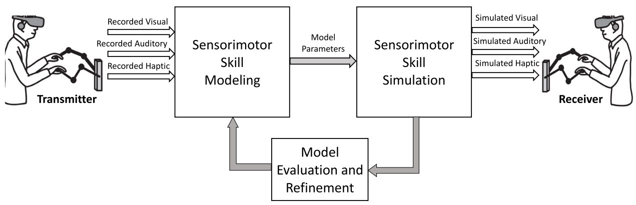

FIGURE 4. **Functional architecture of a sensorimotor skills communication system.**

Sensorimotor skill simulation determines how to respond to sensory inputs to produce a meaningful interaction with the environment. In other words, a sensorimotor skill simulator is a system that accepts sensory inputs and outputs the motor commands. The knowledge of sensorimotor skill is inherently encoded in the simulator. The sensorimotor skill simulator can be used for model-based generation of motor commands, record-playback, or control by modifying the parameters of the system. It can also be directly related to the sensorimotor skill model.

Having a model or simulation of the sensorimotor skill necessitates evaluating it. The model evaluation takes place during the sensorimotor skill simulation and estimates the way the skill is communicated to the environment and how the user benefits from it. That is, we evaluate how far the knowledge, extracted by the model, is relevant to the skill, i.e. the quality of knowledge. Simulation evaluation qualifies to what extent the simulator produces the same skill experience, i.e. the quality of the simulation. One way to evaluate simulation, for example, is to record playback and compute the error between recorded and desired trajectories.

A modeled sensorimotor skill does not have to be simulated and vice versa. For example, a sensorimotor skill model without simulation is used for understanding the skill itself, i.e. learning what are the features of the skill, what are its policies, and rules. Sensorimotor skill models without simulation are implemented for teaching the skill by using other means of communication, for example through audio-visual instructions. A simulation without modeling is taking place when someone just blindly follows the guiding stimulus.

There are parallels between representations of sensorimotor skill communication system from the engineering and neurobiological perspectives. In humans, the signals from the external world are perceived through senses and processed by the Central Nervous System (CNS) to generate movements. The robots perceive the external environment via sensors and this information is processed by the model that sends the commands to motors and actuators in a similar fashion the human sensorimotor system triggers the movement intent that activates contraction of the selected muscles. The recent developments in haptics and infrared thermography [48], [49] made it possible to perceive the physical properties of the external environment, thus making the analogy with the human sensorimotor system more extensive.

From a functional perspective, a sensorimotor skill communication system comprises three sub-systems: Sensorimotor Skills Modeling (SSM), Sensorimotor Skills Simulation (SSS), and Model Evaluation and Refinement (MER) (see Figure 4). The recorded multi-modal (visual, auditory, and haptic) media feed into the SSM module in order to train a model that captures the particular features of a sensorimotor skill. The output of the SSM module is a set of parameters and policies/rules that determine the motor command in response to sensory stimuli. The SSM module shall strive to capture the gross/fine sensorimotor skills in order to provide a highly customizable and personalized communication system. The model parameters are communicated to the SSS in order to generate multi-modal (visual, auditory, or haptic) streams aiming to reproduce the sensorimotor skill experience for a learner to acquire. The output of the SSS module is a simulated (rendered) set of the multi-modal media stream that would be displayed to a learner via appropriate interfaces (visual display, speaker, and a haptic device) with an intention to simulate the specific skill of the expert. Meanwhile, feedback from the SSS module is sent back to the MER module in order to validate, evaluate, and refine the sensorimotor skill model parameters in an effort to improve the communication performance.

III. SENSORIMOTOR SKILL MODELING

Sensorimotor skill modeling aims to capture and understand the skill. A human skill can be considered as a mathematical correspondence between stimuli and responses [1]. In general, a skill can be modeled as a non-linear, non-deterministic, time-variant, generalizable, and decomposable control system, that provides mapping between stimuli and responses [1]. This formulation corresponds to a trajectories encoding approach to modeling the skill in imitation learning, which is a low-level representation of the skill as a mapping between the sensory and motor information [50]. In contrast, the high-level symbolic encoding approach to skill modeling aims to represent a skill as a sequence of action-perception units [50]. From the reinforcement learning perspective, the learning of movement is finding for a given moment of time a task-specific mapping, called policy, between sensory inputs for a given time point (and past time points), and commands sent to the actuators outputs [51], [52]. In this sense, the sensorimotor skill is equivalent to a policy that a robot adapted to perform a specific task. In addition, there is a close relationship between modeling and communicating the skill, that is to say, the clear model (representation) of the skill facilitates its acquisition (learning) [1]. Since the human performance on a repeated task is a stochastic process, in order to acquire (learn) a particular skill one needs to extract from multiple recording of the task some general patterns and characteristics pertaining to the skill [1]. Three main approaches to sensorimotor skill modeling are presented in the following subsections.

A. PRIMITIVE-BASED MODELING

One approach to sensorimotor skill modeling considers a task as a succession of movement primitives, generally speaking, a set of actions leading to the accomplishment of a complete goal-directed behavior [50], [53]–[55]. More precisely, movement primitives are parametrized representations of elementary movements, called policies, that allow each sensorimotor skill to be described as a small set of parameters that can be tuned or learned [56]. This approach leads to learning complexity reduction in multidimensional systems [52]. In sensorimotor skill models a skill is represented as a union of movement templates, i.e. as a combination, arrangement and generalization of a sequence of elemental motions [57]. Each sensorimotor skill, associated with a motor primitive can be learned [51], [52], [58]–[61] and generalized to new situations [61]–[64] before combining them into more complex tasks [57].

Another approach to producing real-time human-like kinematics from a combination of movement primitives obtained by capturing and processing human motions is to use evolutionary algorithms, e.g. for generating arm motions of a humanoid robot [65]. Also, the movement primitives are used in highly dimensional systems for evaluating the policies learned from demonstration [52], [62]. From the neurobiological perspective, motor primitives are closely associated with motor pattern generators (MPGs) [66], [67] – the complete motor circuits including sensory feedback, central pattern generators (CPGs) (circuits that do not require sensory feedback to produce motor activities) and modulations from descending pathways through which the motor signals travel from the brain to lower motor neurons [68]. This approach is used for imitation learning in hierarchical distributed motor control systems, which allows to simplify the perception of a demonstrated movement and facilitate the selection and execution of an optimal action, e.g. [69]–[71].

B. STATE-BASED MODELING

Another approach is to model the sensorimotor skill by considering the task as a set of continuous states (spaces), i.e. encoding the imitation at the trajectory level [72]. The trajectory can be encoded with the Hidden Markov Models (HMMs) and then reproduced by a stochastic algorithm from the transition probabilities [72]. The HMM consists of two stochastic processes – an invisible process of hidden states and a visible process of observable symbols containing a hidden stochastic process [1]. Therefore, HMM is a great tool for modeling stochastic human performance. The hidden process in HMM model of the skill corresponds to an intention, and the produced sequence of observations corresponds to the intended action [1]. Since HMM is a parametric model, the learning of a particular skill is analogous to the optimization of the model's parameters [1]. HMM also accepts different inputs regardless of their modality, thus modeling multimodal sensory inputs.

The probabilistic models allow to analyze the hidden states in the sensorimotor task during imitation learning and correlate them with the corresponding human skill [73]. For example, a robot, based on the probabilistically encoded correlations between the perceived forces and the task's parameters, can decide on which behavior to adapt while reproducing the demonstrated skill [74]. In imitation learning, probabilistic methods consider task as a whole entity (task-level imitation) and are commonly used for extracting common features from multiple demonstrations [75]. Probabilistic inference with Bayesian networks simplifies the learning from demonstration by combining prior kinematic information from a human demonstrator with prior dynamic information and extracts stable motions regardless of the robot's features or properties of the environment [76]. In general, the probabilistic model, trained by an expert returns fewer repeated states, since the expert demonstrates the motion with sufficient variations in each state. The expert is also interested to keep the transitions between different states as similar as possible during the demonstration [73]. Thus, the probabilistic approach to skill modeling allows to distinguish between the skilled and the non-expert demonstrators and evaluate the performance of the latter [73].

To achieve human-like performance in the sensorimotor task it is important first to develop a model of the given skill, focusing on the skill-based performance, and then communicate it (to a robot or a learner). Every sensorimotor task contains uncertainties, which can be represented by the probabilistic models such as Gaussian Mixture Regression (GMR) [77], Gaussian Mixture Model (GMM) [78] and the Hidden Markov Model (HMM) [1], [73], [74], [79]–[85]. This approach results in a sequence of interpretable states, which allows following a generalized trajectory [73], i.e. to perform imitation learning at task level [85]. It is also possible to extract the human skill from the training process in imitation learning [73] and transfer it to a robot or a learner.

C. NEUROLOGY-BASED MODELING

From a neurobiological perspective, the sensorimotor skill communication system is modeled as a composition of mirror neurons [50], [86]. The mirror neurons are located in the premotor and parietal cortices of the human brain and are activated when a person performs a goal-oriented hand movement while observing a similar movement performed by another individual [33], [50], [87], [88]. In fact, the mirror neurons respond not to the specific features, characterizing an action but to the purpose of the action [89].

Current neurobiological models of imitation learning correlate the imitation with the activity of the mirror neurons

[90]–[92] and actively used in learning by demonstration [50], [69], [70], [93]–[95]. For example, in [69] the process of learning from demonstration is inspired by the activation of mirror neurons in human brains. This activation is achieved by the parallel arrangement of multiple pairs of inverse and forward models. In each pair, the inverse model, given a current state and target goal, generates the motor commands. The generated motor commands are sent along with the current state to the forward model to predict the next state. All predicted states from each pair of inverse-forward models are compared to the target goal and the most compatible ones are reinforced, while the others decrease the confidence of the corresponding behavior. This architecture allows both to generate optimal motor commands for achieving the target goal and to recognize which actions have been demonstrated. A detailed review of mirror systems for action recognition and imitation can be found in [96]. Understanding the neurobiological processes underlying the mirror neurons is crucial for facilitating communication of sensorimotor skills [33].

A relatively new area, neurorobotics, focuses on the computational models of sensorimotor skills and complex behaviors, emulating the neurobiological processes in the nervous systems of humans and animals. Apart from developing robots which are manipulated by models, adapting the principles of natural neural computation, neurorobotics contributes to the study of the functioning of complex biophysical systems as the human brain [97]. One approach to emulate the sensorimotor loop of the human nervous system utilizes recurrent neural networks with parametric bias (RNNPB) [98]–[100] to encode sensorimotor trajectories of a humanoid robot, trained to handle balls and blocks by a demonstration from a human teacher [100]. This approach models the neuronal mechanisms responsible for adapting the robot's behavior to different scenarios.

