

# Teleoperation of Humanoid Robots: A Survey

Kourosh Darvish<sup>1\*</sup>, Luigi Penco<sup>2</sup>, Joao Ramos<sup>3</sup>, Rafael Cisneros<sup>4</sup>,  
Jerry Pratt<sup>2</sup>, Eiichi Yoshida<sup>6</sup>, Serena Ivaldi<sup>5</sup>, Daniele Pucci<sup>1</sup>

**Abstract**—Teleoperation of humanoid robots enables the integration of the cognitive skills and domain expertise of humans with the physical capabilities of humanoid robots. The operational versatility of humanoid robots makes them the ideal platform for a wide range of applications when teleoperating in a remote environment. However, the complexity of humanoid robots imposes challenges for teleoperation, particularly in unstructured dynamic environments with limited communication. Many advancements have been achieved in the last decades in this area, but a comprehensive overview is still missing. This survey paper gives an extensive overview of humanoid robot teleoperation, presenting the general architecture of a teleoperation system and analyzing the different components. We also discuss different aspects of the topic, including technological and methodological advances, as well as potential applications. A web-based version of the paper can be found at <https://humanoid-teleoperation.github.io/>.

**Index Terms**—Humanoid robot; Teleoperation.

## I. INTRODUCTION

There are many situations and environments where we need robots to replace humans at the site. Despite the recent progress in robot cognition based on AI techniques, fully autonomous solutions are still far from producing socially and physically competent robot behaviors; that is why teleoperating robots (Fig. 1) acting as physical avatars of human workers at the site is the most reasonable solution. In environments like construction sites, chemical plants, contaminated areas and space, teleoperated robots could be extremely valuable, relieving humans from any potential hazard. Contrary to other conventional robotic platforms, humanoids' structure is a better fit for environments and tasks that are designed for and performed by humans. The operational versatility of these robots makes them suitable for work activities that require



Fig. 1: Examples of humanoid robot teleoperation; from top left to bottom right corner: [1], [2], [3], [4], [5], [6], [7], [8].

a variety of complex mobility and manipulation skills, such as inspection, maintenance, and interaction with human operators. In certain contexts such as telenursing, where human subjects are expected to interact with a teleoperated robot, the human-likeness factor is important since it increases the acceptability, social closeness, and legibility of its intentions [9].

In the literature, different attempts have been made to “*deploy the senses, actions, and presence of a human to a remote location in real-time, leading to a more connected world*” [10]. Inspired by a visit to the Tachi Lab, the XPRIZE Foundation has recently launched the ANA Avatar XPRIZE global competition [10]. Previously, in response to the 2011 Fukushima Daiichi nuclear disaster, the DARPA Robotics Challenge (DRC) was launched to promote innovation in human-supervised robotic technology. In space applications, rovers and mobile manipulators were teleoperated from aboard the International Space Station (ISS), in the context of ME-TERON and Kontur-2 projects [11]. In 2019, the humanoid Skybot F-850 was rocketed to the ISS [12]; however, it turned out to have a design that did not work well, demonstrating that there is still work to do to get humanoids into space.

Humanoid robot teleoperation involves many multidisciplinary and interleaved challenges, ranging from dynamics and control to communication and human psychophysiology. Uniquely, due to their resemblance to human appearance, societal expectations are high as well; they are expected to do a wide range of tasks that are not expected from other types of robots. They are highly redundant with nonlinear, hybrid, and underactuated models. While doing dynamic and agile motions with the feet like walking, running, or stepping over

<sup>1</sup> Artificial and Mechanical Intelligence, Center for Robotics and Intelligent Systems, Istituto Italiano di Tecnologia, Genoa, Italy, Kourosh.Darvish@gmail.com, Daniele.Pucci@iit.it

<sup>2</sup> Florida Institute for Human and Machine Cognition, 40 S Alcaniz St, Pensacola, FL 32502, United States lpenco@ihmc.org, jpratt@ihmc.org

<sup>3</sup> Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, USA, jlramos@illinois.edu

<sup>4</sup> CNRS-AIST JRL (Joint Robotics Laboratory), IRL; National Institute of Advanced Industrial Science and Technology (AIST), Department of Information Technology and Human Factors, Tsukuba, Japan, rafael.cisneros@aist.go.jp

<sup>5</sup> Inria, Loria, Université de Lorraine, CNRS, Nancy, France, serena.ivaldi@inria.fr

<sup>6</sup> Department of Applied Electronics, Faculty of Advanced Engineering, Tokyo University of Science, Japan, eiichi.yoshida@rs.tus.ac.jp

\* Corresponding author

This work has received funding from the European Union's Horizon 2020 research and innovation programmes under grant agreement No. 731540 (An.Dy), No. 869855 (SoftManBot), No. 101070596 (EuRobin), & the Italian National Institute for Insurance against Accidents (INAIL) ergoCub project.

obstacles, they are supposed to perform dexterous power and precision manipulation. At the same time, they are expected to work alongside humans, be safe, friendly, and socially interact with others. On the other hand, teleoperation interfaces and techniques should be designed such that the human operator receives minimal, effective, and informative haptic feedback from the humanoid robot, to cover for human errors, overcome communication delays, and above all, be telepresent. Along with these challenges, the field is new and due to its high resource demand for development, not many laboratories have been working on it.

Many efforts from the robotics community have been devoted to studying humanoid robots, teleoperation, evaluation metrics, or human-robot interaction. Among them, the book on humanoid robotics [13] studied comprehensively different aspects of humanoid robots, including their history, design, mechanics, control, simulation, and interaction. Several survey papers likewise studied specific aspects of humanoid robots, for example humanoid dynamics [14], control [15], motion generation [16], or robot teleoperation interface design and metrics [17]. A primary work on bilateral teleoperation techniques has been presented by [18] as well. Another seminal survey [19] covers many aspects of interactive robots, including their design, autonomy level, and human factors that helped us in articulating the current manuscript. Another interesting survey [20] highlights several aspects of humanoid teleoperation and autonomy. However, [20] is a decade old, and an up-to-date survey on the topic is missing, especially considering that humanoid teleoperation is a far-from-solved challenge and a highly active field of research where new solutions are proposed each year. Following the workshop in [21], this survey paper presents the latest results in humanoid robot teleoperation and draws in detail the challenges that the research community faces to effectively deploy such systems toward real-world scenarios.

Starting from what emerged from the workshop, we conducted a survey on teleoperation of humanoid robots. We present here the systems and devices that have been adopted so far to teleoperate humanoids (Sec. II) and how these robots have been modeled, retargeted, and controlled (Sec. III). We also examine a promising case of teleoperation in which the robot assists the user in accomplishing a desired task (Sec. IV). Later, we discuss complications along with some compensating solutions that arise due to non-ideal communication channels (Sec. V). We explain the evaluation of teleoperation systems prior to development to meet the users' needs (Sec. VI). Finally, discussions on current and potential applications and the associated challenges follow (Sec. VII).

## II. TELEOPERATION SYSTEM AND DEVICES

In the literature, the terms teleoperation and telepresence have been used indistinctly in different contexts. Telepresence refers to the technology that allows human to virtually exist in a remote location through an avatar, experiencing real-time sensations from the remote site [22]. Both the remote environment and the avatar can be real or virtual, but in this article we only consider a real environment and a surrogate humanoid robot as avatar. Telepresence has also been

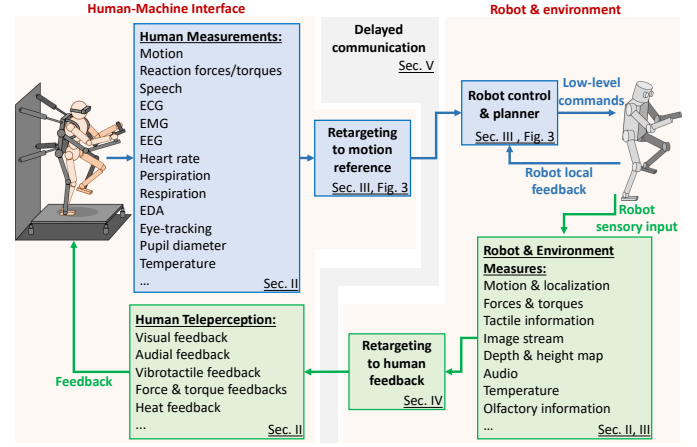


Fig. 2: Schematic architecture for teleoperating a humanoid.

referred to as telepresence in the literature [22]. The concept of teleoperation, on the other hand, still refers to a human operator remotely controlling a robot, but the focus is mainly put on performing tasks that require high dexterity in the remote location. We use these terms interchangeably throughout this survey. From another perspective, the teleoperation setup represents an interactive system where the robot imitates the human's actions to reach a common objective [19].

In a teleoperation setup, the user is the person who teleoperates the humanoid robot and identifies the teleoperation goal, i.e. intended outcome, through interfaces. The interfaces are the means of the interaction between the user and the robot. The nature of the exchanging information, constraints, the task requirements, and the degree of shared autonomy determine different preferences on the interface modalities, according to specific metrics which will be discussed later. Moreover, the choice of the interfaces should make the user feel comfortable, hence enabling a natural and intuitive teleoperation.

### A. Teleoperation Architecture

Fig. 2 shows a schematic view of the architecture for teleoperating a humanoid robot. First, human kinematic and dynamic information are measured and transmitted to the humanoid robot motion for teleoperation. More complex retargeting methods (i.e., mapping of the human motion to the robot motion), employed in assisted teleoperation systems (Sec. IV), may need the estimation of the user reaction forces/torques. There are cases in which also physiological signals are measured in order to estimate the psychophysiological state of the user, which can help enhancing the performance of the teleoperation. On the basis of the estimated states, the retargeting policy is selected and the references are provided to the robot accordingly. Teleoperation systems are employed not only for telemanipulation scenarios but also for social teleinteractions (i.e. remotely interacting with other people). In this case, the robot's anthropomorphic motion and social cues such as facial expressions can enhance the interaction experience. Therefore, rich human sensory information is indispensable.

To effectively teleoperate the robot, the user should make proper decisions; therefore he/she should receive various feedback from the remote environment and the robot. In many

cases, the sensors for perceiving the human data and the technology to provide feedback to the user are integrated together in an interface. In the rest of the section, we will discuss different available interfaces and sensor technologies in teleoperation scenarios, whereas their design and evaluation will be discussed later in Sec. VI.

The retargeting block (Fig. 2) maps human sensory information to the reference behavior for the robot teleoperation, hence the human can be considered as master or leader, and the robot as slave or follower. We can discern two retargeting strategies: unilateral and bilateral teleoperation. In the unilateral approach, the coupling between the human and robot takes place only in one direction. The human operator can still receive haptic feedback either as a kinesthetic cue not directly related to the contact force being generated by the robot or as an indirect force in a passive part of her/his body not commanding the robot. But in bilateral systems, human and humanoid robot are directly coupled. The choice of bilateral retargeting depends on the task, communication rate, and degree of shared autonomy, as will be discussed in Sec. IV.

The human retargeted information, together with the feedback from the robot, are streamed over a communication channel that could be non-ideal. In fact, long distances between the operator and the robot or poor network conditions may induce delays in the flow of information, packet loss and limited bandwidth, adversely affecting the teleoperation experience. Sec. V details the approaches in the literature to teleoperate robots in such conditions. Finally, the robot local control loop generates the low level commands- i.e., joints position, velocity, or torque references- to the robot, taking into account the human references (Sec. III).

## B. Human Sensory Measurement Devices

In this section, we provide a detailed description of the technologies available to sense various human states, including the advances to measure the human motion, the physiological states, and the interaction forces with the environment. We report in TABLE I the various measurement devices adopted so far for humanoid robot teleoperation.

1) *Human kinematics and dynamics measurements*: To provide the references for the robot motion, we need to sense human intentions. For simple teleoperation cases, the measurements may be granted through simple interfaces such as keyboard, mouse, or joystick commands [23]. However, for a more complex system like a humanoid robot, those simple interfaces may not be enough, especially when the user wants to exert a high level of control authority over the robot. Therefore, the need for natural interfaces for effectively commanding the robot arises. We can consider solutions benefiting from the similarity of the human and the humanoid robot anthropomorphic geometries, i.e., providing the references to the robot limbs with spatial analogies to those of the human (natural mapping). Therefore, the need to measure the human kinematics emerges.

Different technologies have been employed in the literature to measure the motion of the main human limbs, such as legs, torso, arms, head, and hands. An ubiquitous option are the Inertial Measurement Unit (IMU)-based wearable technologies.

In this context, two cases are possible, a segregated set of IMU sensors are used throughout the body to measure the human motion or an integrated network of sensors is used throughout the human body [2], [3]. The former case normally provides the raw IMU values, while the latter provides the information of the human limb movements. In the second case, a calibration process computes the sensors transformations with respect to body segments [24]. This technology is especially of interest because of the high accuracy and frequency of the retrieved human motion information without the occlusion problem. However, its accuracy may suffer from disturbances caused by the magnetic field and displacement of the sensors with respect to their initial emplacement. Yet, another recent wearable technology to capture the hand pose is a stretchable soft glove embedded with distributed capacitive sensors [25]. A review about the textile-based wearable technology used to estimate the human kinematics is provided in [26].

Optical sensors are another technology used to capture the human motion, with active and passive variants. In the active case, the reflection of the pattern is sensed by the optical sensors. Some of the employed technologies include depth sensors and optical motion capture systems. In the case of passive sensors, RGB monocular cameras (regular cameras) or binocular cameras (stereo cameras) are used to track the human motions. Thanks to the optical sensors, a skeleton of the human body is generated and tracked in 2-D or 3-D Cartesian space. The main problems with these methods are the occlusion and the low portability of the setup.

To track the users' motion in bilateral teleoperation scenarios, exoskeletons are often used. In this case, the exoskeleton model and the encoder data are fed to the forward kinematics to estimate the human link's poses and velocities [27], [4].

The previously introduced technologies can be used to measure the human gait information with locomotion analysis. Conventional or omnidirectional treadmills are employed in the literature for this purpose [28]. While the treadmill can be used for even terrains, it would not work to retarget locomotion on uneven terrains. To respond to this shortcoming, a cockpit-like teleoperation setup has been recently proposed [4]. To estimate the interaction forces between the human and the teleoperation setup, force-torque sensors measuring the human wrenches can be integrated in ground plates, shoes, or exoskeletons. Richer information can be obtained by distributed capacitance sensors that measure the pressure manifold.