IV. SENSORIMOTOR SKILL SIMULATION

Sensorimotor skill simulation aims to reproduce the skill. Three main categories of sensorimotor skills simulation are presented here.

A. IMITATION-BASED SIMULATION (RECORD-PLAYBACK)

The idea of human-to-human skill transfer using the haptic and visual playback was first introduced in [101]. The proposed WYSIWYF (What You See Is What You Feel) concept ensures a correct visual/haptic registration in record-playback strategy when an expert performs his/her sensorimotor skill in the virtual environment while the system records all available data for further simulation of the skill [101]. Currently, a large number of studies focus on record-playback-based simulation. For example, haptic-based engineering solutions for sensorimotor skill communication rely on haptic playback, which is the ability to reproduce the force or position trajectories of a particular sensorimotor skill, pre-recorded by an expert with the help of a tracking device and force sensors [102]. However, the accuracy of tracking in the haptic playback system does not signify increased effectiveness of sensorimotor skill communication, indicating the presence of other factors, affecting the successful sensorimotor skill transfer [102]. Also, some of the haptic playback techniques are depersonalized, i.e. they overlook the differences in user-specific dynamics and prevent deviating from expert's strategy [103]. Advanced approaches to haptic playback involve a progressive scheme adaptable to the user's performance. Under this scheme, the haptic playback decreases or grows depending on whether the user's performance increases or decreases. The most effective acquisition of sensorimotor skills with haptic guidance happens at the early stage of the task with progressive removal of haptic guidance at more advanced stages of the task [103].

The haptic playback is widely used in haptic-based assistive systems for teaching handwriting [47], [104]–[114]. In these systems, the pre-recorded handwriting sensorimotor skill is communicated to a learner via a haptic interface that reproduces a pre-recorded force feedback to guide the user's hand along a predefined trajectory [104]. However, there exist bi-directional haptic devices offering a real-time (online) haptic guidance, controlled remotely by an expert [23], [24]. The successful achievement of learning objectives with haptic-based handwriting systems depends on the playback method, for example, novices benefit greatly from partial haptic playback, which signals the deviation from the desired trajectory, facilitating the acquisition of general motor skills [108]. However, the advanced students honed their fine sensorimotor writing skills with the help of a full haptic playback. Also, several studies report the positive effects of implementing the haptic-based handwriting assistive systems in occupational therapy to facilitate the integration and acquisition of sensorimotor skills for handwriting in stroke patients [115] or in children with different learning difficulties [116]–[118].

Haptic playback techniques combined with virtual reality (VR) systems are widely implemented in dental simulations, particularly for training periodontal procedures [12], [119]–[123]. These simulators are designed to teach students how to perform periodontal probing and treatment. In general, simulators display the 3D model of the human mouth along with the haptic feedback to imitate real tactile sensations while probing the teeth, gingiva, and calculi with virtual dental tools [121]. These systems are characterized by high fidelity of stimulation, full immersion in the environment, and the ability to standardize the learning process. The periodontal simulators are also capable of displaying different gingival/health scenarios, record the student's performance, provide feedback, and replay those cases that the student had difficulties with. For instance, the Haptodont periodontal simulator consists of bi-manual haptic interaction integrated with a virtual environment, providing simulation of the custom grip with both the dental probe and the mirror instruments. Periodontal simulators provide a safe and customizable environment for the communication of sensorimotor skills for medical students before real-life clinical applications.

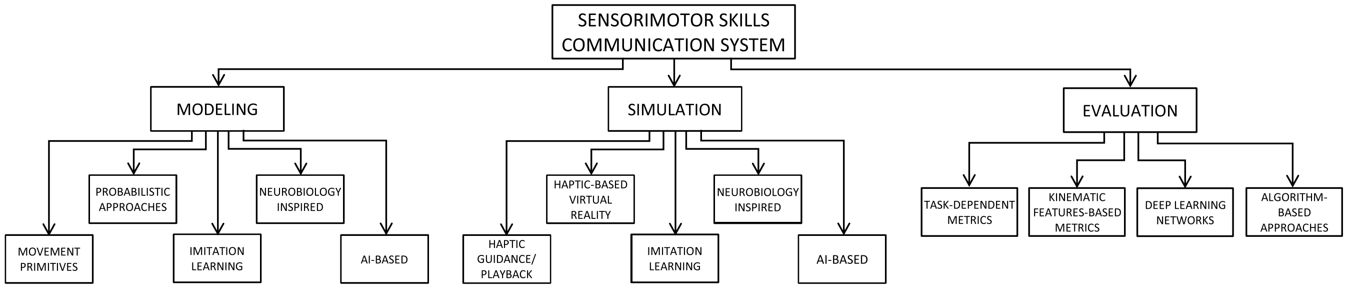

FIGURE 5. The proposed taxonomy of methods for sensorimotor skill communication system.

Other examples of Haptic-based virtual reality (VR) systems for sensorimotor skill communication can be found in medical training and rehabilitation [103]. For instance, a haptic interface simulating a Virtual Haptic Back (VHB) is developed for training students to perform a palpatory diagnosis [124]. Similar haptic-based VR systems are deployed in the neuropsychological evaluation of sensorimotor skills and rehabilitation of brain-injured patients [125]. Haptics-enabled tele-rehabilitation systems provide a bi-directional haptic interaction in a virtual environment that allows the therapist to remotely guide the patient by applying nonpassive nonlinear assistive/resistive forces in response to their movements [27], [28]. The incorporation of VR into robotics devices used in conventional therapy also allows to estimate the progress in patients' performance more accurately and thus, to adjust the treatment by generating an appropriate haptic feedback [126]. The tests demonstrate the efficiency of the assistive haptic feedback in comparison with the fixed impedance controller and other conventional devices for rehabilitation therapy that provide a fixed force playback [126].

B. CONTROL-BASED SIMULATION (SHARED CONTROL/GUIDANCE)

The initial record-playback approach in WYSIWYF (What You See Is What You Feel) display for sensorimotor skill communication [101] has been extended to a more general mechano-media concept, i.e. using robotic mechanisms as media for transferring the motion/kinesthetic information from an expert to a learner [127]. In contrast to the tele-operation system where an expert controls a remote robot, human-human haptic interaction allows bi-directional haptic interaction in a virtual environment [128]. Thus the human-human haptic interaction (mechano-media) can be considered as an extension of the tele-operation framework [127], [128] since it allows to exchange the sensorimotor skill between participating individuals via a haptic device, and forms the foundations of shared-control/guidance systems.

The shared-control/guidance systems are actively used in tele-operation, particularly for surgical training, where the shared control between an expert and a learner facilitates training while performing a robotic surgical task [22]. The Dual-User Tele-operated System (see [22]) adaptively controls the level of involvement of the expert and the learner in the real surgical procedure depending on the latter's expertise and on the expert's recommendations. The system also provides adaptive guidance virtual fixtures that give the expert an additional option to place force cues to guide the learner and allows to conduct the evaluation of the student's expertise based on his/her force profile [22]. The shared guidance scheme is also used to allow experts to follow and correct the learner's movements in real time [23], [24]. Similarly, human-human haptic interaction-based systems can be used for teaching sensorimotor skills for command games, such as playing basketball in a networked haptic basketball game [128]. In summary, it is hard to underestimate the growing popularity of shared-control/guidance systems for human sensorimotor skill communication and acquisition.

Two approaches for haptic guidance are common to support sensorimotor skill development/performance, namely, proactive (or full) guidance, and retroactive (or partial) guidance [108]. In proactive guidance, the haptic interface leads the movement along a desirable trajectory while the user follows the movement. However, in retroactive guidance, the learner is free to move along the desirable trajectory but will experience haptic feedback only when a significant error between the actual and desirable trajectories is observed. Proactive (or full) haptic guidance is described by equation (1) where $\mathbf{F}(t)$ is the calculated force, K_{max} is the stiffness of the haptic interface and $\Delta\mathbf{u}$ is the displacement. Equation (2) outlines how the retroactive (or partial) haptic guidance force is calculated, where C_p, C_i and C_d are the proportional, integral and differential gains respectively. $\mathbf{e}(t)$ is the error between the current position ($\mathbf{x}_{cur}$) and the desired position ($\mathbf{x}_{des}$).

$$\mathbf{F}(t) = K_{max}\Delta\mathbf{u} \tag{1}$$

$$\begin{aligned} \mathbf{F}(t) &= C_p\mathbf{e}(t) + C_i\int_{\Delta T}\mathbf{e}(t)dt + C_d\frac{d\mathbf{e}(t)}{dt} \\ \mathbf{e}(t) &= \mathbf{x}_{cur} - \mathbf{x}_{des} \end{aligned} \tag{2}$$

C. MACHINE-LEARNING-BASED SIMULATION

Machine Learning models have been used to simulate sensorimotor skills. Some implementations of Machine-Learning-based simulations of sensorimotor skills utilize artificial neural networks (ANNs), trained on the examples provided by an expert. To learn the skill, the ANN must be able to construct the skill representation from the training data and extend the initially learned skill through incremental learning [129]. To meet these requirements instead of traditional Multilayer Perceptrons (MLPs) it is advisable to use the local reception fields-based networks such as Radial-Basis Function Networks (RBF) [129], [130]. As universal approximators, RBFs are able to learn from the examples and support incremental learning due to their ability to utilize time-delays for processing spatio-temporal data [129]. The effectiveness of the RBF-based approach was demonstrated with tasks such as peg-into-hole insertion and opening a door [129].

The ANN-based approach is also widely used in embedded artificial sensorimotor skill communication systems, similarly to imitation learning. For instance, a robot with an embedded artificial sensorimotor system can perform simple sensorimotor tasks online (e.g. audio-visual tracking of a person's head). After being trained by an expert, the robot is capable of recognizing the object and updating weights of the embedded network without further supervision, utilizing auditory signals and information about color, luminance, and motion [131].

The sensorimotor data potentially can be simulated with generative models that learn the data distribution from the training set. Later this distribution can be used to better understand the data and to generate the new samples. For example, the Generative Adversarial Networks (GANs) [132] train the competing generative and discriminative models simultaneously. The generative model produces the new data samples following the same distribution as the training data, and the discriminative model determines whether the newly-generated sample is coming from the same distribution. In other words, the objective of a generative model is to maximize the probability of the discriminator making a mistake, while the discriminative model maximizes the probability of assigning the correct label to both training and generated samples [132], [133].