2) *Human physiological measurements*: Among the different sensors available to measure human physiological activities, we briefly describe those that have been mostly used in teleoperation and robotics literature. An electromyography (EMG) sensor provides a measure of the muscle activity, i.e., contraction, in response to the neural stimulation action potential [42]. It works by measuring the difference between the electrical potential generated in the muscle fibres by employing two or more electrodes. There are two types of EMG sensors, the surface EMG and the intramuscular EMG. The former records the muscle activity from above the skin (therefore, noninvasive), while the latter measures the muscle activity by inserting needle electrodes into the muscle (invasive). The main problem with EMG sensors, especially the

TABLE I: Main works and technologies related to the teleoperation of humanoids.

ref	robot	teleoperation devices										retargeting & planning			stabilizer		whole-body controller			low-level joint control			
		human motion measurement			feedback							GUI-based planner	retargeting & motion generation			ZMP	contact wrench, force distribution	IK/velocity IK	momentum-based inverse dynamics	QP method	joint position	joint torque	
		motion-capture suit	optical-tracking	exoskeleton	mouse, keyboard, joystick, treadmill	visual		haptic					footstep motion generation	upper-body retargeting	whole-body retargeting								
						AR/VR headset	Stereo, mono display, GUI	whole-body exoskeleton	dual-arm exoskeleton	glove	balance												vibrotactile
[2]	iCub Genova04	✓			✓		✓						✓		✓			✓	✓	✓	✓	✓	
[29]	iCub3	✓	✓		✓		✓				✓		✓	✓				✓	✓	✓	✓	✓	
[8]	HRP-4CR	✓			✓		✓						✓	✓	✓			✓	✓	✓	✓	✓	
[28]	iCub Genova04		✓		✓		✓						✓	✓				✓	✓	✓	✓	✓	
[5]	Valkyrie				✓		✓						✓				✓	✓	✓	✓		✓	
[30]	Atlas				✓		✓						✓				✓	✓	✓	✓		✓	
[31]	DRC-Hubo		✓		✓		✓						✓		✓			✓	✓	✓	✓	✓	
[3]	iCub Nancy01	✓			✓		✓						✓					✓	✓	✓	✓	✓	
[32]	DRC-Hubo Beta				✓		✓						✓					✓	✓	✓		✓	
[33]	HRP-2KAI				✓		✓						✓	✓				✓	✓	✓	✓	✓	
[4]	JAXON			✓			✓			✓							✓	✓	✓	✓	✓	✓	
[1]	little HERMES			✓			✓			✓			✓					✓	✓	✓	✓	✓	
[7]	MAHRU	✓					✓			✓			✓	✓				✓	✓	✓	✓	✓	
[34]	NimbRo Avatar			✓	✓		✓			✓				✓				✓	✓	✓		✓	
[35]	HRP-2			✓									✓					✓	✓	✓	✓	✓	
[36]	JAXON		✓				✓											✓	✓	✓	✓	✓	
[37]	TELESAR VI		✓				✓											✓	✓	✓	✓	✓	
[38]	TORO	✓																✓	✓	✓	✓	✓	
[6]	TORO			✓						✓								✓	✓	✓	✓	✓	
[39]	COMAN		✓										✓					✓	✓	✓	✓	✓	
[40]	HRP-4	✓																✓	✓	✓	✓	✓	
[41]	Fujitsu HOAP-3		✓																✓	✓	✓	✓	

surface one, is the low signal to noise ratio, which is the main barrier for a desirable performance [43]. The EMG signals are used in the literature for teleoperating a robot or a prosthesis, for estimation of the human effort and muscle forces [42], or for estimation of the muscle stiffness. The EMG signals can anticipate human motions by measuring the muscle activities within a few milliseconds in advance of force generation; this could be exploited to anticipate the human operator's motion, enhancing the teleoperation.

Electroencephalography (EEG) sensors can be employed to identify the user mental state. They are most widely used in non-invasive brain-machine interface (BMI) and they monitor the brainwaves resulting from the neural activity of the brain. The measurement is done by placing several electrodes on the scalp and measuring the small electrical signal fluctuations [44]. Other sensors that could be employed in a telepresence scenario for an advanced estimation of the human psychophysiological state include the Heart Rate Monitor sensors, which estimate the maximal uptake of the oxygen and the heart rate variability [45], useful to measure the user's fatigue, Electro-Oculography (EOG) sensors or video-based eye-trackers, which estimate of gaze position based on the pupil or iris position [46], important for the visual feedback given by Virtual Reality (VR) or Augmented Reality (AR) goggles; and capacitive thin-film humidity sensors, which measure the humidity of the gas-flow of the human skin [47], a good indicator of the human emotional stimuli and stress level.

### C. Feedback Interfaces: Robot to Human

A crucial point in robot teleoperation is to sufficiently inform the human operator of the states of the robot and

its work site, so that she/he can comprehend the situation or feel physically present at the site, producing effective robot behaviors. TABLE I summarizes the different feedback devices that have been adopted for humanoid teleoperation.

1) *Visual feedback*: A conventional way to provide situation awareness to the human operator is through visual feedback. Visual information allows the user to localize themselves and other humans or objects in the remote environment. Graphical User Interfaces (GUIs) were widely used by the teams participating at the DRC to remotely pilot their robots through the competition tasks [30], [33], [32]. Not only the information coming from the RGB cameras of the robot but also other information such as depth, LIDAR, RADAR, and thermographic maps of the remote environment was displayed to the user.

An alternative way to give visual feedback to the human operator is through VR headsets, connected to the robot cameras. Although this has been demonstrated to be effective in several robotic experiments [7], [36], [2], [3], during locomotion the users often suffer from motion sickness, since the images from the robot cameras are not stabilized, while the images perceived by the human eyes are automatically stabilized on the retina thanks to compensatory eye reflexes. This aspect could be improved by adopting digital image stabilization techniques or AR [48]. Another issue concerning the visual feedback is related to the limited bandwidth of the communication link, which can delay the stream of information. Also, human reaction time to visual input is inherently slow ( $\approx 250$ -300 ms), so a higher delay in the stream of information can be perceived by the user and further aggravate the motion sickness. If also haptic feedback is streamed to the operator, then even lower

delays can be disturbing. In fact, human reaction to haptic information is much faster (around 100-150ms).

2) *Haptic feedback*: Visual feedback is not often sufficient for many real-world applications, especially those involving power manipulation (with high forces) or interaction with other human subjects, where the dynamics of the robot, the contact forces with the environment, and the human-robot interaction forces are of crucial importance. In such scenarios also the haptic feedback is required to exploit the human operator's motor skills in order to augment the robot performance.

There are different technologies available in the literature to provide haptic feedback to the human. Force feedback, tactile and vibro-tactile feedback are the most used in teleoperation scenarios. The interface providing kinesthetic force feedback can be similar to an exoskeleton [49] or can be cable driven. The latter only provides a tension force feedback, while the former provides the force feedback on different directions. Dual-arm exoskeletons have been proposed to guide the teleoperated robot during manipulation tasks while receiving haptic feedback through the same actuated exoskeleton arms [22], [6] and recently also a whole-body exoskeleton cockpit has been proposed to teleoperate the JAXON humanoid robot during heavy manipulation tasks and stepping on uneven terrains [4], getting a force feedback on the whole limbs.

To convey the sense of touch, tactile displays have been adopted in the literature [50]. These can also provide temperature feedback to the user. Other employed haptic feedbacks are vibrotactile and air pressure ones, used as a sensory substitution to transmit senses of touch, texture, forces, suggesting directions, or to catch the attention of the user [39].

All these types of haptic feedback are combined in the telexistence system TELESAR V [22], which has been developed to provide complete cutaneous sensations to the human operator. The idea is that different physical stimuli give rise to the same sensation in humans and are perceived as identical. This is due to the fact that human skin has limited receptors and can perceive only force, vibration, and temperature, which in [51] are defined as *haptic primary colors*. It is thus sufficient to combine these *colors* in order to reproduce any cutaneous sensation without actually touching the real object.

3) *Balance feedback*: Haptic feedback can also be used to transmit to the operator a sense of the robot's balance. The idea behind the balance feedback is to transfer to the operator the information about the *effect of disturbances* over the robot dynamics or stability instead of directly mapping to the human the disturbance forces applied to the robot. In [39], Brygo *et al.* proposed to provide the feedback of the robot's balance state by means of a vibro-tactile belt. Also, Peternel and Babic [41] proposed a cable-driven haptic interface that maps the state of the robot's balance to the human demonstrator. Alternatively, Abi-Farrajil *et al.* [6] introduced a task-relevant haptic feedback interface composed by two light-weight robotic arms that receives high-level informative haptic cues informing the user about the impact of her/his potential actions on the robot's balance. These studies do not investigate the case of dynamic behaviors, but are rather limited to double support scenarios. In [1] instead, simultaneous stepping is enabled via bilateral coupling between the human operator, wearing a Balance

Feedback Interface (BFI), and the robot. The BFI is composed of an passive exoskeleton which captures human body posture and a parallel mechanism which applies feedback forces to the operator's torso.

4) *Auditory feedback*: Auditory feedback is another means of communication. It is mainly provided to the user through headphones, single or multiple speakers. Auditory information can be used for different purposes: to enable the user to communicate with others in the remote environment through the robot, to increase the user situational awareness, to localize the sound source by using several microphones, or to detect the collision of the robot links with the environment. The user and the teleoperated robot may also communicate through the audio channel; e.g., for state transitions.

#### D. Graphical User Interfaces (GUIs)

GUIs are used in the literature to provide both feedback to the user and give commands to the robot. In the DRC, operators were able to supervise the task execution through a task panel, using manual interfaces in case they needed to make corrections. The main window consisted of a 2D and 3D visualization environment, the robot's current and goal states, motion plans, together with other perception sensor data such as hardware driver status [52], [32]. Due to the limited robot cognitive skills, perception tasks were shared among the users and the robot.

A common approach adopted by the different teams was to guide the robot perception algorithms to fit the object models to the point cloud, used to estimate the 3D pose of objects of interest in the environment. For example, operators were annotating search regions by clicking on displayed camera images or by clicking on 3D positions in the point cloud. Following that, markers were used to identify the goal pose of the robot arm end effectors [30], the robot configuration, or with a higher autonomy level to define the goal pose of objects for manipulation tasks [52]. In these cases, robots tried to find an obstacle-free path (related to Sec. III-B1), and show the generated path to the operator for verification prior to execution [52]. Throughout this process, the robots' lower-body teleoperation was treated differently. When the robot's desired base/CoM goal or footsteps were marked by the user, an obstacle-free path (Sec. III-B1) and footsteps trajectories were automatically generated (Sec. III-B2). In this process, footsteps were adjusted to uneven terrain given the estimated height-field data [52].

GUIs have also been used to command frequent high-level tasks to the robot by encoding them as state machines or as task sequences [52]. In DRC open-source software tools, such as RViz and Choreonoid, were commonly used [32], [52]. Custom functionalities were added to them using software plugins. Today, many of these functionalities can be integrated with VR and AR devices.

### III. HUMANOID ROBOT RETARGETING AND CONTROL

This section describes the retargeting and control techniques for unilateral teleoperation of humanoid robots. We can define the retargeting and control as a mapping  $\mathbb{H} : A \rightarrow A'$  where



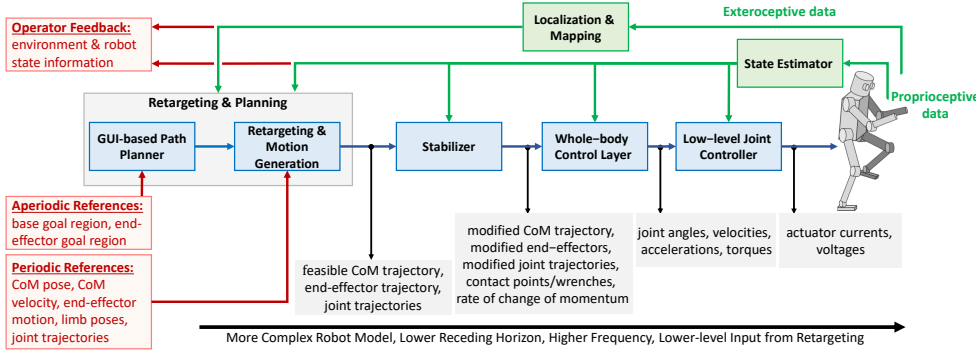


Fig. 3: Flowchart of a humanoid robot retargeting, planning, and control architecture (red color: references/feedback of human; blue color: retargeting, motion generation, and control; green color: perception and estimation).

$A$  is the domain of the perceived human actions (kinematic and dynamic trajectories) and  $A'$  is the space of robot actions. In this definition, the mapping principle identifies the robot autonomy and the human authority in the teleoperation scenario. This mapping should be identified such that it minimizes the difference between the human intent and the robot actions while respecting the constraints in the teleoperation scenario.

Humanoid robot retargeting and control introduces many challenges to teleoperation scenarios, namely due to the non-linear, hybrid, and underactuated dynamics of humanoids with high degrees of freedom, as well as imprecise robot dynamical model and control, and partially-known environment dynamics. All these challenges together with the fact that the robot retargeting and controller blocks (in Fig. 3) usually run online, make the problem even more complicated. To overcome the introduced challenges, model-based optimal control architectures are the foremost technique used in the literature. During DRC, most of the finalist architectures converged to a similar design [53], [30] where the robot trajectories and stability are achieved by retargeting and control blocks connected in cascade. This architecture is conceptually demonstrated in Fig. 3, where its blocks are characterized in the following sections. This architecture needs a humanoid model (Sec. III-A), human references coming from teleoperation devices, and the robot environment information. Finally, challenges unique to humanoid robot retargeting, control, and possible future directions are discussed in Sec. III-G.

### A. Modeling

1) *Notations and complete humanoid robot models*: Most humans and humanoid robots are modeled as multi-body mechanical systems with  $n + 1$  rigid bodies, called links, connected with  $n$  joints, each with one degree of freedom. The configuration of a humanoid robot with  $n$  joints depends on the robot shape, i.e., joint angles  $\mathbf{s} \in \mathbb{R}^n$ , the position  ${}^{\mathcal{I}}\mathbf{p}_B \in \mathbb{R}^3$ , and orientation  ${}^{\mathcal{I}}\mathbf{R}_B \in SO(3)$  of the floating base (usually associated with the pelvis,  $B$ ) relative to the inertial or world frame  $\mathcal{I}^1$ . The robot configuration is indicated by  $\mathbf{q} = ({}^{\mathcal{I}}\mathbf{p}_B, {}^{\mathcal{I}}\mathbf{R}_B, \mathbf{s})$ . The pose of a frame attached to a robot link  $\mathcal{A}$  is computed via  $({}^{\mathcal{I}}\mathbf{p}_A, {}^{\mathcal{I}}\mathbf{R}_A) = \mathcal{H}_A(\mathbf{q})$ , where  $\mathcal{H}_A(\cdot)$  is the geometrical forward kinematics. The velocity of the model is summarized in the vector  $\boldsymbol{\nu} = ({}^{\mathcal{I}}\dot{\mathbf{p}}_B, {}^{\mathcal{I}}\boldsymbol{\omega}_B, \dot{\mathbf{s}}) \in \mathbb{R}^{n+6}$ , where  ${}^{\mathcal{I}}\dot{\mathbf{p}}_B, {}^{\mathcal{I}}\boldsymbol{\omega}_B$  are the base linear and rotational (angular) velocity

of the base frame, and  $\dot{\mathbf{s}}$  is the joints velocity vector of the robot. The velocity of a frame  $\mathcal{A}$  attached to the robot, i.e.,  ${}^{\mathcal{I}}\mathbf{v}_A = ({}^{\mathcal{I}}\dot{\mathbf{p}}_A, {}^{\mathcal{I}}\boldsymbol{\omega}_A)$ , is computed by *Jacobian* of  $\mathcal{A}$  with the linear and angular parts, therefore  ${}^{\mathcal{A}}\mathbf{v}_{\mathcal{I}} = \mathcal{J}_A(\mathbf{q})\boldsymbol{\nu}$ . Finally, the  $n + 6$  robot dynamics equations, with all  $n_c$  contact forces applied on the robot, are described by [54]:

$$\mathbf{M}(\mathbf{q})\dot{\boldsymbol{\nu}} + \mathbf{C}(\mathbf{q}, \boldsymbol{\nu})\boldsymbol{\nu} + \mathbf{g}(\mathbf{q}) = \mathbf{B}\boldsymbol{\tau} + \sum_{k=1}^{n_c} \mathcal{J}_k^T(\mathbf{q})\mathbf{f}_k^c, \quad (1)$$

where  $\mathbf{M}(\mathbf{q})$  is the symmetric positive definite inertia matrix of the robot,  $\mathbf{C}(\mathbf{q}, \boldsymbol{\nu})$  is the vector of Coriolis and centrifugal terms,  $\mathbf{g}(\mathbf{q})$  is the vector of gravitational terms,  $\mathbf{B} = (\mathbf{0}_{n \times 6}, \mathbf{1}_n)^T$  is a selector matrix,  $\boldsymbol{\tau}$  is the vector of actuator joint torques, and  $\mathbf{f}_k^c$  is the vector of the  $k$ -th contact wrenches acting on the robot. More information about the humanoid robot modeling can be found in [14].

2) *Simplified humanoid robot models*: Various simplified models are proposed in the literature in order to extract intuitive control heuristics and for real-time lower order motion planning. The most well-known approximation of humanoid dynamics is the inverted pendulum model [55]. In this model, the support foot is connected through a variable-length link to the robot center of mass (CoM). Assuming a constant height for the inverted pendulum [55], one can derive the equation of motion of the *Linear Inverted Pendulum Model* (LIPM) by:

$$\ddot{\mathbf{x}} = \frac{g}{z_0}(\mathbf{x} - \mathbf{x}_b), \quad (2)$$

where  $g$  is the gravitational acceleration constant,  $z_0$  is the constant height of the CoM,  $\mathbf{x} \in \mathbb{R}^2$  is the CoM coordinate vector, and  $\mathbf{x}_b \in \mathbb{R}^2$  is the base of the LIPM coordinate vector.

The base of the LIPM is often assumed to be the Zero Moment Point (ZMP) (equivalent to the center of pressure, CoP) of the humanoid robot [56]. The LIPM dynamics can be divided into stable and unstable modes, where the unstable mode is referred as (instantaneous) Capture Point [57], or Divergent Component of Motion (DCM) [58], [59] in the literature. The DCM dynamics is characterised by a first-order system as:

$$\dot{\boldsymbol{\xi}} = \mathbf{x} + b\dot{\mathbf{x}}, \quad (3)$$

where  $\boldsymbol{\xi}$  and  $b = \sqrt{\frac{z_0}{g}}$  are the DCM variable and the time constant. Equation 3 shows that the CoM follows the DCM. Differentiating Eq. (3) and replacing into Eq. (2) results in:

$$\dot{\boldsymbol{\xi}} = \frac{1}{b}(\boldsymbol{\xi} - \mathbf{x}_b). \quad (4)$$

<sup>1</sup>With the abuse of notation, we will drop  $\mathcal{I}$  in formulas for simplicity.

Equations (3) and (4) together represent the LIPM dynamics.

### B. Retargeting & Planning

The goal of this block in Fig. 3 is to morph the human commands or measurements coming from the teleoperation devices into robot references. It comprises the reference motions for the robot links as well as the robot locomotion references, i.e., alternate footstep locations, timings, and allocating a given footstep to the left or right foot to follow the user's commands. The input to this block may vary according to the task and system requirements, and consequently the design choice and retargeting policy differ. Besides, the retargeting policy can even be determined online as a classification or regression problem using the human speech and psychophysiological state. For example in [60], [61], the authors developed attentive systems for real-time adaptation of the retargeting policy and the robot autonomy level. As we go from the left to the right in Fig. 3, the level of automation and the input frequency increase, and higher communication bandwidth with lower delays is required. The approaches presented in this part are model-based, while learning-based techniques for retargeting will be discussed later in Sec. III-G.

Teleoperation devices may provide different types of input to retargeting and planning. The inputs can be *i)* the desired goal pose (region) in the workspace for the base and the end-effectors using GUIs (high-level), *ii)* the CoM velocity, the base rotational velocity, the end-effector motion, the user's desired footstep contacts, or whole-body motion (low-level). While at a low-level, the user is in charge of the obstacle avoidance, the planning and retargeting at the high-level deals with finding the path leading to the desired goal. The output of the high-level goes to the lower level in order to finally compute the robot reference joint angles, footstep contacts and timings as the output of the block. We structured the rest of this subsection according to the input category.

*1) GUI-based Path Planner:* When the user provides the goal region of the robot base or arms through GUIs, the humanoid robot not only should plan its footsteps or arm motion but also should find a *feasible* path for reaching the goal, if any exists. Related to the footsteps, contrary to wheeled mobile robots, a feasible path here refers to an obstacle-free path or one where the humanoid robot can traverse the obstacles by stepping over. Primary approaches for solving the high-level path planning are search-based methods and reactive methods. The first step toward finding a feasible path, i.e., regions where the robot can move, is to perceive the workspace by means of the robot perception system. Following the identification of the workspace, a feasible path is planned.

Search-based algorithms try to find a path from the starting point to the goal region by searching a graph. The graph can be made using a grid map of the environment or by a random sampling of the environment. Some of the methods used in the literature to perform footstep planning are A\* [62], D\* Lite [63], RRT variations [64], and dynamic programming techniques [65]. The *completeness*, *global optimality*, and the ability of real-time *replanning* of the path in dynamic environments are the important features of these search-based

algorithms when selecting a proper method. These algorithms are not very efficient for real-time execution when an exhaustive search is done, hence, to enhance the efficiency a *heuristic* is chosen in order to prune the search space and perform a greedy search.

Reactive methods for high-level kinematic retargeting and path planning problems can be addressed as an optimization problem or as a dynamical system. In [66] the problem has been tackled with a Model Predictive Control (MPC) approach. It allows finding the foot poses as a continuous decision-making problem by formulating it as a Quadratic Programming (QP) optimization problem. However, when the end-effector rotation or the obstacle-avoidance is added to the problem, the optimization problem becomes non-convex, therefore, there is no guarantee on completeness and global optimality [67]. However, although there are approaches to relax the non-convex optimization problems, their computational complexity is still a challenge [67]. Moreover, the problem of path planning and obstacle avoidance for a humanoid robot can be viewed as a simplified dynamical system control approach, for example, by using potential fields [68].

*2) Retargeting & Motion Generation:* Given the continuous measurements from the teleoperation devices, we can divide the retargeting approaches into three groups: *lower-body footstep motion generation*, *upper-body retargeting*, and *whole-body retargeting*.

*Lower-body footstep motion generation:* The role of this block is to plan the footstep motion and find the sequence of foot locations and timings given the CoM position or velocity, and the floating base orientation or angular velocity provided by teleoperation devices. One possible approach to address this problem is based on the instantaneous capture point [57]. Given the reference CoM position and velocity, one can compute the next foot contact point using the capture point relation in (3). In simple cases, the contact sequences can be preliminarily identified by the user for example by a finite state machine, and the desired footstep locations are modified to the left or right side of the capture point according to the nature of the foot contact (left or right). Another way to find the sequence of footsteps is to formulate an optimization problem, where the cost function is decided based on commonsense heuristics. The footstep timing can be found from the CoM velocity such that the total gait cycle duration corresponds to the average CoM velocity and gait length. This approach uses minimal information to generate footstep motion and does not enforce the user and the robot motion similarity.

*Upper-body retargeting:* In this method, the retargeting is done either in task space or configuration space. In the task space retargeting, the Cartesian pose (or velocity) of some human limbs is mapped to corresponding values for the robot limbs. Later, the inverse kinematic problem is solved by minimizing a cost function on the basis of the robot model while considering the robot constraints [2]. Different authors considered disparate limbs as the target of the mapping. A popular approach is to map the motion of the human wrist to that of the robot end effectors [28], [36]. A commonly used mapping in the literature is to perform an identity map between the rotational motion of the human and the robot,

whereas in the case of translational motion a fixed gain (due to differences in the geometry) is used [28]. A more complicated approach may identify this rotational and translational gain as a function of the human intention and ongoing task. To enhance the similarity of the robot motion to the human one, the retargeting of the elbow motion with a lower priority in the optimization problem is suggested in [69]. To overcome the manual morphing problem, [70] suggested solving an optimization problem to identify geometric parameters of the morphing function.

Configuration space retargeting refers to the mapping in the joint space from human to robot. This morphing is normally used when the human and robot have similar joint orders. In this technique, human measurements and model are used in an inverse kinematics problem. The joint angles and velocity of the human joints are identified and mapped to the corresponding joints of the robot [69]. When the human joint ranges differs from those of a humanoid, the robot constraints should also be applied in the morphing function.

*Whole-body retargeting:* This method measures the whole-body motion of human and kinematically retargets to the robot motion, similar to the upper-body retargeting approaches. Given the environment information, this approach may consider the contact constraints in the retargeting phase. Thus, this approach yields footstep locations, sequences, and timings. Later, outputs of this technique are provided to the *stabilizer*, to deliberate on the feasibility and enhance the robot stability [36]. In [3], authors measured the normalized ground projection of human CoM from an arbitrary foot and retargeted it to the equivalent robot CoM ground projection, on a line connecting the two robot feet. This approach can be extended toward retargeting the heel-to-toe motion (orthogonal to the feet line) and multi-contact scenarios [71].

### C. Stabilizer

The main goal of the *stabilizer* is to implement a control policy that *dynamically adapt* input references to enhance the stability and balance of the centroidal dynamics of the robot. Because of the complexity of robot dynamics, classical approaches are limited to examine the stability of a closed-loop control system. Therefore, other insights such as ZMP criteria and DCM dynamics are tailored in order to evaluate how far the robot is about to fall [72]. The *stabilizer* gets inputs from the *retargeting & planning* level, as shown in Fig. 3. However, these reference trajectories may destabilize robot's behavior, therefore the *stabilizer* adapts those references based on different criteria. Accordingly, the output of the *stabilizer* are references for the CoM position, the end-effector poses, joint angles, contact points, contact wrenches, and/or the rate of change of the momenta. Next, we provide an overview of stabilization approaches according to different criteria adopted so far in the literature.

1) *ZMP approach:* This approach is based on the idea that the robot's ZMP should remain inside of the support polygon of the base [56], [72], as shown in Fig. 4. Given the footstep trajectories provided by the kinematic retargeting, this approach computes the desired ZMP. While walking, in

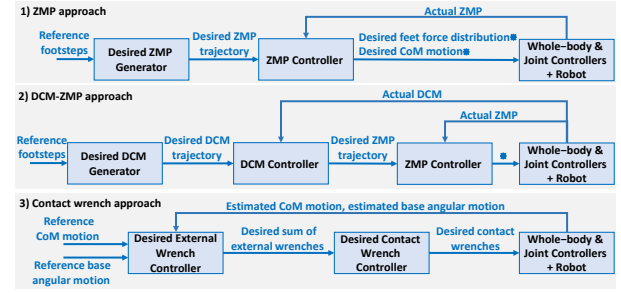


Fig. 4: Different stability criteria approaches.

single support the desired ZMP remains in the middle of the support polygon (e.g., left foot), and during double support the desired ZMP moves smoothly from the previous support leg to the middle of the new support polygon (e.g., from the left foot to the right foot). Following that, a ZMP controller makes sure that the real ZMP tracks the desired ZMP trajectory. However, the ZMP controller introduces several limitations. If the robot is subject to high disturbances, it does not adapt online to footstep locations to avoid the robot from falling. Moreover, this approach hardly extends to non-flat terrain or when the support foot rotates or slides.

2) *DCM-ZMP approach:* Another approach that is often employed with the simplified model retargeting and control is the DCM [58], which can be viewed as an extension of the capture point concept to the three-dimensional case for uneven terrain [59]. Eq. (4) shows that the DCM dynamics is unstable, i.e., it has a strictly positive eigenvalue, and using equation of LIPM (2) it can be shown that the CoM dynamics converges to the DCM value. Therefore, the main goal of this controller is to implement a control law that stabilizes the unstable DCM dynamics (4) as well as the ZMP as mentioned previously. To stick to the formulation presented in Sec. III-A, the DCM-ZMP dynamic retargeting is explained with flat floor assumption; however, for the extension of the controller to 3D, one can refer to [59]. The block diagram of the DCM-ZMP stabilizer is also shown in Fig. 4. DCM-ZMP approach has been deployed for whole-body retargeting in [36]. In order to enhance the stability of the system, the authors have introduced the predicted support region where the time-delayed DCM of the robot is kept inside that region.