V. SENSORIMOTOR SKILL EVALUATION

To optimize the skill modeling or simulating in sensorimotor skill communication systems, a set of criteria for evaluating how close is the model to the actual skill is desirable. Currently, there is no specific goodness measure for sensorimotor skills. In closely related areas such as imitation learning, the goodness of a model is determined by measuring the performance of the learner (typically a robot) and comparing it to the performance of the expert. For a good model, the learner's performance must be as close as possible to the performance of the expert, i.e. the learner must perform most likely to the expert [1]. In other words, it is necessary to define a metrics of imitation performance that determines the importance of reproduction each of the components of the learned skill, acts as a cost function, and can be optimized [50], [134]. A good metrics of imitation performance quantifies the expert's intents while demonstrating the task and learner's ability for accurate replication of these intentions [50]. The metrics is usually learnt from several demonstrations of the same sensorimotor skill under different conditions and analyzing the variance between demonstrations [135], which allows to extricate features that remain the same over the course of demonstrations [50], [75], [136]–[139]. This approach is also useful for evaluating stochastic models since it takes into account uncertainties in human performance and the environmental noise [1].

The model of a given sensorimotor skill can be implicitly evaluated by measuring the learner's performance after being trained with the system, implementing this model, and comparing it with the expert's performance. In this sense, the skill evaluation score is task-dependent. For instance, in a haptically guided virtual dynamic task [103], the measure of performance is the number of hits of the diagonally placed targets (hit count). In the haptic-based VR system for diagnosis and rehabilitation of patients suffered Traumatic Brain Injury (TBI), the sensorimotor skills are assessed based on the Rey-Osterrieth Complex Figure (ROCF) test and other parameters such as physical performance, drawing time, accuracy, and placement [125].

For the evaluation of handwriting skills acquired using a haptic device, the algorithmic approach based on the local similarity between images is adopted [105], [106]. The fidelity of the written letters was tested by transforming a produced letter into invariant features, that were used to match points between the letters written by the learner and expert. This approach is known as SIFT (Scale Invariant Feature Transform) algorithm [140]. To avoid too many false outliers the extracted matches processed with RANSAC (RANdom SAmple Consensus) algorithm, that finds a transformation which minimizes the sum of squared perpendicular distances, given that the valid points do not deviate by a specified number of units [105], [141]. Thus, the improved algorithmic approach provides an objective, efficient, and more accurate evaluation of the acquired handwriting skill.

In VR-based surgical simulators, multiple performance metrics are defined to assess surgical skills in different scenarios, ranging from some basic kinematic descriptions, including average speed, smoothness (root-mean-square of movement error), and path length to scores specifically defined for linear and circular paths [142]. Metrics as the duration of the period of inactivity (idle time), smoothness, working volume, and force can also serve as indicators of expertise. For example, professional surgeons achieved shorter idle times and smaller working volumes compared to learners [142]–[146]. A similar approach is adopted for evaluating the effectiveness of teaching the handwriting skill by considering its kinematic parameters, for example, average velocity, number of velocity peaks, and number of breaks while performing the task. A successful acquisition of skill results in a decrease of the motion duration, number of breaks, and the number of velocity peaks [112]. Thus, the metrics-based descriptive analysis provides a good insight into the expert's skill.

The main drawback of the metric-based evaluation approach is that they require manual feature engineering. To avoid this limitation, machine learning techniques are recently utilized to evaluate sensorimotor skills [147], [148]. For example, deep learning networks can be applied for surgical skill assessment [149]. Due to their hierarchically organized layers, deep learning networks are capable of inferring intrinsic patterns in motion, learning latent features from the multivariate motion data, and discovering abstract representations of the skill [149]. Another example involves the application of deep learning networks to assess surgical skills by mapping multivariate motion data into three proficiency levels (novice, intermediate, expert) [149]. The algorithms utilizing deep learning have the potential for sensorimotor skill evaluation, for example, by establishing a regression model to quantitatively evaluate the quality of the sensorimotor skill.

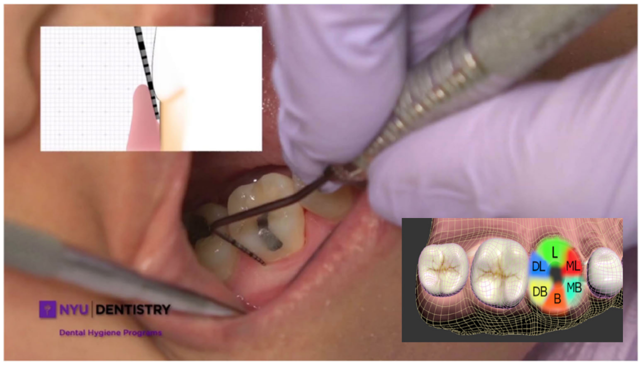

FIGURE 6. Probing technique with correct probe orientation and grip pose. Different tooth regions are color-coded: ML, L, DL, MB, B, and DB correspond to mesial lingual, lingual, distal lingual, mesial buccal, buccal, and distal buccal regions (with permission from the College of Dentistry of New York University).

VI. CASE STUDY

The case study aims to demonstrate how a periodontal probing, being a sensorimotor skill, can be modeled/simulated using a haptic-based dental simulation system called Haptodont.

A. PERIODONTAL PROBING

Periodontal probing is a process of diagnosing a periodontal disease caused by a bacterial infection of the tissue around the teeth (periodontium). Due to the bacterial activity, a sticky colorless 'plaque' gradually surrounds the teeth resulting in gingival inflammation of the periodontium [150]. Calcifying, this 'plaque' forms calculus which porous structure attracts more bacteria, thus aggravating the development of the periodontal disease. If untreated, the periodontal disease leads to the formation of periodontal pockets in the gingival sulcus (space between the gingiva and the tooth surface).

Diagnosing and treatment of periodontal disease involve several steps and procedures as detecting periodontal pockets, measuring their depth, locating and removing calculus. All of these procedures heavily rely on tactile feedback obtained from the dental tools while probing [151]. The presence of periodontal pockets is detected by a procedure called periodontal probing, involving the insertion of a periodontal probe in the gingival sulcus and measuring the depth from the gingival margin to the bottom of the pocket or sulcus (Figure 6). The probe's first circumferential marking above the gingiva gives an estimate of a sulcular depth, which is about 2-3 mm for the healthy gingival sulcus. Depths deeper than 3 mm signal the development of periodontitis. Also for depths greater than 3 mm it is important to distinguish between the gingival pocket (no bone loss) and periodontal pocket (when bone loss is present). The periodontal probing task involves probing of all six regions of each tooth for pockets, namely the mesial lingual, lingual, distal lingual, mesial buccal, buccal, and distal regions (see Figure 6). While probing, the maintenance of the probe's proper angulation and adaptation (orientation) is crucial for reaching the bottom of the pocket or sulcus. The probe can also be rotated to follow the anatomy of the tooth so that its tip remains touching the base of the gingival sulcus or pocket while moving along the tooth circumference. Usually, to improve the precision of the probing procedure, a finger rest placed 1-4 teeth away from the probing tooth is used.

B. EXPERIMENTAL SETUP

We deploy the Haptodont periodontal simulator [12] to record the haptics data for the periodontal probing task. The main setup (Figure 7) consists of two 3D Geomagic Touch haptic devices simulating the probe and mirror interactions using the custom grips made from real dental instruments to enhance the tactile experience of grasping the tools, a Novint Falcon device to provide a finger support for the fulcrum, and a VR headset (Oculus Rift) to provide an immersive visual experience.

The Haptodont simulation system makes a perfect test-bed for capturing and evaluating the sensorimotor skill of periodontal training. For the case study we use a VR headset (Oculus Rift) to visualize a model and only one 3D Systems Geomagic Touch haptic device for simulating the probe instrument. We perform the periodontal probing task on a mandibular low jaw 3D model with rendered teeth and gingiva. In consultation with domain experts (dental professors from NYU College of Dentistry and UoT School of Dentistry), we place markers labeling different stages of the periodontal task. To ensure the continuity of the skill, we integrated a Microsoft speech recognition system (Windows Desktop Speech [152]), accepting two commands – "START" and "STOP" to label the beginning and the end of the current stage of the probing task. We use Chai3D C++ framework [153] to render the haptic force feedback by attaching to the tool a probe model with a single point haptic interaction [154]. The synchronized visual and haptic rendering in Chai3D provides a realistic periodontal probing experience which is important for developing an accurate model of the periodontal probing skill.

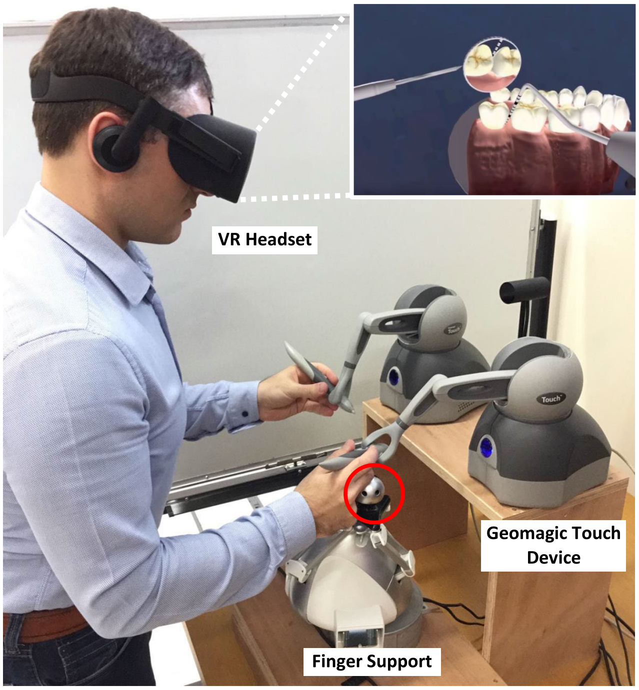

FIGURE 7. Experimental setup of the Haptodont system.

C. DATA

The recorded haptic data include time in seconds, the position and rotation of the haptic interaction point along with the linear and angular velocity components, and the exerted forces. We also record the position and rotation matrices of the dental (teeth-gingiva) model for reconstructing and playing back the tool tip trajectory in reference to the position/orientation of the dental model. These data are used to infer the angulation of the tool tip. The Workspace Scale Factor, which is the scale factor between the physical workspace of the haptic device and the virtual workspace defined for the tool is also included in the haptics data. Finally, we label the trajectory corresponding to the skill of interest (0 – no skill recording, 1 – skill is recorded) with the voice commands interpreted by the speech recognition module. These labels are necessary for segregating the skill path from the other recorded trajectories.