3) *Contact wrench approach:* This approach finds the desired contact wrenches at each contact point such that it enhances the robot stability. A schematic block diagram of this controller is shown in Fig. 4. This controller has been initially introduced in [73] and later considered by [74] as a stability augmentation criteria when introducing the contact constraints. This approach is similar to the momentum-based controller and will be explained in more detail later. However, to be thorough, we have mentioned it here as well.

### D. Whole-Body Control Layer (WBC)

The *WBC* block in Fig. 3 gets as input the retargeted human references corrected by the *stabilizer*, and provides as an output the robot joint angles, velocities, accelerations, and/or the joint torques. The whole-body control problem can be formalized with different cost functions and be solved as a QP



problem or other approaches such as Linear Quadratic Regulator (LQR) [75], [76] and MPC [77], [78]. In the following, different whole-body control approaches are presented.

1) *Whole-body inverse instantaneous velocity kinematics control*: The problem of inverse instantaneous or velocity kinematics is to find the configuration state velocity vector  $\nu(t)$  for a given set of task space velocities using the Jacobian relation. One common approach is to formalize the controller as a constrained QP problem with inequality and equality constraints. Conventional solutions of redundant inverse kinematics are founded on the pseudo-inverse of the Jacobian matrix [79].

2) *Whole-body inverse kinematics control*: The problem of Inverse Kinematics (IK) is to find the configuration space vector  $q(t)$  given the reference task space poses. This problem can be sometimes solved analytically for determined robots; however, this solution is not scalable to different architectures and for a redundant humanoid robot (with a high degree of freedom), it is not feasible. Differently from inverse instantaneous velocity kinematics, to solve the IK problem a non-linear constrained optimization problem is defined using the geometrical forward kinematics relation. Solving this problem might be time-consuming and the results can be discontinuous as well. To solve these challenges, a common approach in the literature is to transform the whole-body IK into a whole-body inverse velocity kinematics problem.

3) *Whole-body inverse dynamics control*: The Inverse Dynamics (ID) refers to the problem of finding the joint torques of the robot to achieve the desired motion given the robot's constraints such as joint torque limits and feasible contacts. There are different techniques to solve the ID problem in the literature [80]. To reach the desired end-effector poses and contact wrenches, one can formulate the ID problem as an optimization problem, minimizing the error on metrics such as the motion tasks and contact wrenches. Some of the constraints are dynamics equation of motion (1), joint angle, velocity, acceleration, or torque limits, collision avoidance constraints, and non-sliding contact constraints. The details of the constraints can be found in [40], [80].

4) *Momentum-based control*: This controller finds the configuration space acceleration and ground reaction forces such that the robot follows the given desired rate of change of whole-body momentum [81], [54]. According to the Newton-Euler's laws of motion and (1), the rate of change of centroidal momentum (the whole-body momenta of the humanoid robot about the CoM) is equal to the sum of all external wrenches applied to the robot. Using that, one can write the momentum-based controller in a QP fashion [82]. Eventually, external wrenches and joint accelerations are used with an inverse dynamics algorithm to compute the robot joint torques.

#### E. Low-level Joint Controller

One of the main challenges in deploying a humanoid robot controller is the different behavior obtained in simulation and on the real robot of the low-level joint torque tracking. The low-level joint controller is in charge of generating motor commands to ensure the tracking of the higher-level control

inputs. The input to the joint controller can be the desired joint positions, velocities, torques, or a mixture of them. The output of this controller is the current or voltage to the motors driving the joints. While the joint position can be directly measured by the encoder sensors, the joint velocity feedback is obtained by differentiating in time the encoder values, hence it can be noisy. Joint torque sensors or whole-body estimation algorithms equipped with distributed joint-torque sensors can estimate the generated joint torques. The torque control of the robot joints is more challenging because of the dynamics of the joints which are subject to friction and the mechanical power transmission from the motor to the joint shaft. On the other hand, compliance control of the joints allows for more robust locomotion and a safer interaction of the robot with the humans and the environment. For example, ankle joint torque control allows for conforming the foot to the ground in case of small obstacles or a mismatch between the ground slope and its estimation. However, the compliance control with only joint torques may lead to poor velocity or position tracking, therefore a mixture of them is proposed in the feedback and feed-forward terms of the joint controller in [30], [76], [54].

#### F. State Estimator, Localization & Mapping

These blocks in Fig. 3 receive measurements from the robot sensors and estimate the necessary information for other blocks in Fig. 3 or to the human as shown in Fig. 2 (for assisted teleoperation). A family of well-known model-based estimation techniques commonly used in robotics is the Kalman filters. The joint encoders are used to estimate the joint angles, velocities, and acceleration. These joint states accompanied by the robot dynamics model in Eq. (1) and force/torque sensors can be used to estimate joint torques and external forces [83]. To estimate the joint torques, the actuation system model can be learned or identified as well [84]. If a humanoid robot is equipped with tactile sensors, the external wrenches and its point of application can be estimated likewise [85]. To estimate the ground reaction wrenches and ZMP of a humanoid, its feet are normally equipped with force/torque sensors [86]. Moreover, a combination of proprioceptive and exteroceptive sensory information allows a legged robot to estimate better the terrain characteristics and eventually elevate the control robustness [87]. Finally, calculated joint angles and velocities are used to estimate the robot link poses and velocities through the forward kinematics. Combining those values with the robot link inertia, one can compute the CoM position and velocity [76]. Moreover, considering the uncertainties of the link inertia and joint measurement noises, CoP and force/torque sensors are used to estimate CoM in [88].

One of the challenges specific to humanoid robots is the estimation of the robot base, and eventually robot localization in an environment. For the estimation, either proprioceptive sensors (i.e., joint encoders and IMU sensor attached to a robot link) or a combination of proprioceptive and exteroceptive sensors (such as cameras, GPS) are employed. When only proprioceptive sensors are used, a common approach is to assume that at each time instant at least one of the robot links is in contact with the ground, therefore the frame attached to

the contact link has zero velocity and no slippage. With this assumption and taking into account the floating base frame pose in the kinematic chain, one can estimate the floating base velocity given the joint velocity vector, and eventually the base pose by integrating those velocities. However, the error of estimation is propagated over time due to kinematic modeling errors. To enhance the accuracy and limit the uncertainty of state estimation endowed with the odometry, IMU measurements are fused in the estimation process [89], [90]. Nevertheless, exploiting only the proprioceptive sensors does not lead to observability in the yaw axis (parallel to gravity vector) and the absolute position of the robot [90]; thereby exteroceptive sensors can facilitate to overcome this difficulty. Exteroceptive data such as camera information allows finding feasible regions for the humanoid robot foot locations as well as a map of the environment and obstacles.

#### G. Challenges & Future Directions for Retargeting & Control

Humanoid robot teleoperation is a new field, and many challenges to put together whole-body coordinated motion retargeting, planning, stability, and control are not addressed effectively yet. For example, dividing the retargeting and planning problems of humanoid robot teleoperations appears to be useful but not effective in performing agile teleoperation tasks outside of the lab (in the real world) and in unstructured environment as it is the case for many hazardous environments and disaster response scenarios. Many of the techniques adopted so far use simplified models of humanoid robots. Although these approaches are computationally efficient, they are limiting due to the adopted simplifying assumptions, e.g., fixed robot CoM height, ignorance of human or robot rotational motion, the existence of at least one contact point with no slippage between the robot and environment.

An alternative approach to solving simultaneously whole-body coordinated retargeting and planning problems is using the MPC technique and defining them as optimization problems with equality and inequality constraints. However, they are non-linear and non-convex optimization problems with large input and state spaces; therefore, they are computationally demanding, and the optimization may suffer from local minima. This issue intensified when retargeting the rotational motion and angular momentum of the human to the robot, while biomechanical studies show their importance as one major underlying component of the human-like coordinated motion [91]. To this goal, an *angular excursion* index has been introduced by [92] relating the whole-body orientation, angular velocity, and centroidal angular momentum. When this idea integrates with linear momentum, it enables highly agile motion with considerable rotational behaviors [93], [92]. However, retargeting the human angular excursion to the robot is an open problem. A possible trade-off solution between simplified and whole-body motion retargeting and planning approaches is to adopt the centroidal dynamics, the terrain map, and whole-body kinematics of the robot given the human measurements [94], [95]. This approach benefits from centroidal dynamics constraints and whole-body kinematics in collision avoidance and reachability computations. However, to adopt MPC in

humanoid teleoperation scenarios, major challenges remain to predict future human motion and the difference between the surrounding terrain of human and robot.

While the MPC approach can be a valid solution for the humanoid robot teleoperation, it is not sufficient when deploying the robot in the real world to perform various tasks in an unstructured and dynamic environment with varying compliance and slippery characteristics. An approach to resolve these challenges is to introduce new sets of manual heuristics and constraints to the optimization problem for every scenario and environment. However, it is tedious and inefficient to generalize over different situations. To overcome these drawbacks, data-driven approaches with special attention to the robot dynamics and stability indications have shown promising results in the learning and mobile robot communities. [96], [97], [98] are some of the successful examples of leveraging neural networks and reinforcement learning (RL) techniques. In retargeting problems, safety considerations can be explicitly enforced as well [99]. A novel approach blending an unsupervised learning technique with forward kinematics is proposed by [100]. It relies on the cycle consistency principle, i.e., motions retargeted to the avatar should generate the original motions of the human when retargeted back. The general trend in whole-body planning of robots is that many robots simultaneously learn how to perform agile and dynamic motion robustly in a simulated environment, by incrementally introducing new complex terrain difficulties, dynamic obstacles, and terrain with different parameters. To relax the simulation to real-world gap, [84] suggested learning robot actuator dynamics from the real robot and incorporating them with simulated robots. However, none of those works considered at the same time the whole-body coordinated motion retargeting of a human to a humanoid robot. We speculate that adding a reward term in an RL problem or using transfer learning techniques can impose the similarity of the human and humanoid robot motion during the training phase. While in some cases, human motion can be tracked by the humanoid robot, in other cases, the humanoid robot may perform step adjustments to keep its balance, for example, based on some behavioral selection techniques [95].

## IV. ASSISTED TELEOPERATION

In many teleoperation scenarios, controlling the robot as explained in the previous section, is not the only viable solution. In fact, many tasks can achieve higher performances by sharing the autonomy among the human and robot. This section provides more details about these assisted teleoperation strategies.

#### A. Shared Control

Delegating robot's full control to the human operator's experience can often limit the efficiency of the teleoperation, resulting in clunky motions, failures, or numerous attempts before being able to accomplish a given task. This applies particularly to humanoid robots where the operator has to control many aspects at once via teleoperation (e.g., the pose of both hands, the feet location, balance) and can fail very easily without significant robot autonomy being used

simultaneously. In shared-control teleoperation, some robot autonomy is used to assist the user in accomplishing the desired task, potentially making teleoperation easier and more seamless. Generally, the operator's input is modified according to specific metrics by sharing the control authority between the robot and the operator to enhance performance or safety [101]. For example, in [102], Rakita *et al.* teleoperated the upper-body of a humanoid using a shared-control approach, providing on-the-fly assistance to help the user complete tasks more easily, enhancing the end-effectors control while performing bi-manipulation of objects. Similarly, in [103], Rahal *et al.* designed a shared control approach to assist the human operator by enforcing different nonholonomic-like constraints representative of the cutting kinematics. In other shared-control approaches, the user provides an input  $\mathbf{u}$ , which enables the robot to predict human intent, and assist her/him in the task by adjusting the motion or by executing a pre-optimized version of that motion [101]. Then, a blending policy arbitrates the user input  $\mathbf{u}$  and the enhanced robot motion  $\mathbf{r}$ , determining the final reference:

$$\mathbf{u}^* = (1 - \alpha)\mathbf{u} + \alpha\mathbf{r}, \quad (5)$$

where  $\alpha$  can be any scalar function. A common choice is the confidence in the prediction of the user intent:

$$\alpha = \max(0, 1 - d/D), \quad (6)$$

where  $d$  the distance to the goal and  $D$  some threshold past which the confidence is 0. In this case the closer the robot gets to a predicted goal, the more likely that this goal is the correct one, and the input  $\mathbf{r}$  is preferred over  $\mathbf{u}$ . The prediction of the user intent has also been successfully used to provide haptic guidance through a master device to teleoperate a robot manipulator [104], and could be applied to humanoid robots by using exoskeletons as input devices. The haptic information can also be used to enhance the user's comfort during teleoperation [105].

### B. Supervised and Safeguarded Teleoperation

When full robot autonomy is available for a given task, the operator can simply act as a supervisor. By monitoring the robot, the operator can then identify and react to unexpected problems and intervene in a timely manner by controlling the robot directly to handle "uncovered" situations. This was a common approach in the DRC, where team operators could guide the robot to achieve complex tasks through failure when needed [30]. Similarly, in [106], the user monitored multiple robots interacting with passersby in a shopping mall. The robots performed their own speech, gesture, and motion planning autonomously, and the role of the human was only to provide occasional sensor inputs. In rare cases, an operator had to control the robot directly to handle unexpected questions from a customer or to re-plan the robot's path to avoid unmodeled obstacles.

A mirrored approach can also be adopted when teleoperating robots. For example, in [107], the operators shared control with a safeguarding system onboard the robot. In

benign situations, the operator had full control of the vehicle's motion, while in hazardous situations, the safeguarder modified or overrode the operator's commands to maintain safety. The safeguarder, therefore, exhibits many characteristics of autonomous systems, such as perception and command generation. In [107], the safeguarder not only had the function of preventing collision and rollover but also monitoring the system's health (e.g., vehicle power, motor stall). A similar approach could be used for humanoid robots to prevent them from falling or reaching singular configurations while being operated by the user.

### C. Bilateral Teleoperation

Bilateral teleoperation techniques have been widely used in the literature for robot manipulators. In this approach, not only the robot receives kinodynamic references from the human, but the operator also receives kinesthetic feedback (force, pressure, vibration, etc) from the robot. This feedback informs the operator about the robot's performance reproducing the commanded motion or about external disturbances applied to the machine. The approaches described next focus on teleoperating the upper- or whole-body, which in the latter case couples altogether the human operator and the robot at upper- and lower-body levels.

1) *Upper-body bilateral teleoperation:* A simpler strategy that has been adopted on humanoid robots consists in teleoperating the upper-body in a bilateral fashion, while using a separate balancing controller on the lower-body to regulate balance [108], [109]. To control the upper limbs of the HRP-2, Peer *et al.* [109] adopted a common control scheme for position-controlled robots: the *admittance* controller. Such controllers define the reference through the differential equation of a virtual mass-spring-damper system. Under time delay, the parameters of the controller should be selected appropriately to guarantee stability of the overall teleoperation system, as done in [110] while teleoperating the HRP-2 located in Tsukuba (Japan) from Munich (Germany).