D. PRELIMINARY RESULTS

We tested the ability of the system to distinguish the stages of the periodontal probing task. Particularly, we asked the dental professional to perform a probing task on three teeth from each side (32, 31, 30 from the lower right buccal and 17, 18, 19 from the lower left lingual). For testing purposes, instead of probing each of the six regions (see Figure 6), the dental specialist probed only the lower right buccal side of the jaw (facing the cheeks), including the mesial buccal, buccal and distal buccal regions and the lower left lingual side (facing the tongue) – mesial lingual, lingual and distal lingual regions of each tooth. Thus, for each tooth, the probing task includes moving around the circumference of the tooth spanning these three regions and checking for pockets. The sequence of "START" and "STOP" commands clearly labels ends of the probing task paths for each tooth, ensuring the continuity of the task trajectory. Later, the paths of each tooth were partitioned into three regions with the unsupervised learning algorithm (k-means) and the boundaries of partitions were manually adjusted in consultation with dental professionals. In total, three recordings of each side were collected, each one containing paths around 3 teeth with 3 regions per tooth on the corresponding side.

To demonstrate the ability for spatial segmentation of the periodontal probing task trajectories according to the probed regions, univariate feature selection with Mutual Information (MI) on all three recordings is performed to determine the set of significant features. Out of an initial list of 24 features, we selected the top 8 most significant ones. MI estimates how much knowing the value of one variable reduces the uncertainty on the other [155]. It is invariant to the data transformations and can capture any kind of dependency between variables and targets, including nonlinear relationships [156]. According to the MI, the positional coordinates x, y of the periodontal probe, the probe rotation matrix components $R_{1,1}, R_{1,3}, R_{2,2}, R_{3,1}, R_{3,2}$ and $R_{3,3}$ are the top significant features. To validate the performance of SVM classifier with the optimal value of the inverse of the standard deviation of the RBF kernel $\gamma = 0.01$ and penalty term $C = 100$, we iterated over three recordings, consecutively training the SVM classifier on two recordings and testing it on the remaining one.

The SVM classifier was able to achieve high accuracy of spatial segmentation of the probe path (see average normalized confusion matrices on Figure 8 for buccal a.) and lingual b.) sides). The average accuracy was 0.86 and 0.78 for buccal and lingual regions, respectively. Precision and recall values averaged over three iterations were 0.88 and 0.85 for the buccal side correspondingly. For the lingual side, the average precision and recall values were 0.82 and 0.78, respectively. The recorded path, as well as the predicted path generated by the trained SVM model, are presented in Figure 9. The model is capable to detect which region is probed with very high accuracy. The identified regions are always indistinguishable from the actual regions on both sides (Figure 9). It is worth noting that the incorrectly identified points are found on the parts of an "approaching" path laying far outside the space between the gingiva and teeth, and also near the boundary areas of regions.

The model can be extended to identify if there are pockets in the probed regions and to estimate the depth of the probed pocket. It is also important to investigate whether the model, trained on one specialist's recordings will be capable of inferring which region is being probed from another specialist's/student's recordings. Furthermore, the model can be utilized to distinguish between expert and novice users. For this purpose, additional features may be incorporated, such as force projections and linear/angular velocities.

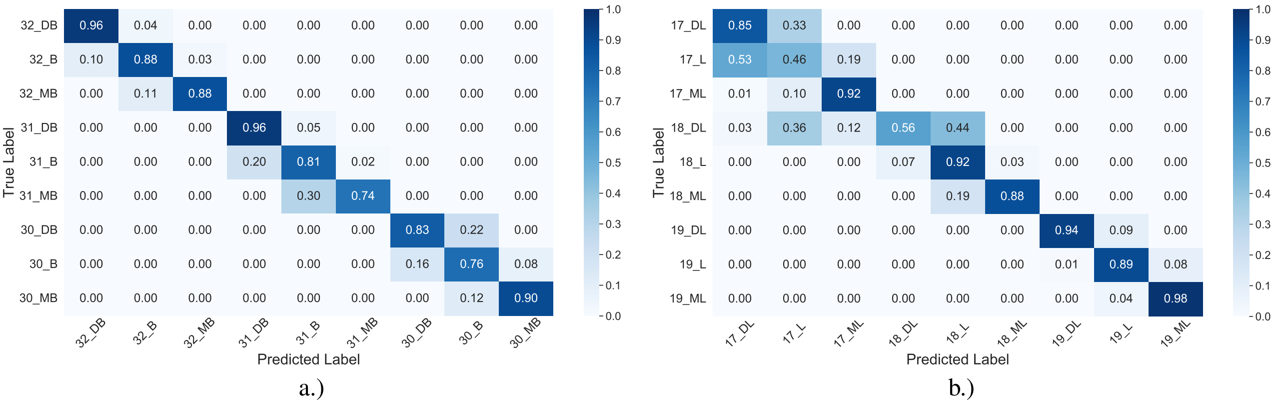

FIGURE 8. Average of normalized confusion matrices of SVM classifier for predicting probing regions, trained on the first two recordings and tested on the third one: a.) lower right buccal side b.) lower left lingual side.

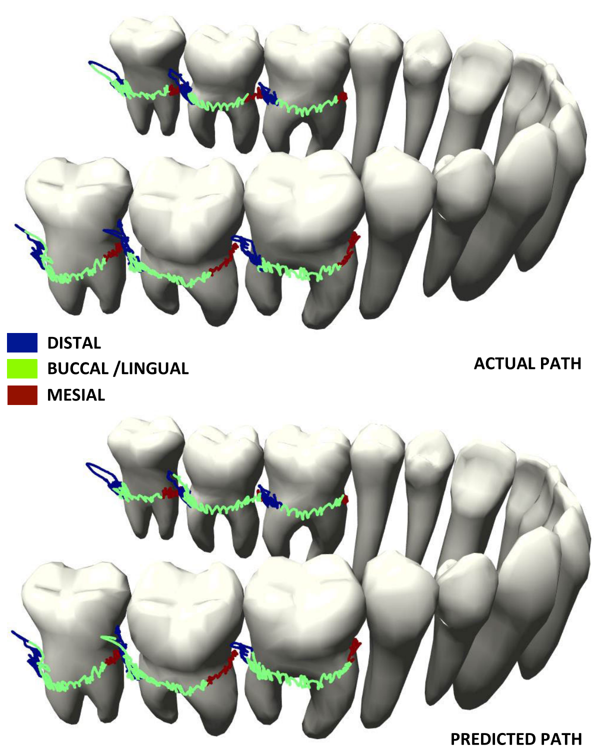

FIGURE 9. Recorded and predicted probe trajectories (upper and lower panels). The model was trained with 8 features on the first and second recordings and the regions are accurately identified for the third recording. The gingiva model is removed to clearly visualize the probe trajectory.

VII. CONCLUSION

A. SUMMARY OF FINDINGS

Our main findings are summarized below:

- A representation of brain areas and interactions between these areas during sensorimotor skill communication is shown in Figure 3. Sensory information is routed from the visual, auditory, and somatosensory cortices into the posterior parietal cortex, which in turn routes relevant data to the dorsolateral prefrontal cortex, the supplementary motor area, the premotor cortex, and the primary motor cortex. The primary motor cortex is tasked with providing motor commands to generate a response.
- A complete sensorimotor skill communication system is yet to be fully developed. A complete sensorimotor skill communication system should be able to process multimodal sensory information, model the sensorimotor skill, decode the modeled skill into a set of parameters and transfer these parameters to a remote user in order to reproduce the skill. Furthermore, most of the currently available systems do not provide a generalized model of sensorimotor skills. The quality of generalization of the modeled sensorimotor skill's decreases with the growing complexity of the task. Finally, most existing literature covers off-line sensorimotor skill communication (non-realtime).
- A sensorimotor skill communication system has three functional elements: skill modeling, skill simulation, and skill evaluation. These elements must be able to support multimodal interaction, perform either online or offline and are tunable for gross and fine sensorimotor skills. Furthermore, a taxonomy of methods commonly utilized to develop sensorimotor skill communication systems is proposed in Figure 5.
- The quality of a sensorimotor skill communication systems can be evaluated based on six criteria; generalization (whether is the approach extendable to the more general skills or is task-specific or user-specific), explainability (to what extent the approach analyzes and reflects the skill features), multimodality (if the proposed approach supports multiple modalities), data hunger (how much data does the system need to train itself), complexity (how complex the approach is), and quality of communication (how good this approach is in communicating the sensorimotor skill). Comparing the existing literature against the evaluation metrics is summarized in Table 1.
- Properly modeling a sensorimotor skill may not demand ultra high speed, ultra high bandwidth communication infrastructure since it involves the communication of

the model parameters. For instance, a parametrized skill model, that is used for generalization of the skill, can be reproduced in novel situations by changing the parameters of the model [157]. This approach will eliminate the necessity of transferring the pre-recorded skill data over the Internet, thus speeding up the sensorimotor skill communication.

B. CHALLENGES OR LIMITATIONS

Some of the prominent challenges that are derived from this literature review are listed below. Many of these challenges inspire perspectives for future work.

- Multimodal nature of sensorimotor skill: Even though incorporating multimodal features, including visual, auditory, and haptic, has clear benefits, the multimodal nature of sensorimotor skills introduces additional challenges for recording and rendering multimodal skill. For example, actuation systems do not match human motor skill capabilities (including Degrees Of Freedom (DOF), workspace, force, torque, etc.). Future developments in science and technology will make it possible to utilize other modalities resulting in more accurate modeling of sensorimotor skills.
- Online (realtime) Sensorimotor Skill Communication Systems: Most existing literature covers off-line sensorimotor skill communication. Only a few studies explored the real-time (online) modeling/simulation of a sensorimotor skill such as handwriting skills [23], [24] and tele-operation systems [22]. The paucity of online sensorimotor skill communication systems is partially attributed to the high bandwidth requirements (e.g. 1000 Hz for haptic media) for a proper representation of the skill, resulting in colossal amounts of data. The recent emergence of high-speed low latency 5G networks makes it possible to transfer the skill data without delays that will lead to the growing interest in real-time sensorimotor skill communication.
- Modeling environmental conditions: In the process of transferring the skill, the conditions in the remote environment are different and the sensorimotor skill communication system must be able to adapt the communicated skill to the environmental conditions at the receiver side. One way to deal with environmental noise is to apply machine learning techniques (such as reinforcement learning or recurrent neural networks) for tuning and refining the control strategies without human intervention.
- Gross and fine sensorimotor skill communication: A complete sensorimotor skill communication system should be capable of clearly distinguishing between gross and fine sensorimotor skills and model them separately. There are not so many works that clearly discriminate between the gross and fine sensorimotor skills, partially because it requires using different sensory modalities for controlling and different actuators to simulate the gross and fine motions. Visual modality can be used for controlling the gross skills [167], while the force profile collected via haptic/tactile inputs can be used for fine skills representation [73].
- Complete Sensorimotor Skill Communication System: A complete SSC system is yet to be fully developed. The currently available systems do not incorporate simultaneous modeling, simulation, and evaluation of the sensorimotor skill. Among the available research, some focus purely on modeling to better understand the skill, e.g. neurobiology-inspired models of sensorimotor skill [66], [67] while others focus on purely simulation-based systems e.g. for rehabilitation or teaching [104], [117], [166]. The deep learning methods are very promising for implementing future sensorimotor skill communication systems since they do not require the knowledge about the task and potentially, will open an opportunity for creating task-independent sensorimotor skill communication systems.
- Conventional metrics may not be sufficient for evaluating the quality of sensorimotor skill communication. For example, quantitative measures such as RMSE are not reliable to predict the skill quality.
- Most approaches are not able to consistently generalize sensorimotor skills. Generalization quality decreases as the complexity of the task increases.
- Security and Privacy: While a sensorimotor skill communication system provides immense opportunities for the skill transfer, it is prone to data tampering. For instance, if the haptic data is intercepted by a third party, it might be used to infer personal information about the user from their skill patterns (dominant hand, gender, age group, etc.). Furthermore, having the genuine data the attacker can learn the distribution and generate new samples almost indistinguishable from the genuine data, e.g. using Generative Adversarial Networks (GANs) [132] to launch Men-in-the-Middle attacks. Finally, the copyright protection problem is even more challenging once sensorimotor skill communication systems become widely available.