2) *Whole-body bilateral teleoperation:* An extension of the conventional idea of bilateral teleoperation is starting to be investigated to dynamically couple human and robot at a whole-body level [1]. This strategy consists of mapping the whole-body kinematic (joints position, velocity, etc) and dynamic (contact forces, joint torques, etc) references from humans to robots, while providing the operator with feedback regarding the robot's whole-body dynamics, as shown in Fig. 5. However, the naturally unstable dynamics of humanoid robots poses an additional challenge to the whole-body teleoperation: the robot must balance while reproducing human movement.

The strategies for whole-body bilateral teleoperation utilize kinesthetic feedback to inform the operator about the robot's dynamics and stability in real-time. The strategy usually focuses on the CoM dynamics, and other condensed information about the robot. For instance, the cable driven feedback interface in [41] exerts forces on the demonstrator's waist corresponding to the state of the robot's CoM. This feedback allows the human to teach the robot how to compliantly interact with the environment. In [49], the feedback force

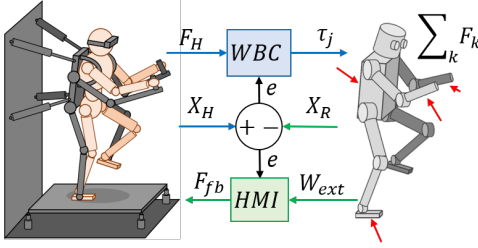


Fig. 5: General concept for whole-body bilateral teleoperation. The robot WBC computes joint torques  $\tau_j$  using the reference interaction forces  $F_H$  from the operator and the error  $e$  between the human state  $X_H$  and the robot state  $X_R$ . The external contact forces  $F_k$  applied to the robot create a net wrench  $W_{ext}$ , which is used by the human-machine interface (HMI) to compute a proportional kinesthetic feedback  $F_{fb}$  to the operator.

applied to the human's torso is proportional to how close the robot is from tipping over. This is estimated by considering the distance between the robot's CoP and the edge of the support polygon. The closer the robot is from tipping over in one direction, the larger the feedback force applied to the operator in the opposite direction. A similar strategy is used in [39] by providing discrete vibration levels to the operator using a belt with vibrotactile feedback. In [4], the force feedback device TABLIS, a powered exoskeleton, applies forces to the operator's feet to indicate that the robot is stepping onto an obstacle. This enables the operator to control the robot to navigate over objects. Finally, in [27] the force feedback is utilized to dynamically synchronize human and machine. The force feedback generates drag (negative feedback) if the robot cannot keep up with the operator's movement, or it speeds up human motion (positive feedback) if the robot moves faster than the operator. The high-level expression for the force feedback is given by:

$$\mathbf{f}_{fb} = k_H [(\dot{\mathbf{x}}'_R - \dot{\mathbf{x}}'_H) + \mathbf{f}'_{ext}], \quad (7)$$

where  $\dot{\mathbf{x}}'_i$  is the dimensionless CoM velocity of the human ( $H$ ) and robot ( $R$ ) [57],  $\mathbf{f}'_{ext}$  is dimensionless external force vector applied to the robot, and  $k_H$  is a scaling factor proportional to the operator's size and body mass. This strategy enables human and robot to dynamically take simultaneous steps.

In general, during whole-body bilateral teleoperation, the kinesthetic feedback provided to the operator is proportional to some kinematic or dynamic discrepancy between human and robot with respect to the task at hand and/or the balance regulation. The assisted teleoperation principle arises from the fact that the robot must prioritize between following the human motion command to perform a task and maintain its own bipedal stability. Some strategies shift the balancing authority to the robot's autonomous controller. This means that the robot is responsible for predicting if a given command will jeopardize stability and deciding the best course of action to override the motion. For instance, in [36], the robot's controller uses stability considerations from the LIP model to modify the CoM trajectory commanded by the human, preventing the operator from destabilizing the robot. This represents a more conservative approach that guarantees stability of the movement of the robot, but prevents the operator from exploring motions that would go beyond the boundaries of the stability metric employed. In contrast, other approaches, such as [111], rely

on the human operator to actively regulate the robot's balance. In this sense, the robot follows the human's movement with only minor regulation from the robot's autonomous controller; then, the operator must perceive the robot's destabilization through the kinesthetic feedback and mitigate it by adapting the commanded movement. Although riskier, this strategy frees the operator from abiding to the boundaries of predefined stability metrics, which are challenging to be mathematically defined in the context of legged robots. The goal of this approach is to eventually allow the operator to learn how to cope with the robot's dynamics and to create new motions on the fly.

#### D. Impedance Control

A similar concept to admittance control can be employed in non-bilateral systems to provide the robot the ability to dexterously interact with the remote environment. In [108], Brygo *et al.* implement an autonomous *impedance* controller that regulates the robot's joints stiffness and damping according to the manipulation loading conditions. Also in this case a virtual mass-spring-damper system is employed, but the main difference between admittance control and impedance control is that the former controls motion after a force is measured, while the latter controls force after motion (or deviation from a set point) is measured [112]. In [108], the COMAN robot performs free-space motions with compliant limbs to ensure safe interactions during unforeseen collisions on the whole-arm. During the manipulation task, the controller stiffens the humanoid arm joints according to the external force sensed at the end-effector, permitting to handle the task loads.

Alternatively, the impedance profile can be sent to the robot through a BMI applied to the operator's arm, using for example non-intrusive position and EMG measurements, technique also known as *tele-impedance* [113].

### V. COMMUNICATION CHANNEL

In teleoperation, human and humanoid robot transmit information through a *communication channel* [114]. However, the communication channel introduces complexities that impact the stability and performance of the teleoperation, namely, transmission delays, and distortion of the information.

#### A. Transmission delays

The interest for the transmission delays in teleoperation emerged back in the 60s. The first research to consider the effect of these transmission delays on the performance of the teleoperation was published by Ferrel [115]. He realized that when a delay is inserted most operators adopt an effective *move-and-wait* strategy. However, as he did not consider force reflection there was no problem of instability. Only when force reflection was considered, instability became apparent.

As a consequence, most of the early research on teleoperation targeting delays was based on supervisory and software-based teleoperation, without *closing the loop* in force. Some examples are the employment of high-level commands and stored subroutines, or a predictive display that gives the operator a prediction of the system response [114].

Although old, these techniques are still in use when dealing with extremely complex robots, as humanoid robots are. As an example, the robots that participated during the DRC had to deal with a communication channel represented by two links: (a) a bi-directional low bandwidth (9600 bps) continuous communication path, or (b) a unidirectional high bandwidth (300 Mbps) communication path with intermittent connection. The purpose was to boost the autonomy of the robots. As a result, most of the teams employed high-level commands, visual aids, and no bilateral teleoperation [53]. The effect of the “*move-and-wait*” strategy was apparent, and greatly influenced the operation speed.

A real sense of telepresence can only be achieved by kinesthetically coupling the operator to the remote environment by introducing force reflection [18]. However, time delay can represent a destabilizing factor unless the motion bandwidth is severely reduced or more advanced solutions are considered. Shared-control solutions can also be adopted to handle communication delays [116]. Even round-trip delays around 100 ms (typical of the Internet) in fact, can induce the instability of the teleoperation system either indirectly, due to human operator’s overcompensations of the delayed perceived errors [117], or directly in the control law in the case of bilateral systems.

Advanced techniques to deal with delays appeared in the mid 80s. Particularly, using *network theory* and the concept of passivity allowed to achieve stable time-delayed teleoperation assuming constant-time delays [118] [119]. The *constant-time delay* assumption is limited and valid for simple communication channels. Packet switched networks, as it is the case of Internet, introduce additional difficulties: (a) random varying delays that can reach very high values, (b) a discrete-time exchange of information, (c) the effect of quantization, and (d) loss of packets [18]. Varying-time delays appear due to several factors like traffic congestion, bandwidth, and information loss. The latter mainly caused by transmission time-outs, transmission errors, and a limited buffer size. To deal with the varying-time delay, one can estimate an upper bound of the delay through networks statistics, then use it to dissipate energy [120], to emulate a virtual constant-time delay [121], or to use it to extend the horizon of a model predictive control [114].

### B. Distortion

Distortion is another effect introduced by the communication channel. In case of packet switched networks, a delay of discrete-time velocity information results in a distortion of velocity and position drift, hence, degrading the performance. Solutions to this problem can include the exchange of position information together with the velocity, or time-stamping the information [122]. Another source of distortion is due to the policies introduced when there is information loss: (a) to treat the lost packet as a null packet, (b) to use the last valid packet, or (c) to use interpolation [18].

### C. Network-based Model

The standard model of the communication channel is based on network theory. By using an analogy between mechanical

and electrical systems, we can represent the teleoperation system as a network of interconnected  $n$ -ports. An  $n$ -port is characterized by the relationship between *effort*  $f$  (voltage or force) and *flow*  $v$  (current or velocity).

Assuming that we have flows and efforts on both sides of a 2-port network representing a Linear Time-Invariant (LTI) system, the relationship between them can be written (in the frequency domain) by a hybrid matrix. The idea is to use control to modify the characteristics of this matrix in order to overcome the difficulties imposed by the communication channel [18]. In the topic of bilateral teleoperation represented by a nonlinear system, which is the case for humanoid robots, the frequency-domain approach is no longer valid. However, Anderson [118] demonstrated that by using a Hilbert network, with efforts and flows belonging to a Hilbert space, equivalent analyses could be performed.

### D. Passivity

The passivity formalism provides a simple and robust tool to analyze the stability of a nonlinear system [119]. A passive system may dissipate energy ( $E(t)$ ) but it cannot increase its total energy [123]. For an  $n$ -port, we can take forces as inputs and velocities as outputs, and define the *power* (not necessarily a physical one) entering to the system as the scalar product between input and output ( $P(t) = f^T(t)v(t)$ ). Then, for an  $n$ -port to be passive, the power is either stored or dissipated:

$$\int_0^t f^T(\tau)v(\tau)d\tau = E(t) - E(0) + \int_0^t P_{\text{diss}}d\tau \geq -E(0), \quad (8)$$

where  $P_{\text{diss}}$  is a non-negative power dissipation function. If no power is dissipated, the  $n$ -port is lossless [119].

An important tool that can be used to analyze the passivity is *scattering theory*, which relates the effort and flow of a network through a *scattering operator* [123], [18]. In this respect, a passivity-based system imitates a physical system that obeys energy conservation [119]. An important property of passive systems is that the stored energy and power dissipation of a combined system is equal to the sum of individual stored energies for each system [119]. Therefore, if we assume that the operator and the environment behave passively, the entire system can behave passively if each  $n$ -port is passive.

### E. Stability of Time-delayed Humanoid Robot Teleoperation

Passivity and stability are related as a result of considering the expression of the stored energy as a Lyapunov function [119]. If the  $n$ -port corresponding to the communication channel is simply determined by a delay, then it can be demonstrated that a pure delay introduces power [123]. However, it is possible to modify the behavior of the communication channel by using control. A naive solution is to add damping, but there is no stability guarantee, and it affects the performance [119]. Alternatively, Niemeyer et al [119] used the power and energy to define the concept of *wave variables*, where input and output of wave variables are a combination of efforts and flows. Then, a system is passive if the energy of the output waves is less than the energy of input waves. In a recent work by [124] on bilateral teleoperation of legged



robot mobile manipulators, to ensure the passivity and stability they used *energy tanks* as energy observers and included that as a constraint to the robot whole-body control layer. The problem is that achieving passivity does not guarantee good performance [122], and stability and transparency are conflicting objectives [18].

More specifically to the humanoid robot's unilateral and bilateral teleoperation, the delay is relevant to the robot's balance and velocity. These challenges have not yet been addressed in the literature. However, in many applications the delay is low, hence does not affect highly the robot balance through teleoperation. In the case of large delays, a higher autonomy level for teleoperation is required. In the case of mid-range delay, a commonly used strategy is *move-and-wait* to overcome the delay problem, however, this reduces the robot velocity and efficiency in task performance. In the case of bilateral teleoperation and manipulation tasks, the use of this approach is highly limited. With the rise of machine-learning techniques, an envisioning approach to overcome the delay is to use predictive approaches, both from the human and the robot side. On one hand, the human motion should be predicted and mapped to the robot, and on the other hand, the robot motion and interaction forces with the environment, as well as the stream of the visual feedback, should be predicted and sent to the human operator. Finally, to speculate the humanoid robot's balance through teleoperation with a lower level of autonomy, we may consider the human operator, and retargeting and control strategy, all together, as a complex control system to balance the humanoid robot. In this case, disturbances on the robot can be controlled through the whole teleoperation pipeline and human commands. By speculating on the simplified robot model, i.e., LIPM, studies shows that if the time delay of the system is higher than a critical time delay, the robot destabilizes [125]. The critical time delay  $\tau_c$  is relevant to the robot CoM height, i.e.,  $\tau_c = \beta \sqrt{\frac{z_0}{g}}$  from Eq. (2). The lower the CoM height, the lower becomes  $\tau_c$ . This is related to the fact that a lower CoM height implies faster dynamics of the LIPM, hence a lower time delay is tolerated to keep the humanoid robot's balance. Here  $\beta$  can be related to different parameters such as the human performance, the retargeting and control strategy, and the robot actuators.

## VI. DESIGN & EVALUATION OF HUMANOID TELEOPERATION SYSTEM

In order to deploy the teleoperation system for *real users*, such systems should meet the users' needs and be evaluated prior to deployment. These needs and considerations should be anticipated in the design and development phase. A human-centered design, consisting in an iterative process aiming at properly allocating the tasks (functions) between the user and the teleoperation system, is required. Designers need metrics to evaluate the teleoperation systems, to share the knowledge, and compare their findings. These metrics are highly task-dependant, i.e., various tasks impose different functional requirements that the system should meet [126], [19]. These metrics not only determine inherent problems and limitations of the humanoid robot teleoperation system, but also provide

guidelines for the design and development, reducing the cost and the time to design such a system. The rest of this section adopts some metrics proposed in relevant fields that help design and evaluate humanoid robot teleoperation systems.