C. TRENDS AND FUTURE WORK

Even though there have been significant efforts to develop systems for modeling, simulating, or evaluating sensorimotor skills, combining all these sub-systems to create a complete sensorimotor skill communication system remains in its infancy. Potential future perspectives include the followings:

- Artificial Intelligence: There is a growing interest in machine learning to build sensorimotor skill communication systems, including modeling, simulation, and evaluation of the skill. For example, reinforcement learning (RL) algorithms can be used for searching the task-specific policy that maps the sensory input about the user behavior and environment into the motor output, where the complex movement is represented as a sequence of simplistic motor primitives. Model-free RL policy search methods open an opportunity of learning a wide range of skills but require multiple demonstrations (data hungry). Model-based RL methods are more efficient in simpler tasks, but for more complex tasks they can lead to learning the sub-optimal policy. Therefore, a combination of model-based and model-free RL methods seems very promising for modeling and simulation of sensorimotor skills. Finally, generative models (e.g. GANs) can be trained on expert data to generate a skill that is indistinguishable from the genuine recordings.

- Haptics for Sensorimotor Skill Communication Systems: Given the crucial role that touch plays in sensorimotor skill development and communication, future work should focus on building haptic technologies to record and playback position, movement, force, torque, and temperature, and techniques to synchronize haptic media with other modalities such as visual and/or auditory.
- Online Sensorimotor Skill Communication Systems: An online sensorimotor skill communication system continuously learns and refines its model to more accurately deliver the sensorimotor skill. It will be imperative to create systems that can learn novel skills or refine already learned ones 'online', which may necessitate human intervention to guide the online learning/modeling process.
- Study Human Learning and Behaviour: Once completed, a sensorimotor skill communication system can be utilized to study how humans develop and communicate sensorimotor skills. It would also be a platform for studying human behavior involving sensorimotor interactions.
- Environment-Aware Sensorimotor Skill Communication Systems: This involves developing systems that accommodate for environmental disturbances or noises during skill communication.

REFERENCES

[1] J. Yang, Y. Xu, and C. S. Chen, "Hidden Markov model approach to skill learning and its application to telerobotics," *IEEE Trans. Robot. Autom.*, vol. 10, no. 5, pp. 621–631, Jun. 1994.

[2] D. Wolpert, J. Diedrichsen, and J. Flanagan, "Principles of sensorimotor learning," *Nature Rev. Neurosci.*, vol. 12, pp. 739–751, Dec. 2011.

[3] C. S. Green and D. Bavelier, "Action video game modifies visual selective attention," *Nature*, vol. 423, no. 6939, pp. 534–537, May 2003.

[4] S. Most, B. Scholl, E. Clifford, and D. Simons, "What you see is what you set: Sustained inattentional blindness and the capture of awareness," *Psychol. Rev.*, vol. 112, pp. 217–242, Feb. 2005.

[5] M. Land and P. McLeod, "From eye movements to actions: How batsmen hit the ball," *Nature Neurosci.*, vol. 3, pp. 1340–1345, Jan. 2001.

[6] M. Land and B. Tatler, "Looking and acting: Vision and eye movements in natural behaviour," *Looking Acting, Vis. Eye Movements Natural Behav.*, vol. 4, pp. 1–288, Jan. 2012.

[7] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[8] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-enabled tactile Internet," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 460–473, Dec. 2016.

[9] *The Tactile Internet, ITU-T Technology Watch Report*, ITU-T, Geneva, Switzerland, 2014.

[10] M. Liyanage, A. Braeken, P. Kumar, and M. Ylianttila, *IoT Security: Adv. Authentication*. Hoboken, NJ, USA: Wiley, 2019.

[11] M. Dohler, T. Mahmoodi, M. A. Lema, M. Condoluci, F. Sardis, K. Antonakoglou, and H. Aghvami, "Internet of skills, where robotics meets ai, 5G and the tactile Internet," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, 2017, pp. 1–5.

[12] G. Karafotias, G. Korres, D. Sefo, P. Boomer, and M. Eid, "Towards a realistic haptic-based dental simulation," in *Proc. IEEE Int. Symp. Haptic, Audio Vis. Environ. Games (HAVE)*, Oct. 2017, pp. 1–6.

[13] H. H. Clark, *Using Language*. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[14] F. Donnarumma, H. Dindo, and G. Pezzulo, "Sensorimotor communication for humans and robots: Improving interactive skills by sending coordination signals," *IEEE Trans. Cognit. Develop. Syst.*, vol. 10, no. 4, pp. 903–917, Dec. 2018.

[15] G. Pezzulo, F. Donnarumma, and H. Dindo, "Human sensorimotor communication: A theory of signaling in online social interactions," *PLoS ONE*, vol. 8, no. 11, Nov. 2013, Art. no. e79876.

[16] C. Vesper, R. P. R. D. van der Wel, G. Knoblich, and N. Sebanz, "Are you ready to jump? Predictive mechanisms in interpersonal coordination," *J. Experim. Psychol., Hum. Perception Perform.*, vol. 39, no. 1, pp. 48–61, 2013.

[17] M. Candidi, A. Curioni, F. Donnarumma, L. M. Sacheli, and G. Pezzulo, "Interactional leader-follower sensorimotor communication strategies during repetitive joint actions," *J. Roy. Soc. Interface*, vol. 12, pp. 1–12, Sep. 2015.

[18] G. Pezzulo, "Shared representations as coordination tools for interaction," *Rev. Philosophy Psychol.*, vol. 2, no. 2, pp. 303–333, Jun. 2011.

[19] F. Leibfried, J. Grau-Moya, and D. A. Braun, "Signaling equilibria in sensorimotor interactions," *Cognition*, vol. 141, pp. 73–86, Aug. 2015.

[20] T. Salter, K. Dautenhahn, and R. T. Boekhorst, "Learning about natural human–robot interaction styles," *Robot. Autom. Syst.*, vol. 54, no. 2, pp. 127–134, Feb. 2006.

[21] A. Hussein, M. M. Gaber, E. Elyan, and C. Jayne, "Imitation learning: A survey of learning methods," *ACM Comput. Surv.*, vol. 50, no. 2, pp. 1–35, Jun. 2017.

[22] M. Shahbazi, S. F. Atashzar, and R. V. Patel, "A dual-user teleoperated system with virtual fixtures for robotic surgical training," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 3639–3644.

[23] N. Pedemonte, T. Laliberte, and C. Gosselin, "A bidirectional haptic device for the training and assessment of handwriting capabilities," in *Proc. World Haptics Conf. (WHC)*, Apr. 2013, pp. 599–604.

[24] N. Pedemonte, T. Laliberté, and C. Gosselin, "Bidirectional haptic communication: Application to the teaching and improvement of handwriting capabilities," *Machines*, vol. 4, no. 1, p. 6, Feb. 2016.

[25] J. G. W. Wildenbeest, D. A. Abbink, C. J. M. Heemskerk, F. C. T. van der Helm, and H. Boessenkool, "The impact of haptic feedback quality on the performance of teleoperated assembly tasks," *IEEE Trans. Haptics*, vol. 6, no. 2, pp. 242–252, Apr. 2013.

[26] B. Weber and C. Eichberger, "The benefits of haptic feedback in telesurgery and other teleoperation systems: A meta-analysis," in *Proc. Int. Conf. Univ. Access Hum.-Comput. Interact.*, 2015, pp. 394–405.

[27] S. F. Atashzar, I. G. Polushin, and R. V. Patel, "Networked teleoperation with non-passive environment: Application to tele-rehabilitation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 5125–5130.

[28] S. F. Atashzar, I. G. Polushin, and R. V. Patel, "A small-gain approach for nonpassive bilateral telerobotic rehabilitation: Stability analysis and controller synthesis," *IEEE Trans. Robot.*, vol. 33, no. 1, pp. 49–66, Feb. 2017.

[29] G. Liu, X. Geng, L. Liu, and Y. Wang, "Haptic based teleoperation with master-slave motion mapping and haptic rendering for space exploration," *Chin. J. Aeronaut.*, vol. 32, no. 3, pp. 723–736, Mar. 2019.

[30] A. Bolopion and S. Regnier, "A review of haptic feedback teleoperation systems for micromanipulation and microassembly," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 3, pp. 496–502, Jul. 2013.

[31] C. D. Pham, C. Spiten, and P. J. From, "A control allocation approach to haptic control of underwater robots," in *Proc. IEEE Int. Workshop Adv. Robot. Social Impacts (ARSO)*, Mar. 2015, pp. 1–6.

[32] B. Riemann and S. Lephart, "The sensorimotor system, Part I: The physiologic basis of functional joint stability," *J. Athletic Training*, vol. 37, pp. 9–71, Feb. 2002.

[33] J. Pinel and S. Barnes, *Introduction to Biopsychology*. London, U.K.: Pearson, 2018.

[34] M. Davare, A. Kraskov, J. Rothwell, and R. Lemon, "Interactions between areas of the cortical grasping network," *Current Opinion Neurobiol.*, vol. 21, pp. 565–570, Jun. 2011.

[35] C. Wilson, D. Gaffan, P. Browning, and M. Baxter, "Functional localization within the prefrontal cortex: Missing the forest for the trees?" *Trends Neurosci.*, vol. 33, pp. 40–533, Dec. 2010.