### A. Evaluation Metrics

1) *Usability Assessment*: According to ISO 9241 [127], *usability* is defined as “*the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use*”. In line with this definition and the teleoperation context, effectiveness is the degree of accuracy and completeness in which the user reaches the teleoperation goal, whereas efficiency is related to the resources being used (e.g., time, cost, human effort, etc.) with respect to the achieved outcome [127]. Effectiveness and efficiency are considered as the objective usability measures, which depend on the teleoperation task. On the other hand, satisfaction determines the degree in which the users' needs and expectations are met as a result of the use of the teleoperation system [127]. It is determined according to the user's emotional, physical, and cognitive responses. The user's perceived usability, which is a subjective measure, has a direct relation with the user's satisfaction; the more the perceived usability, the higher the satisfaction [128]. Moreover, usability assessment is relevant to the individual user in terms of the frequency of use and familiarity with the system. Some criteria to measure the effectiveness of a task execution are the task completeness, objective achievement, and task errors [129]. The efficiency of a task execution is established according to the task time, cost-effectiveness, energy consumption, productive time ratio, and unnecessary actions (just a few to mention) [129]. For the subjective measure of the usability there are some standard empirical and informal techniques to examine the user's perceived usability including concurrent think aloud, retrospective think aloud, concurrent probing, and retrospective probing. Among these, the most famous one is the System Usability Scale (SUS) [130] retrospective probing technique which assesses two factors including usability as well as learnability for a variety of tasks. Nevertheless, a subjective measure of the usability is in relation with the user perception; therefore, it is affected by the situational awareness, which will be discussed in the next section. An attempt to provide a holistic taxonomy of usability evaluation in robot teleoperation scenarios is done in [131].

2) *Situational Awareness (SA)*: SA correlates the user capabilities, training, experiences, preconceptions, and objectives with the ongoing task workload together [132]. Endsley identified *situation awareness* with three levels as “*the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*” [132]. Accordingly, SA tries to understand the process which leads to decision making, considering a variety of elements.

High SA promotes the probability of a good performance, and the poor performance of the task is normally the result of SA loss. SA is lost when the user's knowledge is incomplete

or inaccurate, the user attention is narrowed to some elements, or when the mental model of the user diverges from the reality [132], [133]. Loss of SA is also correlated to the workload; when the workload is high the user's attention is drawn from the main task and therefore the user does not give importance to the rest of the elements.

The main works in the literature measuring SA are based on the probe technique, physiological measurements, implicit inference, process indices, and observer ratings [132], [133]. Subjective ratings of the SA, where the users fill a form, are limiting because of the inaccurate evaluation of the users about their SA and the biased evaluation depending on the task outcome [132]. To assess SA, EEG and eye-tracking physiological signals are employed as well in the literature. SA can be inferred indirectly and implicitly through other related measurable metrics, like the task performance [133]. Finally, probe techniques can be both retrospective, after task execution, or concurrent, by collecting data with a questionnaire while executing the task. Among the concurrent ones, the Situation Awareness Global Assessment Technique (SAGAT) proposed by Endsley in [132] is a freeze-probe approach that measures SA objectively. This method interrupts the experiment at some random points and halts the user interface, then a number of questions are asked from the users to evaluate their knowledge about the current and future situation including the user's perception, comprehension, and actions. Later, the responses are compared with the real values collected from the scenario or from experts' responses. While SAGAT is very reliable, it is intrusive to the natural flow of the task execution.

3) *Workload*: Workload relates the resources demanded by a task to the available resources supplied by the human operator. There is no consensus about the definition of the workload [134]; however, [135] identifies the *workload* as "*the amount of work that is loaded on an individual, the time pressure in which a task is performed, the level of effort exerted, the success in meeting the task requirements, physiological and psychological*". Workload is related to the human operator's (subjective) experience in response to the task objectives. Especially in time critical decision making tasks, a high workload can lead to user errors or to a delay in the decision making.

Workload presents high variability depending on the human operator and the teleoperation task, hence being difficult to measure. In fact, the workload is multi-dimensional, and various subject- and task-related factors (such as mental, physical, information, perceptual, and communication loads) should be considered [134], [135]. Different approaches have been proposed in the literature for measuring the workload on the human operator, namely questionnaires and experts' reports (subjective), physiological techniques, and performance-based approaches [134]. Specifically, for a reliable mental workload assessment a mixture of them is suggested in [134]. Among the subjective measurement methods, the NASA Task Load Index (NASA-TLX) and the Subjective Workload Assessment Technique (SWAT) are the most famous multi dimensional subjective rating methods used in the literature [135], [136]. Continuous physiological measurements such as cardiac ac-

tivity, eye activity, respiratory activity, speech measures, and brain activity can provide a reliable assessment of the task physical or mental load as well [134]. To measure the physical workload, oxygen consumption estimation can directly provide an index, while heart rate can be used as an indirect method. Moreover, since the workload and task performance have a causal relationship, the task execution performance can play as another indicator of the workload, when time-shared or difficulty of tasks are exploited. An extensive review of the methods to measure the workload can be found in [134].

4) *Engagement, Immersion, Involvement, and Presence*: Some metrics concerning robot teleoperation, specifically related to the subjective experience of the user are engagement, immersion, flow, involvement, and presence. In different fields, engagement is defined and measured distinctly. In the context of teleoperation, the definition and measurement approaches of engagement are especially relevant in gaming and virtual reality applications. According to [137], engagement is "*the willingness to have emotions, affect, and thoughts directed towards and aroused by the mediated activity in order to achieve a specific objective*". It relies on the user's activity and expectations. The user is engaged when her/his perceptual, intellectual, and interactional expectations are met. In this context, *presence* is characterized as "*the subjective experience of being in one place or environment, even when one is physically situated in another*" [138]. Presence is a multifaceted concept that is related to *involvement*, a psychological state depending on attention to remote environment stimuli, and *immersion*, a psychological state of perceiving oneself as a part of the remote environment stimulus flow [138]. Here, stimulus flow is a dynamic stream of sensory information. There are several factors contributing toward the sense of presence including control, sensory information, distraction, and realism factors. More information about *presence* can be found in [138].

Even if presence is a subjective sensation, there are both subjective and objective methods to measure it. Subjective measures are based on self-report questionnaires like the ITC-sense of presence inventory [139] or the Witmer & Singer Presence Questionnaire [138]. For objective measures, both behavioral (e.g., startle response) and physiological measures (e.g., heart-rate changes, skin conductance/temperature) are suggested in the literature [140]. Similar methods have also been exploited to measure the user engagement [141].

## B. Interface Design and Human-centered Autonomy

The interface design impacts the user's decisions while teleoperating the humanoid robot, being the medium of communication between human and robot. Good design choices leverage the user experience and result in a successful teleoperation system. The diagnostic information provided by different measurement techniques can eventually enhance the design and the user support. Following the explanation of different metrics provided before, we can identify four types of measurement methods, namely, questionnaire-based subjective ratings, performance-based objective ratings, physiological measures, and behavioral responses. The subjective rating evaluation is quick, and can be done either online (while

performing the task) or retrospectively. Nevertheless, studies show that retrospective evaluation, while not being intrusive, may be subject to information loss because of the user's bias and poor subject ability to recall. Physiological measurements are partly related to the concept of *activation* or *arousal* in the field of psychophysiology, which relates the changes of mental effort and its effect on the physiological activity of the human operator [142]. Finally, all the metrics provided are interrelated; for example, [143] has studied the effect of SA and task difficulty level on the human performance and the user's sense of telepresence.

The described metrics can be used to identify the robot behaviors that should be automated. When designing a teleoperation system, one of the critical parameters that impacts the success of the system and the user experience is the autonomy level. It affects the SA, the user workload, the system usability and the user's sense of presence [144]. While high degree of automation degrades the SA, low autonomy in complex tasks intensifies the workload. High level of autonomy promotes the user's performance and diminishes the workload in the normal situations [145], while during system failures, a low level of autonomy increases the SA and enhances the human manual performance. Therefore an intermediate level of autonomy in which the user is in the control loop, i.e., human-centered autonomy, enhances the SA, reduces the workload, and improves the performance against failures [145].

### C. Humanoid Robot Teleoperation Design, Evaluation Challenges & Future Directions

The design of humanoid robots and teleoperation systems are inherently iterative, where the system goes under different evaluations. However, to improve the design, not many studies in the literature have been published to evaluate such systems systematically. Table II provides an overview of the works where different metrics are used to enhance and evaluate the teleoperation system. This table demonstrates that fewer behavioral and physiological measurements are employed to evaluate the developed system. We speculate that this is due to the lack of competencies in those fields in the robotics community. More recently, ANA Avatar XPRIZE [10] took the lead in the evaluation of humanoid robot teleoperation systems by assessing both task performance and subjective measures in locomotion, manipulation, and social interaction scenarios. These evaluations can help the developers to choose proper teleoperation devices and adjust the autonomy level, more specifically retargeting and control approaches for humanoid teleoperation according to the architecture presented in Sec. 3.

While developing a humanoid robot teleoperation system, different hardware and software components are interleaved in order to perform a task. In the case of poor performances, it is important to consider the interconnection of different components and try to find the sources of the problems. Another challenge related to teleoperation system development might be latency, which can be related to perception, communication, control, or actuation. An experimental demonstration provided by [146] showed a detrimental impact of the time delay to the human operator's increased workload as well as the performance. Moreover, the discrepancy of the time delay among

TABLE II: Examples of evaluation metrics for the design and evaluation of robot teleoperation setups.

ref	evaluation metrics				autonomy attentive system	measurements			
	usability	situational awareness	workload	engagement, presence		subjective measures	objective & task performance	physiological	behavioural
[145]		✓	✓		✓	✓	✓		
[146]	✓		✓			✓	✓	✓	
[143]		✓		✓		✓	✓		
[141]				✓	✓	✓		✓	
[147]	✓		✓			✓	✓		
[148]	✓			✓		✓	✓		
[60]			✓		✓	✓		✓	✓
[149]		✓				✓	✓		
[150]	✓						✓		

different components, e.g., arms, torso, locomotion, hand, and social cues (such as facial expression), can further degrade the teleoperation performance and the user's experience, highlighting the importance of synchronized and coordinated motion, and social cues for the teleoperated humanoid robot. Unique to humanoid robot teleoperation, having an anthropomorphic humanoid robot body, and a proper selection of teleoperation interfaces, algorithms, and low latency can lead to strong telepresence, or so-called tele-embodiment [151].

A final challenge in humanoid teleoperation is to build a system that can be easily mastered by non-expert users. State of the art does not often consider any usability assessment or carry out any user study, strongly relying on the expertise and training of a single human operator. From this perspective, we hope that this section helps the reader in identifying the key points and right evaluation metrics for deploying a teleoperation system that meets the user's needs.

## VII. APPLICATIONS AND PERSPECTIVES

Future applications of teleoperated humanoid robots are very promising but further progress is required to employ these robots in real world scenarios. For successful teleoperation, the theoretical foundations are provided in the previous sections. Here, we overview major applications of the humanoid robot teleoperation, and discuss their challenges and opportunities.

### A. Telexistence and Telepresence

The COVID-19 pandemic has either canceled many conferences and meetings around the world or led them to be transformed into virtual events. Moreover, many people cannot see their beloved ones frequently. The current substitution for such cases currently is to use communication mediums, allowing the people to see and hear each other. However using these means, many social cues -equally important in social interaction- are not yet conveyed. Instead, humanoid robot teleoperation with anthropomorphic shape and motion, similar to a human, would allow for a better experience. Moreover, it would allow people to interact physically with each other. Currently, the ANA Avatar XPRIZE competition [10] is aiming at this application. The main challenges in this

context are to allow the operator and the humans who interact with the robot to engage in a scenario, feel the presence of the operator, and enable manipulation in the remote environment.

#### B. Teleoperation in Hazardous Environments

One of the main applications urging robot teleoperation is in disaster-response scenarios where the environment is potentially hazardous to humans. This need arose, for example, in the Chernobyl and Fukushima Daiichi Nuclear Power Plants crises [152]. As a response to these disasters, robots were needed to perform surveillance, search and rescue, and manipulation tasks. However, to do that, a robot should reach the point of interest passing possibly through a dynamic and unstructured environment to perform a given task. Bipedal locomotion overcomes other locomotion means in terms of mobility, agility, and a wide range of motion of the robot; therefore, reliable humanoid robot teleoperation can assist the human in such cases. In these scenarios, the mission requirements allow to identify necessary robot features such as hardware reliability, communication protocols, additional hardware and sensors, autonomy level, and so on. Normally, in natural disasters the time is very critical; therefore, another important aspect to consider is the response time and the technology readiness level. For example, in the Fukushima crisis where the human could not enter because of the radiation levels, it took more than three months to deliver a robot that could perform the first mission [152]. However, in cases such as Urban Search and Rescue Tasks, where humans' lives are jeopardized, the delay for intervention is not admissible [153].

#### C. Teleoperation in Manufacturing & Research Environments

Humanoid robots, despite being attractive because of their anthropomorphism and potential to operate in environments designed for humans, are still expensive and not affordable for many industries and research institutes. So far, this inconvenience has strongly limited their study to a restricted robotics community. Moreover, very often either the robot hardware or the underlying control software (or both) are not designed to comply with control failures, loss of balance or wrong interactions with the environment, which can lead to costly breakage. In an example of teleoperation in a manufacturing environment by Kheddar *et al.* [154], the humanoid robots TORO and HRP-4 have been deployed in an aircraft manufacturing environment for assembly operations. Another example in construction sites can be found in [155].

#### D. Telenursing

Front-line healthcare workers are exposed to infectious diseases and are at higher risk of infection compared with the general community. This was evident during disease outbreaks, as experienced during the recent COVID-19 pandemic. Nurses physically interact with their patients and the environment for a wide variety of tasks such as manipulating objects, taking measurements, and interacting with them. In this respect, teleoperated robots can facilitate the situation and improve the nurses' safety by performing some of their tasks, exemplified

in [156]. However, robots should not only perform the nurses' tasks but also avoid patients' discomfort or avoid limiting other health worker activities. In this respect, the teleoperation of humanoid robots can potentially overcome these challenges when being deployed in clinical environments with the supervision of professional operators and other healthcare personnel.

#### E. Space Applications

Space robotics has many applications, including satellite on-orbit servicing, maintenance of the ISS, performing experiments there, and interplanetary exploration and construction [21], [157]. Some of these tasks cannot be performed by humans due to cost, safety, and the increased complexity of the required system. On the other hand, the reliability and robustness of autonomous robots are not yet sufficient to perform such tasks autonomously. Therefore, teleoperation of space robots with different degrees of autonomy is required. Space applications are more challenging due to communication latency and bandwidth, possible unknown kinematics and dynamics properties of the target objects of manipulation, and human factors [157]. In the case of bilateral teleoperation, the round trip communication delay matters; however, time-domain passivity control approaches can compensate to some extent for earth-orbiting robots with the cost of degrading the efficiency [158]. However, for application with higher distances, it is a compromise between autonomy, risk, and efficiency. To overcome that, [159] proposed a supervised autonomy framework with a natural interface to astronauts in the ISS for teleoperating the SUPVIS Justin robot on Earth, simulating an interplanetary solar panel service and maintenance task.

#### F. Service Robotics Application

Another area of use of humanoid robot teleoperation is in domestic environments such as houses, supermarkets, schools, and hotels with a diverse range of goals such as giving care to elderly people or housekeeping [160], restocking the market shelves, teleducation, guiding visitors in hotels, or teletourism. These environments are intrinsically unstructured and built for humans' use; therefore, a humanoid robot is likely to be deployed for such applications. These applications will become more evident when the operator of the humanoid robot cannot be present in the target environment. In other words, workforces who teleoperate the humanoid robot can be at any place. The main requirements for such applications to be acceptable are the safety of humanoid robots with the people whom they are interacting with, and the ability to manipulate and modify remote locations.

#### REFERENCES

- [1] J. Ramos and S. Kim, "Humanoid dynamic synchronization through whole-body bilateral feedback teleoperation," *IEEE Trans. on Robotics*, vol. 34, no. 4, pp. 953–965, 2018.
- [2] K. Darvish, Y. Tirupachuri *et al.*, "Whole-body geometric retargeting for humanoid robots," in *IEEE/RAS Int. Conf. on Humanoid Robots*, 2019.
- [3] L. Penco, N. Scianca *et al.*, "A multimode teleoperation framework for humanoid loco-manipulation: An application for the iCub robot," *IEEE Robotics Automation Mag.*, vol. 26, no. 4, pp. 73–82, 2019.

- [4] Y. Ishiguro, T. Makabe *et al.*, “Bilateral humanoid teleoperation system using whole-body exoskeleton cockpit TABLIS,” *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6419–6426, 2020.
- [5] S. J. Jorgensen, M. W. Lanighan *et al.*, “Deploying the nasa Valkyrie humanoid for IED response: An initial approach and evaluation summary,” in *IEEE/RAS Int. Conf. on Humanoid Robots*, Oct 2019.
- [6] F. Abi-Farraj, B. Henze *et al.*, “Humanoid teleoperation using task-relevant haptic feedback,” in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct 2018.
- [7] D. Kim, B.-J. You, and S.-R. Oh, *Whole Body Motion Control Framework for Arbitrarily and Simultaneously Assigned Upper-Body Tasks and Walking Motion*. Springer, 2013, pp. 87–98.
- [8] R. Cisneros, M. Benallegue *et al.*, “Team JANUS humanoid avatar: A cybernetic avatar to embody human telepresence,” in *Toward Robot Avatars: Perspectives on the ANA Avatar XPRIZE Competition, RSS Workshop*, 2022.
- [9] A. D. Dragan, K. C. T. Lee, and S. S. Srinivasa, “Legibility and predictability of robot motion,” in *ACM/IEEE Int. Conf. on Human-Robot Interaction*, Mar 2013.
- [10] ANA Avatar XPRIZE. [Online]. Available: <https://avatar.xprize.org/prizes/avatar>
- [11] N. Lii, C. Riecke *et al.*, “The robot as an avatar or co-worker? an investigation of the different teleoperation modalities through the KONTUR-2 and METERON SUPVIS Justin space telerobotic missions,” in *Int. Astronautical Cong.*, Oct 2018.
- [12] E. Ackerman. (2019) Russian Humanoid Robot to Pilot Soyuz Capsule to ISS This Week. [Online]. Available: <https://spectrum.ieee.org/automaton/robotics/space-robots/russian-humanoid-robot-to-pilot-soyuz-capsule-to-iss-this-week>
- [13] A. Goswami and P. Vadakkepat, *Humanoid robotics: a reference*. Springer, 2019.
- [14] T. Sugihara and M. Morisawa, “A survey: dynamics of humanoid robots,” *Advanced Robotics*, vol. 34, no. 21-22, pp. 1338–1352, 2020.
- [15] K. Yamamoto, T. Kamioka, and T. Sugihara, “Survey on model-based biped motion control for humanoid robots,” *Advanced Robotics*, vol. 34, no. 21-22, pp. 1353–1369, 2020.
- [16] Y. Tazaki and M. Murooka, “A survey of motion planning techniques for humanoid robots,” *Advanced Robotics*, vol. 34, no. 21-22, pp. 1370–1379, 2020.
- [17] P. G. De Barros and R. W. Lindeman, “A survey of user interfaces for robot teleoperation,” *Worcester Polytechnic Institute*, 2009.
- [18] P. F. Hokayem and M. W. Spong, “Bilateral teleoperation: An historical survey,” *Automatica*, vol. 42, no. 12, pp. 2035–2057, 2006.
- [19] M. A. Goodrich and A. C. Schultz, “Human-robot interaction: a survey,” *Foundations and Trends® in Human-Computer Interaction*, vol. 1, no. 3, pp. 203–275, 2007.
- [20] M. A. Goodrich, J. W. Crandall, and E. Barakova, “Teleoperation and beyond for assistive humanoid robots,” *Reviews of Human factors and ergonomics*, vol. 9, no. 1, pp. 175–226, 2013.
- [21] K. Darvish, D. Pucci *et al.* Workshop on teleoperation of humanoid robots. [Online]. Available: [https://dic-iit.github.io/WS\\_teleopreration\\_humanoids/](https://dic-iit.github.io/WS_teleopreration_humanoids/)
- [22] S. Tachi, *Telexistence*. World Scientific 2nd edition, 2015.
- [23] T. Ando, T. Watari, and R. Kikuuwe, “Master-slave bipedal walking and semi-automatic standing up of humanoid robots,” in *IEEE/SICE Int. Symp. on System Integration*, Jan 2020.
- [24] D. Roetenberg, H. Luinge, and P. Slycke, “Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors,” *Xsens Motion Technologies BV, Tech. Rep.*, vol. 1, 2009.
- [25] O. Glauser, S. Wu *et al.*, “Interactive hand pose estimation using a stretch-sensing soft glove,” *ACM Trans. on Graphics*, vol. 38, no. 4, pp. 1–15, 2019.
- [26] J. Wang, C. Lu, and K. Zhang, “Textile-based strain sensor for human motion detection,” *Energy & Environmental Materials*, vol. 3, no. 1, pp. 80–100, 2020.
- [27] J. Ramos and S. Kim, “Dynamic locomotion synchronization of bipedal robot and human operator via bilateral feedback teleoperation,” *Science Robotics*, vol. 4, no. 35, 2019.
- [28] M. Elobaid, Y. Hu *et al.*, “Telexistence and teleoperation for walking humanoid robots,” *IntelliSys*, vol. 1038, pp. 1106–1121, 2019.
- [29] S. Daffarra, K. Darvish *et al.*, “iCub3 avatar system,” *arXiv preprint arXiv:2203.06972*, 2022.
- [30] M. Johnson, B. Shrewsbury *et al.*, “Team IHMC’s lessons learned from the DARPA Robotics Challenge: Finding data in the rubble,” *J. of Field Robotics*, vol. 34, no. 2, pp. 241–261, 2017.
- [31] J. Chagas Vaz, D. Wallace, and P. Y. Oh, “Humanoid loco-manipulation of pushed carts utilizing virtual reality teleoperation,” in *ASME Int. Mechanical Engineering Congress and Exposition*, Nov 2021.
- [32] M. Zucker, S. Joo *et al.*, “A general-purpose system for teleoperation of the DRC-HUBO humanoid robot,” *J. of Field Robotics*, vol. 32, 2015.
- [33] R. Cisneros, M. Morisawa *et al.*, “Enabling a teleoperated humanoid robot to pass through debris-filled terrain using manipulation,” in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2016.
- [34] M. Schwarz, C. Lenz *et al.*, “Nimbro avatar: Interactive immersive telepresence with force-feedback telemanipulation,” in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Sept-Oct 2021.
- [35] O. E. Ramos, N. Mansard *et al.*, “Dancing humanoid robots: Systematic use of OSID to compute dynamically consistent movements following a motion capture pattern,” *IEEE Robotics Automation Mag.*, vol. 22, no. 4, 2015.
- [36] Y. Ishiguro, K. Kojima *et al.*, “High speed whole body dynamic motion experiment with real time master-slave humanoid robot system,” in *IEEE Int. Conf. on Robotics and Automation*, May 2018.
- [37] S. Tachi, Y. Inoue, and F. Kato, “Telesar vi: Telexistence surrogate anthropomorphic robot vi,” *Int. J. of Humanoid Robotics*, vol. 17, no. 05, 2020.
- [38] K. Hu, C. Ott, and D. Lee, “Online human walking imitation in task and joint space based on quadratic programming,” in *IEEE Int. Conf. on Robotics and Automation*, June 2014.
- [39] A. Brygo, I. Sarakoglou *et al.*, “Humanoid robot teleoperation with vibrotactile based balancing feedback,” in *Int. Conf. on Human Haptic Sensing and Touch Enabled Computer Applications*, Jun 2014.
- [40] A. D. Fava, K. Bouyarmane *et al.*, “Multi-contact motion retargeting from human to humanoid robot,” in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2016.
- [41] L. Pernel and J. Babič, “Learning of compliant human–robot interaction using full-body haptic interface,” *Advanced Robotics*, vol. 27, 09 2013.
- [42] D. Staudenmann, K. Roeleveld *et al.*, “Methodological aspects of SEMG recordings for force estimation—a tutorial and review,” *J. of electromyography and kinesiology*, vol. 20, no. 3, pp. 375–387, 2010.
- [43] R. H. Chowdhury, M. B. Reaz *et al.*, “Surface electromyography signal processing and classification techniques,” *Sensors*, vol. 13, no. 9, 2013.
- [44] E. Niedermeyer and F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.
- [45] J. Achten and A. E. Jeukendrup, “Heart rate monitoring,” *Sports medicine*, vol. 33, no. 7, pp. 517–538, 2003.
- [46] Y. Sugano and A. Bulling, “Self-calibrating head-mounted eye trackers using egocentric visual saliency,” in *ACM Symp. on User Interface Software & Technology*, Nov 2015.
- [47] T. Ohhashi, M. Sakaguchi, and T. Tsuda, “Human perspiration measurement,” *Physiological measurement*, vol. 19, no. 4, p. 449, 1998.
- [48] Y. G. Ryu, H. C. Roh *et al.*, “Digital image stabilization for humanoid eyes inspired by human VOR system,” in *IEEE Int. Conf. on Robotics and Biomimetics*, Dec 2009.
- [49] A. Wang, J. Ramos *et al.*, “The HERMES humanoid system: A platform for full-body teleoperation with balance feedback,” in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2015.
- [50] C. R. Wagner, S. J. Lederman, and R. D. Howe, “Design and performance of a tactile shape display using RC servomotors,” *Haptics-e*, vol. 3, no. 4, pp. 1–6, 2004.
- [51] S. Tachi, K. Mlnamlzawa *et al.*, “Haptic media construction and utilization of human-harmonized ‘tangible’ information environment,” in *Int. Conf. on Artificial Reality and Telexistence*, Dec 2013.
- [52] S. Nakaoka, M. Morisawa *et al.*, “Task sequencer integrated into a teleoperation interface for biped humanoid robots,” in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2015.
- [53] S. Feng, E. Whitman *et al.*, “Optimization-based full body control for the DARPA robotics challenge,” *J. of Field Robotics*, vol. 32, no. 2, pp. 293–312, 2015.
- [54] R. Cisneros-Limon, M. Morisawa *et al.*, “An inverse dynamics-based multi-contact locomotion control framework without joint torque feedback,” *Advanced Robotics*, vol. 34, no. 21-22, pp. 1398–1419, 2020.
- [55] S. Kajita, F. Kanehiro *et al.*, “The 3D linear inverted pendulum mode: A simple modeling for a biped walking pattern generation,” in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct-Nov 2001.
- [56] M. Vukobratović and J. Stepanenko, “On the stability of anthropomorphic systems,” *Mathematical biosciences*, vol. 15, no. 1-2, 1972.
- [57] J. Pratt, J. Carff *et al.*, “Capture point: A step toward humanoid push recovery,” in *IEEE/RAS Int. Conf. on Humanoid Robots*, Dec 2006.