[36] M. I. Sereno and R.-S. Huang, "Multisensory maps in parietal cortex," *Current Opinion Neurobiol.*, vol. 24, pp. 39–46, Feb. 2014.

[37] K. Man, A. Damasio, K. Meyer, and J. Kaplan, "Convergent and invariant object representations for sight, sound, and touch: Neural convergence of sight, sound, and touch," *Hum. Brain Mapping*, vol. 36, pp. 3629–3640, Jun. 2015.

[38] J. Wang, Y. Yang, L. Fan, J. Xu, C. Li, Y. Liu, P. T. Fox, S. B. Eickhoff, C. Yu, and T. Jiang, "Convergent functional architecture of the superior parietal lobule unraveled with multimodal neuroimaging approaches," *Hum. Brain Mapping*, vol. 36, no. 1, pp. 238–257, Jan. 2015.

[39] J. B. Hutchinson, M. R. Uncapher, K. S. Weiner, D. W. Bressler, M. A. Silver, A. R. Preston, and A. D. Wagner, "Functional heterogeneity in posterior parietal cortex across attention and episodic memory retrieval," *Cerebral Cortex*, vol. 24, no. 1, pp. 49–66, Jan. 2014.

[40] S. Kuang, P. Morel, and A. Gail, "Planning movements in visual and physical space in monkey posterior parietal cortex," *Cerebral Cortex*, vol. 26, pp. 731–747, Jan. 2015.

[41] A. Wilber, B. Clark, T. Forster, M. Tatsuno, and B. Mcnaughton, "Interaction of egocentric and world-centered reference frames in the rat posterior parietal cortex," *J. Neurosci., J. Soc. Neurosci.*, vol. 34, pp. 5431–5446, Apr. 2014.

[42] J. Rigato, M. Murakami, and Z. Mainen, "Spontaneous decisions and free will: Empirical results and philosophical considerations," *Cold Spring Harbor Symposia Quant. Biol.*, vol. 79, pp. 177–184, Dec. 2014.

[43] M. Siegel, T. J. Buschman, and E. K. Miller, "Cortical information flow during flexible sensorimotor decisions," *Science*, vol. 348, no. 6241, pp. 1352–1355, Jun. 2015.

[44] D. Lee, H. Seo, and M. W. Jung, "Neural basis of reinforcement learning and decision making," *Annu. Rev. Neurosci.*, vol. 35, no. 1, pp. 287–308, Jul. 2012.

[45] T. Pearce and D. Moran, "Strategy-dependent encoding of planned arm movements in the dorsal premotor cortex," *Science*, vol. 337, pp. 984–988, Jul. 2012.

[46] M. P. Olshansky, R. J. Bar, M. Fogarty, and J. F. X. DeSouza, "Supplementary motor area and primary auditory cortex activation in an expert break-dancer during the kinesthetic motor imagery of dance to music," *Neurocase*, vol. 21, no. 5, pp. 607–617, Sep. 2015.

[47] C.-H. Park, W. H. Chang, M. Lee, G. H. Kwon, L. Kim, S. T. Kim, and Y.-H. Kim, "Which motor cortical region best predicts imagined movement?" *NeuroImage*, vol. 113, pp. 101–110, Jun. 2015.

[48] T. Aujeszky, G. Korres, and M. Eid, "Thermography-based material classification using machine learning," in *Proc. IEEE Int. Symp. Haptic, Audio Visual Environ. Games*, Oct. 2017, pp. 1–6.

[49] T. Aujeszky, G. Korres, and M. Eid, "Material classification with laser thermography and machine learning," *Quant. Infr. Thermography J.*, vol. 4, pp. 1–22, Dec. 2018.

[50] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, *Robot Programming by Demonstration*. Berlin, Germany: Springer, 2008, pp. 1371–1394.

[51] S. Schaal, P. Mohajerian, and A. Ijspeert, "Dynamics systems vs. optimal control - a unifying view," in *Computing Neuroscience: Theories Insights into Brain Function*, vol. 165. Amsterdam, The Netherlands: Elsevier, 2007, pp. 425–445.

[52] S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert, "Learning movement primitives," *Int. J. Robotic Res.*, vol. 15, pp. 561–572, Jan. 2003.

[53] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, Jun. 2002, pp. 1398–1403.

[54] S. Schaal, "Is imitation learning the route to humanoid robots?" *Trends Cognit. Sci.*, vol. 3, no. 6, pp. 233–242, Jun. 1999.

[55] S. Schaal, *Dyn. Movement Primitives—A Framework for Motor Control Humans Humanoid Robotics* Tokyo, Japan: Springer, 2006, pp. 261–280.

[56] E. A. Ráckert, G. Neumann, M. Toussaint, and W. Maass, "Learned graphical models for probabilistic planning provide a new class of movement primitives," *Frontiers Comput. Neurosci.*, vol. 6, p. 97, Oct. 2013.

[57] J. Kober and J. Peters, *Learing Motor Skills: From Algorithms to Robot Experiments*, vol. 97. Cham, Switzerland: Springer, 2014.

[58] J. Kober and J. Peters, *Learing Prioritized Control Motor Primitives*. Cham, Switzerland: Springer, 2014, pp. 149–160.

[59] F. Stulp, E. Theodorou, M. Kalakrishnan, P. Pastor, L. Righetti, and S. Schaal, "Learning motion primitive goals for robust manipulation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2011, pp. 325–331.

[60] M. Tamosiunaite, B. Nemec, A. Ude, and F. Wörgötter, "Learning to pour with a robot arm combining goal and shape learning for dynamic movement primitives," *Robot. Autom. Syst.*, vol. 59, no. 11, pp. 910–922, Nov. 2011.

[61] J. Kober, E. Oztop, and J. Peters, "Reinforcement learning to adjust robot movements to new situations," *Robot., Sci. Syst. VI*, vol. 4, pp. 1–8, Jan. 2011.

[62] A. Ijspeert, J. Nakanishi, and S. Schaal, "Learning attractor landscapes for learning motor primitives," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 15, Jan. 2002, pp. 1523–1530.

[63] J. Peters and S. Schaal, "Reinforcement learning of motor skills with policy gradients," *Neural Netw.*, vol. 21, no. 4, pp. 682–697, May 2008.

[64] J. Peters and S. Schaal, "Natural actor-critic," *Neurocomputing*, vol. 71, nos. 7–9, pp. 1180–1190, Mar. 2008.

[65] G.-R. Park, K. Kim, C. Kim, M.-H. Jeong, B.-J. You, and S. Ra, "Human-like catching motion of humanoid using evolutionary Algorithm(EA)-based imitation learning," in *Proc. 18th IEEE Int. Symp. Robot Hum. Interact. Commun.*, Sep. 2009, pp. 809–815.

[66] A. I. Selverston, "Are central pattern generators understandable?" *Behav. Brain Sci.*, vol. 3, no. 4, pp. 535–540, Dec. 1980.

[67] U. Bässler, "On the definition of central pattern generator and its sensory control," *Biol. Cybern.*, vol. 54, no. 1, pp. 65–69, May 1986.

[68] E. Marder, "Motor pattern generation," *Current Opinion Neurobiol.*, vol. 10, no. 6, pp. 691–698, Dec. 2000.

[69] Y. Demiris and M. Johnson, "Distributed, predictive perception of actions: A biologically inspired robotics architecture for imitation and learning," *Connection Sci.*, vol. 15, no. 4, pp. 231–243, Dec. 2003.

[70] Y. Demiris and B. Khadhouri, "Hierarchical attentive multiple models for execution and recognition of actions," *Robot. Auto,. Syst.*, vol. 54, no. 5, pp. 361–369, May 2006.

[71] D. M. Wolpert and M. Kawato, "Multiple paired forward and inverse models for motor control," *Neural Netw.*, vol. 11, nos. 7–8, pp. 1317–1329, Oct. 1998.

[72] T. Cederborg, M. Li, A. Baranes, and P.-Y. Oudeyer, "Incremental local online Gaussian mixture regression for imitation learning of multiple tasks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Oct. 2010, pp. 267–274.

[73] Y. C. Zhao, A. Al-Yacoub, Y. M. Goh, L. Justham, N. Lohse, and M. R. Jackson, "Human skill capture: A hidden Markov model of force and torque data in peg-in-a-hole assembly process," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 655–660.

[74] S. Calinon, P. Evrard, E. Gribovskaya, A. Billard, and A. Kheddar, "Learning collaborative manipulation tasks by demonstration using a haptic interface," in *Proc. Int. Conf. Adv. Robot.*, 2009, pp. 1–6.

[75] T. Sato, Y. Genda, H. Kubotera, T. Mori, and T. Harada, "Robot imitation of human motion based on qualitative description from multiple measurement of human and environmental data," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, vol. 3, Nov. 2003, pp. 2377–2384.

[76] D. B. Grimes, R. Chalodhorn, and R. P. N. Rao, "Dynamic imitation in a humanoid robot through nonparametric probabilistic inference," in *Proc. Robot., Sci. Syst. (RSS)*. Cambridge, MA, USA: MIT Press, 2006.

[77] Z. Ghahramani and M. I. Jordan, "Supervised learning from incomplete data via an EM approach," in *Adv. Neural Inf. Process. Syst. 6*, pp. 120–127, Morgan Kaufmann, 1994.

[78] M. Hersch, F. Guenter, S. Calinon, and A. Billard, "Dynamical system modulation for robot learning via kinesthetic demonstrations," *IEEE Trans. Robot.*, vol. 24, no. 6, pp. 1463–1467, Dec. 2008.

[79] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Adv. Robot.*, vol. 25, no. 5, pp. 581–603, Jan. 2011.

[80] J. Yang, Y. Xu, and C. S. Chen, "Human action learning via hidden Markov model," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 27, no. 1, pp. 34–44, Oct. 1997.

[81] Z. Wang, A. Peer, and M. Buss, "An HMM approach to realistic haptic human-robot interaction," in *Proc. 3rd Joint EuroHaptics Conf. Symp. Haptic Interface Virtual Environ. Teleoperator Syst.*, Mar. 2009, pp. 374–379.

[82] G. E. Hovland, P. Sikka, and B. J. McCarragher, "Skill acquisition from human demonstration using a hidden Markov model," in *Proc. IEEE Int. Conf. Robot. Autom.*, Oct. 1996, pp. 2706–2711.

[83] S. K. Tso and K. P. Liu, "Hidden Markov model for intelligent extraction of robot trajectory command from demonstrated trajectories," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Dec. 1996, pp. 294–298.

[84] J. Aleotti and S. Caselli, "Trajectory clustering and stochastic approximation for robot programming by demonstration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Jun. 2005, pp. 1029–1034.

[85] S. Calinon, F. Guenter, and A. Billard, "On learning the statistical representation of a task and generalizing it to various contexts," in *Proc. IEEE Int. Conf. Robot. Autom.*, Sep. 2006, pp. 2978–2983.

[86] L. Cattaneo and G. Rizzolatti, "The mirror neuron system," *Arch. Neurol.*, vol. 66, pp. 557–560, Jun. 2009.

[87] P. F. Ferrari and G. Rizzolatti, "Mirror neuron research: The past and the future," *Phil. Trans. Roy. Soc. B, Biol. Sci.*, vol. 369, no. 1644, Jun. 2014, Art. no. 20130169.

[88] G. Rizzolatti and L. Fogassi, "The mirror mechanism: Recent findings and perspectives," *Phil. Trans. Roy. Soc. B, Biol. Sci.*, vol. 369, no. 1644, Jun. 2014, Art. no. 20130420.

[89] G. Rizzolatti and C. Sinigaglia, "The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations," *Nature Rev. Neurosci.*, vol. 11, pp. 264–274, Mar. 2010.

[90] M. Iacoboni, "Cortical mechanisms of human imitation," *Science*, vol. 286, no. 5449, pp. 2526–2528, Dec. 1999.

[91] G. Rizzolatti, L. Fogassi, and V. Gallese, "Neurophysiological mechanisms underlying the understanding and imitation of action," *Nature Rev. Neurosci.*, vol. 2, pp. 661–670, Oct. 2001.

[92] J. Decety, T. Chaminade, J. Grèzes, and A. Meltzoff, "A PET exploration of the neural mechanisms involved in reciprocal imitation," *NeuroImage*, vol. 15, pp. 265–272, Feb. 2002.

[93] A. Billard, "Learning motor skills by imitation: A biologically inspired robotic model," *Cybern. Syst.*, vol. 32, nos. 1–2, pp. 155–193, Jan. 2001.

[94] A. Billard and M. J. Mataric, "Learning human arm movements by imitation: Evaluation of a biologically inspired connectionist architecture," *Robot. Autom. Syst.*, vol. 37, no. 2, pp. 145–160, 2001.

[95] S. Schaal, A. Ijspeert, and A. Billard, "Computational approaches to motor learning by imitation," *Phil. Trans. Roy. Soc. London B, Biol. Sci.*, vol. 358, pp. 537–547, Apr. 2003.

[96] E. Oztop, M. Kawato, and M. Arbib, "Mirror neurons and imitation: A computationally guided review," *Neural Netw., J. Int. Neural Netw. Soc.*, vol. 19, pp. 254–271, May 2006.

[97] P. van der Smagt, M. A. Arbib, and G. Metta, *Neurorobotics: From Vision to Action*. Cham, Switzerland: Springer, 2016, pp. 2069–2094.

[98] J. Tani, "Learning to generate articulated behavior through the bottom-up and the top-down interaction processes," *Neural Netw.*, vol. 16, no. 1, pp. 11–23, Jan. 2003.

[99] M. Ito and J. Tani, "On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system," *Adapt. Behav.*, vol. 12, no. 2, pp. 93–115, Jun. 2004.

[100] M. Ito, K. Noda, Y. Hoshino, and J. Tani, "Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model," *Neural Netw.*, vol. 19, no. 3, pp. 323–337, Apr. 2006.

[101] Y. Yokokohji, R. L. Hollis, T. Kanade, K. Henmi, and T. Yoshikawa, "Toward machine mediated training of motor skills. skill transfer from human to human via virtual environment," in *Proc. 5th IEEE Int. Workshop Robot Hum. Commun.*, 1996, pp. 32–37.

[102] M. Kolesnikov and M. Zefran, "Haptic playback: Better trajectory tracking during training does not mean more effective motor skill transfer," in *Proc. Haptics, Generating Perceiving Tangible Sensations*, Jul. 2010, pp. 451–456.

[103] J. C. Huegel and M. K. O'Malley, "Progressive haptic and visual guidance for training in a virtual dynamic task," in *Proc. IEEE Haptics Symp.*, Mar. 2010, pp. 343–350.

[104] M. A. Eid, M. Mansour, A. H. El Saddik, and R. Iglesias, "A haptic multimedia handwriting learning system," in *Proc. Int. workshop Educ. Multimedia*, 2007, pp. 103–108.

[105] A. Teranishi, T. Mulumba, G. Karafotias, J. Aljaam, and M. Eid, "Effects of full/partial haptic guidance on handwriting skills development," in *Proc. IEEE World Haptics Conf.*, Jun. 2017, pp. 113–118.

[106] A. Teranishi, G. Korres, W. Park, and M. Eid, "Combining full and partial haptic guidance improves handwriting skills development," *IEEE Trans. Haptics*, vol. 11, no. 4, pp. 509–517, Oct. 2018.

[107] D. Babu, H. Nagano, M. Konyo, and S. Tadokoro, "A new approach for realistic vibrotactile friction feedback for midair writing systems," in *Proc. JSME Annu. Conf. Robot. Mechatronics*, 2016, pp. 12–19.

[108] W. Park, G. Korres, T. Moonesinghe, and M. Eid, "Investigating haptic guidance methods for teaching children handwriting skills," *IEEE Trans. Haptics*, vol. 12, no. 4, pp. 461–469, Oct. 2019.

[109] K. Henmi and T. Yoshikawa, "Virtual lesson and its application to virtual calligraphy system," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 1998, pp. 1275–1280.

[110] B. Bayart, A. Pocheville, and A. Kheddar, "An adaptive haptic guidance software module for I-TOUCH: Example through a handwriting teaching simulation and a 3D maze," in *IEEE Int. Workshop Haptic Audio Vis. Environ. Appl.*, Oct. 2005, p. 6.

[111] M. Xiong, I. Milleville-pennel, C. Dumas, and R. Palluel-Germain, "Comparing haptic and visual training method of learning chinese handwriting with a haptic guidance," *J. Comput.*, vol. 8, no. 7, pp. 1815–1820, Jul. 2013.

[112] R. Palluel-Germain, F. Bara, A. H. D. Boisferon, B. Hennion, P. Gouagout, and E. Gentaz, "A visuo-haptic device–telemaque–increases kindergarten children handwriting acquisition," in *Proc. 2nd Joint EuroHaptics Conf. Symp. Haptic Interface Virtual Environ. Teleoperator Syst. (WHC)*, Mar. 2007, pp. 72–77.

[113] M. M. Boroujeni and A. Meghdari, "Haptic device application in persian calligraphy," in *Proc. Int. Conf. Comput. Autom. Eng.*, Mar. 2009, pp. 160–164.

[114] M. M. Amin, H. B. Zaman, and A. Ahmad, "Visual haptic approach complements learning process of jawi handwriting skills," in *Proc. 5th Int. Conf. Inf. Commun. Technol. Muslim World (ICT4M)*, Mar. 2013, pp. 1–6.

[115] J. Mullins, C. Mawson, and S. Nahavandi, "Haptic handwriting aid for training and rehabilitation," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, Oct. 2005, pp. 2690–2694.

[116] Y. Kim, Z. Duric, N. L. Gerber, A. R. Palsbo, and S. E. Palsbo, "Poster: Teaching letter writing using a programmable haptic device interface for children with handwriting difficulties," in *Proc. IEEE Symp. 3D User Interface*, Mar. 2009, pp. 145–146.

[117] Y.-S. Kim, M. Collins, W. Bulmer, S. Sharma, and J. Mayrose, "Haptics assisted training (HAT) system for children's handwriting," in *Proc. World Haptics Conf. (WHC)*, Apr. 2013, pp. 559–564.

[118] W. Park, G. Korres, S. Tahir, and M. Eid, "Evaluation of handwriting skills in children with learning difficulties," in *Universal Access Human-Computer Interaction Multimodality Assistive Environments* M. Antona and C. Stephanidis, Eds. Cham, Switzerland: Springer, 2019, pp. 150–159.

[119] D. Wang, Y. Zhang, J. Hou, Y. Wang, P. Lv, Y. Chen, and H. Zhao, "IDental: A haptic-based dental simulator and its preliminary user evaluation," *IEEE Trans. Haptics*, vol. 5, no. 4, pp. 332–343, Oct. 2012.

[120] M. Kolesnikov, M. Zefran, A. D. Steinberg, and P. G. Bashook, "PerioSim: Haptic virtual reality simulator for sensorimotor skill acquisition in dentistry," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 689–694.

[121] C. Luciano, P. Banerjee, and T. DeFanti, "Haptics-based virtual reality periodontal training simulator," *Virtual Reality*, vol. 13, no. 2, pp. 69–85, Jun. 2009.

[122] G. Thomas, L. Johnson, S. Dow, and C. Stanford, "The design and testing of a force feedback dental simulator," *Comput. Methods Programs Biomed.*, vol. 64, no. 1, pp. 53–64, Jan. 2001.

[123] J. F. Ranta and W. A. Aviles, "The virtual reality dental training system: Simulating dental procedures for the purpose of training dental students using haptics," in *Proc. 4th PHANTOM users group workshop*, vol. 4, pp. 67–71, Nov. 1999.

[124] R. L. Williams, M. Srivastava, J. N. Howell, R. R. Conatser, D. C. Eland, J. M. Burns, and A. G. Chila, "The virtual haptic back for palpatory training," in *Proc. 6th Int. Conf. Multimodal Interface*, 2004, pp. 191–197.

[125] Y. Li, D. B. Kaber, L. Tupler, and Y.-S. Lee, "Haptic-based virtual environment design and modeling of motor skill assessment for brain injury PatientsRehabilitation," *Comput.-Aided Des. Appl.*, vol. 8, no. 2, pp. 149–162, Jan. 2011.

[126] G. Burdea, "The role of haptics in physical rehabilitation," in *Proc. Haptic Rendering, Found., Algorithms, Appl.*, Jan. 2008, pp. 517–529,
[127] Y. Yokokohji, Y. Sugawara, J. Kinoshita, and T. Yoshikawa, "Mechano-media that transmit kinesthetic knowledge from a human to other humans," in *Robotics Research*, R. A. Jarvis and A. Zelinsky, Eds. Berlin, Germany: Springer, 2003, pp. 499–512.
[128] M. Sung, Y. Yoo, K. Jun, N.-J. Kim, and J. Chae, "Experiments for a collaborative haptic virtual reality," in *Proc. 16th Int. Conf. Artif. Reality Telexistence–Workshops (ICAT)*, 2006, pp. 174–179.
[129] M. Kaiser and R. Dillmann, "Building elementary robot skills from human demonstration," in *Proc. IEEE Int. Conf. Robot. Autom.*, Dec. 1996, pp. 2700–2705.
[130] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proc. IEEE*, vol. 78, no. 9, pp. 1481–1497, Sep. 1990.
[131] D. D. Lee and H. S. Seung, "Learning in intelligent embedded systems," in *Proc. Workshop Embedded Syst.* Cambridge, MA, USA: USENIX Association, Mar. 1999, pp. 1–5.
[132] I. J. Goodfellow, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, vol. 27, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds. Red Hook, NY, USA: Curran Associates, 2014, pp. 2672–2680.
[133] I. J. Goodfellow, "NIPS 2016 tutorial: Generative adversarial networks," *CoRR*, vol. abs/1701.00160, pp. 1–5, Oct. 2017.
[134] C. L. Nehaniv and K. Dautenhahn, "Of hummingbirds and helicopters: An algebraic framework for interdisciplinary studies of imitation and its applications," in *Interdisciplinary Approaches to Robot Learning*. Singapore: World Scientific, 1999, pp. 136–161.
[135] B. Jansen and T. Belpaeme, "A computational model of intention reading in imitation," *Robot. Autom. Syst.*, vol. 54, no. 5, pp. 394–402, May 2006.
[136] N. Delson and H. West, "Robot programming by human demonstration: Adaptation and inconsistency in constrained motion," in *Proc. IEEE Int. Conf. Robot. Autom.*, Oct. 1996, pp. 30–36.
[137] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 37, no. 2, pp. 286–298, Apr. 2007.
[138] S. Ekvall and D. Kragic, "Learning task models from multiple human demonstrations," in *Proc. 15th IEEE Int. Symp. Robot Hum. Interact. Commun.*, Sep. 2006, pp. 358–363.
[139] M. N. Nicolescu and M. J. Mataric, "Natural methods for robot task learning: Instructive demonstrations, generalization and practice," in *Proc. 2nd Int. joint Conf. Autom. Agents Multiagent Syst.*, 2003, pp. 241–248.
[140] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, p. 91, Nov. 2004.
[141] M. Fishler, "Bolles: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 21, pp. 381–385, Jan. 1981.
[142] F. C. Huang, H. Mohamadipanah, F. A. Mussa-Ivaldi, and C. M. Pugh, "Combining metrics from clinical simulators and sensorimotor tasks can reveal the training background of surgeons," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 9, pp. 2576–2584, Sep. 2019.
[143] H. Mohamadipanah, C. Parthiban, J. Nathwani, D. Rutherford, S. DiMarco, and C. Pugh, "Can a virtual reality assessment of fine motor skill predict successful central line insertion?" *Amer. J. Surg.*, vol. 212, no. 4, pp. 573–578, 2016.
[144] A.-L.-D. D'Angelo, D. N. Rutherford, R. D. Ray, S. Laufer, A. Mason, and C. M. Pugh, "Working volume: Validity evidence for a motion-based metric of surgical efficiency," *Amer. J. Surg.*, vol. 211, no. 2, pp. 445–450, Feb. 2016.
[145] A.-L.-D. D'Angelo, D. N. Rutherford, R. D. Ray, S. Laufer, C. Kwan, E. R. Cohen, A. Mason, and C. M. Pugh, "Idle time: An underdeveloped performance metric for assessing surgical skill," *Amer. J. Surg.*, vol. 209, no. 4, pp. 645–651, Apr. 2015.
[146] A. L. Trejos, R. V. Patel, R. A. Malthaner, and C. M. Schlachta, "Development of force-based metrics for skills assessment in minimally invasive surgery," *Surg. Endoscopy*, vol. 28, no. 7, pp. 2106–2119, Jul. 2014.
[147] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[148] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
[149] Z. Wang and A. Majewicz Fey, "Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 12, pp. 1959–1970, Dec. 2018.
[150] M. Weinberg, C. Westphal, S. Froum, M. Palat, and R. Schoor, *Comprehensive Periodontics for Dental Hygienist*. London, U.K.: Pearson, 2014.
[151] S. Lanning, S. Pelok, B. Williams, P. Richards, D. Sarment, T.-J. Oh, and L. McCauley, "Variation in periodontal diagnosis and treatment planning among clinical instructors," *J. Dental Educ.*, vol. 69, pp. 325–337, Apr. 2005.
[152] *System Speech Recognition Namespace*, Microsoft Corporation, Albuquerque, NM, USA, 2020.
[153] F. Conti, F. Barbagli, R. Balaniuk, M. Halg, C. Lu, D. Morris, L. Sentis, J. Warren, O. Khatib, and K. Salisbury, "The CHAI libraries," in *Proc. Eurohaptics 2003*, Dublin, Ireland, 2003, pp. 496–500.
[154] D. C. Ruspini, K. Kolarov, and O. Khatib, "The haptic display of complex graphical environments," in *Proc. 24th Annu. Conf. Comput. Graph. Interact. Techn.*, 1997, pp. 345–352.
[155] M. Beraha, A. M. Metelli, M. Papini, A. Tirinzoni, and M. Restelli, "Feature selection via mutual information: New theoretical insights," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–9.
[156] J. R. Vergara and P. A. Estévez, "A review of feature selection methods based on mutual information," *Neural Comput. Appl.*, vol. 24, no. 1, pp. 175–186, Jan. 2014.
[157] D. Bruno, S. Calinon, and D. G. Caldwell, "Learning adaptive movements from demonstration and self-guided exploration," in *Proc. 4th Int. Conf. Develop. Learn. Epigenetic Robot.*, Oct. 2014, pp. 101–106.
[158] R. Dillmann, M. Kaiser, and A. Ude, "Acquisition of elementary robot skills from human demonstration," in *Proc. Int. Symp. Intell. Robot. Syst.*, Jul. 1995, pp. 185–192.
[159] R. Dillmann, "Teaching and learning of robot tasks via observation of human performance," *Robot. Autom. Syst.*, vol. 47, nos. 2–3, pp. 109–116, Jun. 2004.
[160] A. M. Schmidts, D. Lee, and A. Peer, "Imitation learning of human grasping skills from motion and force data," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Sep. 2011, pp. 1002–1007.
[161] J. Peters, K. Mülling, J. Kober, D. Nguyen-Tuong, and O. Krömer, "Towards motor skill learning for robotics," in *Robotics Research*, C. Pradalier, R. Siegwart, and G. Hirzinger, Eds. Berlin, Germany: Berlin, 2011, pp. 469–482.
[162] S. Schaal and D. Sternad, "Programmable pattern generators," in *Proc. 3rd Int. Conf. Comput. Intell. Neurosci.*, Research Triangle Park, NC, Oct. 1998, pp. 48–51.
[163] E. Theodorou, J. Buchli, and S. Schaal, "Reinforcement learning of motor skills in high dimensions: A path integral approach," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2010, pp. 2397–2403.
[164] Y. Kuniyoshi, M. Inaba, and H. Inoue, "Learning by watching: Extracting reusable task knowledge from visual observation of human performance," *IEEE Trans. Robot. Autom.*, vol. 10, no. 6, pp. 799–822, Oct. 1994.
[165] J. Aleotti, S. Caselli, and M. Reggiani, "Leveraging on a virtual environment for robot programming by demonstration," *Robot. Autom. Syst.*, vol. 47, nos. 2–3, pp. 153–161, Jun. 2004.
[166] D. Feygin, M. Keehner, and R. Tendick, "Haptic guidance: Experimental evaluation of a haptic training method for a perceptual motor skill," in *Proc. 10th Symp. Haptic Interface Virtual Environ. Teleoperator Syst.*, 2002, pp. 40–47.
[167] R. Kouskouridas, A. Amanatiadis, and A. Gasteratos, "Guiding a robotic gripper by visual feedback for object manipulation tasks," in *Proc. IEEE Int. Conf. Mechatronics*, Apr. 2011, pp. 433–438.

VAHAN BABUSHKIN received the B.Sc. degree in electronics and microelectronics from Slavonic University, in 2010, and the M.Sc. degree in computing and information science from Khalifa University (MIT and Masdar Institute Cooperative Program), Abu Dhabi, in 2013. He is currently pursuing the Ph.D. degree in electrical engineering with the NYU Tandon School of Engineering. He is interested in machine learning applications to different areas of computational science and engineering, including haptics, robotics, and human–computer interaction. His current research at the Applied Interactive Multimedia (AIM) Laboratory is focused on the sensorimotor skills learning for periodontal training (For more information, please visit: http://vaanbabushkin.narod.ru).

MUHAMMAD HASSAN JAMIL received the B.Sc. degree in computer engineering from the COMSATS Institute of Information Technology, Pakistan, and the M.S. degree in computer science from the National University of Computer and Emerging Sciences, Pakistan. He was involved with the design and development of virtual reality, and augmented reality experiences and interfaces. He has also worked previously as a Design Engineer at Open-Silicon Inc. (an ASIC manufacturing company), where he was a part of the Hybrid Memory Cube Controller (HMCC) Design Team. He is currently a part of the research group at the Applied Interactive Multimedia Laboratory, New York University Abu Dhabi (NYUAD), on various projects related to the design and development of immersive haptic interfaces and neurohaptics. His research interests include multimodal haptic interface design, neurohaptics, and immersive haptic systems.

WANJOO PARK received the Ph.D. degree from the Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea, in 2016. From December 2008 to March 2017, he had research experiments at the Korea Institute of Science of Technology (KIST), as a Research Scientist. He is currently a Postdoctoral Associate with the Division of Engineering, New York University Abu Dhabi (NYUAD). His current research interests include haptics, human–computer interaction, brain–computer interface, neuro-rehabilitation, and Internet gaming addiction (For more information, please visit the website: http://wanjoopark.wixsite.com/wanjoo).

MOHAMAD EID (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Ottawa, Canada, in 2010. From June 2008 to April 2012, he was a Teaching and Research Associate with the University of Ottawa. He is currently an Assistant Professor of electrical and computer engineering with the Division of Engineering, New York University Abu Dhabi (NYUAD). He has coauthored the book *Haptics Technologies: Bringing Touch to Multimedia* (Springer, 2011). His academic interest includes multimedia haptics, with emphasis on affective haptics, tangible human–computer interaction, and instrumentation (sensors and actuators). He is the Co-Chair of the Third International IEEE Workshop on Multimedia Services and Technologies for E-health a(MUST-EH 2013), and has been involved in the organization of the Haptic-Audio-Visual Environment and Gaming (HAVE) Workshop for the years 2007, 2008, 2009, 2010, and 2013.