- [58] T. Takenaka, T. Matsumoto, and T. Yoshiike, "Real time motion generation and control for biped robot-1st report: Walking gait pattern generation," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct 2009.
- [59] J. Englsberger, C. Ott, and A. Albu-Schäffer, "Three-dimensional bipedal walking control based on divergent component of motion," *IEEE Trans. on Robotics*, vol. 31, no. 2, pp. 355–368, 2015.
- [60] G. Singh, S. Bermúdez i Badia *et al.*, "Physiologically attentive user interface for robot teleoperation: real time emotional state estimation and interface modification using physiology, facial expressions and eye movements," in *Int. Joint Conf. on Biomedical Engineering Systems and Technologies*, Jan 2018.
- [61] A. Dragan and S. Srinivasa, "Formalizing assistive teleoperation," in *Robotics Science and Systems*, Jul 2012.
- [62] R. J. Griffin, G. Wiedebach *et al.*, "Footstep planning for autonomous walking over rough terrain," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Oct 2019.
- [63] J. Garimort, A. Hornung, and M. Bennewitz, "Humanoid navigation with dynamic footstep plans," in *IEEE Int. Conf. on Robotics and Automation*, May 2011.
- [64] N. Perrin, O. Stasse *et al.*, "Fast humanoid robot collision-free footstep planning using swept volume approximations," *IEEE Trans. on Robotics*, vol. 28, no. 2, 2011.
- [65] J. J. Kuffner, K. Nishiwaki *et al.*, "Footstep planning among obstacles for biped robots," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct–Nov 2001.
- [66] A. Herdt, H. Diedam *et al.*, "Online walking motion generation with automatic footstep placement," *Advanced Robotics*, vol. 24, no. 5–6, pp. 719–737, 2010.
- [67] R. Deits and R. Tedrake, "Footstep planning on uneven terrain with mixed-integer convex optimization," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2014.
- [68] M. Fakoor, A. Kosari, and M. Jafarzadeh, "Revision on fuzzy artificial potential field for humanoid robot path planning in unknown environment," *Int. J. of Advanced Mechatronic Systems*, vol. 6, no. 4, 2015.
- [69] M. V. Liarokapis, P. Artemiadis *et al.*, "Directions, methods and metrics for mapping human to robot motion with functional anthropomorphism: A review," *School of Mechanical Engineering, National Technical University of Athens, Tech. Rep.*, 2013.
- [70] K. Ayusawa and E. Yoshida, "Motion retargeting for humanoid robots based on simultaneous morphing parameter identification and motion optimization," *IEEE Trans. on Robotics*, vol. 33, no. 6, pp. 1343–1357, 2017.
- [71] K. Otani and K. Bouyarmane, "Adaptive whole-body manipulation in human-to-humanoid multi-contact motion retargeting," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2017.
- [72] T. Koolen, T. De Boer *et al.*, "Capturability-based analysis and control of legged locomotion, part 1: Theory and application to three simple gait models," *The Int. J. of Robotics Research*, vol. 31, no. 9, pp. 1094–1113, 2012.
- [73] H. Hirukawa, S. Hattori *et al.*, "A universal stability criterion of the foot contact of legged robots-adios ZMP," in *IEEE Int. Conf. on Robotics and Automation*, May 2006.
- [74] C. Ott, M. A. Roa, and G. Hirzinger, "Posture and balance control for biped robots based on contact force optimization," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Oct 2011.
- [75] S. Mason, L. Righetti, and S. Schaal, "Full dynamics LQR control of a humanoid robot: An experimental study on balancing and squatting," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2014.
- [76] S. Kuindersma, R. Deits *et al.*, "Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot," *Autonomous robots*, vol. 40, no. 3, 2016.
- [77] M. Kelly, "An introduction to trajectory optimization: How to do your own direct collocation," *SIAM Review*, vol. 59, no. 4, 2017.
- [78] M. Posa, C. Cantu, and R. Tedrake, "A direct method for trajectory optimization of rigid bodies through contact," *The Int. J. of Robotics Research*, vol. 33, no. 1, pp. 69–81, 2014.
- [79] O. Kanoun, F. Lamiriaux, and P. Wieber, "Kinematic control of redundant manipulators: Generalizing the task-priority framework to inequality task," *IEEE Trans. on Robotics*, vol. 27, no. 4, 2011.
- [80] L. Righetti and S. Schaal, "Quadratic programming for inverse dynamics with optimal distribution of contact forces," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2012.
- [81] D. E. Orin, A. Goswami, and S.-H. Lee, "Centroidal dynamics of a humanoid robot," *Autonomous Robots*, vol. 35, no. 2–3, 2013.
- [82] T. Koolen, S. Bertrand *et al.*, "Design of a momentum-based control framework and application to the humanoid robot atlas," *Int. J. of Humanoid Robotics*, vol. 13, pp. 1650007–1, 03 2016.
- [83] F. Flacco, A. Paolillo, and A. Kheddar, "Residual-based contacts estimation for humanoid robots," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2016.
- [84] J. Hwangbo, J. Lee *et al.*, "Learning agile and dynamic motor skills for legged robots," *Science Robotics*, vol. 4, no. 26, p. eaau5872, 2019.
- [85] F. J. A. Chavez, J. Kangro *et al.*, "Contact force and joint torque estimation using skin," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3900–3907, 2018.
- [86] S. Kajita, M. Morisawa *et al.*, "Biped walking stabilization based on linear inverted pendulum tracking," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct 2010.
- [87] T. Miki, J. Lee *et al.*, "Learning robust perceptive locomotion for quadrupedal robots in the wild," *Science Robotics*, vol. 7, no. 62, p. eabk2822, 2022.
- [88] C. G. Atkeson *et al.*, "State estimation of a walking humanoid robot," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct 2012.
- [89] N. Rotella, M. Bloesch *et al.*, "State estimation for a humanoid robot," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Sept 2014.
- [90] M. Bloesch, M. Hutter *et al.*, "State estimation for legged robots-consistent fusion of leg kinematics and IMU," *Robotics*, vol. 17, 2013.
- [91] H. Herr and M. Popovic, "Angular momentum in human walking," *J. of experimental biology*, vol. 211, no. 4, pp. 467–481, 2008.
- [92] V. Zordan, D. Brown *et al.*, "Control of rotational dynamics for ground and aerial behavior," *IEEE Trans. on visualization and computer graphics*, vol. 20, no. 10, pp. 1356–1366, 2014.
- [93] S. Kuindersma. (2020, June) Recent progress on atlas, the world's most dynamic humanoid robot. *Robotics Today - A Series of Technical Talks*. [Online]. Available: <https://youtu.be/EGABAx52GKI>
- [94] H. Dai, A. Valenzuela, and R. Tedrake, "Whole-body motion planning with centroidal dynamics and full kinematics," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2014.
- [95] M. Hutter. (2022) Mit robotics seminar: Robots in the wild. [Online]. Available: <https://youtu.be/24uTRT32Cyw>
- [96] Y.-C. Lin, B. Ponton *et al.*, "Efficient humanoid contact planning using learned centroidal dynamics prediction," in *Int. Conf. on Robotics and Automation*, May 2019.
- [97] X. B. Peng, P. Abbeel *et al.*, "Deepmimic: Example-guided deep reinforcement learning of physics-based character skills," *ACM Trans. on Graphics*, vol. 37, no. 4, pp. 1–14, 2018.
- [98] Z. Xie, P. Clary *et al.*, "Iterative reinforcement learning based design of dynamic locomotion skills for Cassie," *arXiv preprint arXiv:1903.09537*, 2019.
- [99] S. Choi and J. Kim, "Cross-domain motion transfer via safety-aware shared latent space modeling," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2634–2641, 2020.
- [100] R. Villegas, J. Yang *et al.*, "Neural kinematic networks for unsupervised motion retargeting," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, Jun 2018.
- [101] A. Dragan and S. Srinivasa, "A policy blending formalism for shared control," *Int. J. of Robotics Research*, vol. 32, no. 7, pp. 790–805, 2013.
- [102] D. Rakita, B. Mutlu *et al.*, "Shared control-based bimanual robot manipulation," *Science Robotics*, vol. 4, no. 30, 2019.
- [103] R. Rahal, F. Abi-Farraj *et al.*, "Haptic shared-control methods for robotic cutting under nonholonomic constraints," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 8151–8157.
- [104] K. T. Ly, M. Poozhiyil *et al.*, "Intent-aware predictive haptic guidance and its application to shared control teleoperation," in *IEEE Int. Conf. on Robot & Human Interactive Communication (RO-MAN)*, Aug 2021.
- [105] R. Rahal, G. Matarese *et al.*, "Caring about the human operator: Haptic shared control for enhanced user comfort in robotic telemanipulation," *IEEE Trans. on Haptics*, vol. 13, no. 1, pp. 197–203, 2020.
- [106] D. Glas, S. Satake *et al.*, "The network robot system: Enabling social human-robot interaction in public spaces," *J. of Human-Robot Interaction*, vol. 1, 01 2013.
- [107] T. W. Fong, C. Thorpe, and C. Baur, "A safeguarded teleoperation controller," in *IEEE Int. Conf. on Advanced Robotics (ICAR)*, Aug 2001.
- [108] A. Brygo, I. Sarakoglou *et al.*, "Tele-manipulation with a humanoid robot under autonomous joint impedance regulation and vibrotactile balancing feedback," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2014.

- [109] A. Peer, S. Hirche *et al.*, "Intercontinental multimodal tele-cooperation using a humanoid robot," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Sept 2008.
- [110] P. Evrard, N. Mansard *et al.*, "Intercontinental, multimodal, wide-range tele-cooperation using a humanoid robot," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct 2009.
- [111] J. Ramos, A. Wang *et al.*, "A balance feedback interface for whole-body teleoperation of a humanoid robot and implementation in the HERMES system," in *IEEE/RAS Int. Conf. on Humanoid Robots*, Nov 2015.
- [112] A. Q. Keemink, H. van der Kooij, and A. H. Stienen, "Admittance control for physical human-robot interaction," *The Int. J. of Robotics Research*, vol. 37, no. 11, pp. 1421–1444, 2018.
- [113] A. Ajoudani, N. Tsagarakis, and A. Bicchi, "Tele-impedance: Teleoperation with impedance regulation using a body-machine interface," *The Int. J. of Robotics Research*, vol. 31, no. 13, pp. 1642–1656, 2012.
- [114] A. Bemporad, "Predictive control of teleoperated constrained systems with unbounded communication delays," in *IEEE Conf. on Decision and Control*, Dec 1998.
- [115] W. R. Ferrell, "Remote manipulation with transmission delay," *IEEE Trans. on Human Factors in Electronics*, vol. 6, no. 1, pp. 24–32, 1965.
- [116] Y.-C. Liu and N. Chopra, "Control of semi-autonomous teleoperation system with time delays," *Automatica*, vol. 49, no. 6, 2013.
- [117] W. R. Ferrell, "Remote manipulation with transmission delay," *IEEE Trans. on Human Factors in Electronics*, vol. HFE-6, no. 1, 1965.
- [118] R. J. Anderson and M. W. Spong, "Asymptotic Stability for Force Reflecting Teleoperators with Time Delay," *The Int. J. of Robotics Research*, vol. 11, no. 2, 1992.
- [119] G. Niemeyer and J.-J. E. Slotine, "Stable adaptive teleoperation," *IEEE J. of Oceanic Engineering*, vol. 16, no. 1, pp. 152–162, 1991.
- [120] R. Lozano, N. Chopra, and M. W. Spong, "Passivation of force reflecting bilateral teleoperators with time varying delay," in *Mechatronics Forum Int. Conf.*, Jun 2002.
- [121] K. Kosuge, T. Itoh, and T. Fukuda, "Scaled telemanipulation with communication time delay," in *IEEE Int. Conf. on Robotics and Automation*, Apr 1996.
- [122] N. Chopra and M. W. Spong, "Bilateral teleoperation over the internet: the time varying delay problem," in *American Control Conf.*, Jun 2003.
- [123] R. J. Anderson and M. W. Spong, "Bilateral control of teleoperators with time delay," *IEEE Trans. on Automatic Control*, vol. 34, no. 5, pp. 131–138, 1989.
- [124] M. Risiglione, J.-P. Sleiman *et al.*, "Passivity-based control for haptic teleoperation of a legged manipulator in presence of time-delays," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Sept 2021.
- [125] J. Milton, J. L. Cabrera *et al.*, "The time-delayed inverted pendulum: implications for human balance control," *Chaos: An Interdisciplinary J. of Nonlinear Science*, vol. 19, no. 2, p. 026110, 2009.
- [126] A. Steinfeld, T. Fong *et al.*, "Common metrics for human-robot interaction," in *ACM SIGCHI/SIGART Conf. on Human-Robot Interaction*, Mar 2006.
- [127] ISO 9241-11:2018(en) Ergonomics of human-system interaction — part 11: Usability: Definitions and concepts. [Online]. Available: <https://www.iso.org/>
- [128] C. Flavián, M. Guinalfú, and R. Gurrea, "The role played by perceived usability, satisfaction and consumer trust on website loyalty," *Information & management*, vol. 43, no. 1, pp. 1–14, 2006.
- [129] N. Bevan, J. Carter *et al.*, "New ISO standards for usability, usability reports and usability measures," in *Int. Conf. on Human-Computer Interaction*, Jul 2016.
- [130] J. Brooke *et al.*, "SUS: A quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.
- [131] G. Adamides, G. Christou *et al.*, "Usability guidelines for the design of robot teleoperation: A taxonomy," *IEEE Trans. on Human-Machine Systems*, vol. 45, no. 2, 2014.
- [132] M. R. Endsley, "Situation awareness global assessment technique (SAGAT)," in *IEEE Nat. Aerospace and Electronics Conf.*, May 1988.
- [133] T. Nguyen, C. P. Lim *et al.*, "A review of situation awareness assessment approaches in aviation environments," *IEEE Systems J.*, vol. 13, no. 3, 2019.
- [134] S. Miller, "Workload measures," *National Advanced Driving Simulator. Iowa City, United States*, 2001.
- [135] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Advances in psychology*, 1988, vol. 52, pp. 139–183.
- [136] G. B. Reid and T. E. Nygren, "The subjective workload assessment technique: A scaling procedure for measuring mental workload," in *Advances in psychology*, 1988, vol. 52, pp. 185–218.
- [137] P. Bouvier, E. Lavoué, and K. Sehaba, "Defining engagement and characterizing engaged-behaviors in digital gaming," *Simulation & Gaming*, vol. 45, no. 4-5, pp. 491–507, 2014.
- [138] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, 1998.
- [139] J. Lessiter, J. Freeman *et al.*, "A cross-media presence questionnaire: The ITC-Sense of presence inventory," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, 2001.
- [140] G. Riva, F. Davide, and W. IJsselstein, "7 Measuring Presence: Subjective, behavioral and physiological methods," *Being there: Concepts, effects and measurement of user presence in synthetic environments*, pp. 110–118, 2003.
- [141] P. Rani and N. Sarkar, "Operator engagement detection and robot behavior adaptation in human-robot interaction," in *IEEE Int. Conf. on Robotics and Automation*, Apr 2005.
- [142] A. H. Roscoe, "Assessing pilot workload. why measure heart rate, HRV and respiration?" *Biological psychology*, vol. 34, no. 2-3, 1992.
- [143] J. M. Riley, D. B. Kaber, and J. V. Draper, "Situation awareness and attention allocation measures for quantifying telepresence experiences in teleoperation," *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 14, no. 1, pp. 51–67, 2004.
- [144] Y. Deng, J. Shirley *et al.*, "Influence of dynamic automation function allocations on operator situation awareness and workload in unmanned aerial vehicle control," in *Int. Conf. on Human Factors and Systems Interaction*, Jul 2019.
- [145] D. B. Kaber, E. Onal, and M. R. Endsley, "Design of automation for telerobots and the effect on performance, operator situation awareness, and subjective workload," *Human factors and ergonomics in manufacturing & service industries*, vol. 10, no. 4, pp. 409–430, 2000.
- [146] S. Lu, M. Y. Zhang *et al.*, "Workload management in teleoperation of unmanned ground vehicles: Effects of a delay compensation aid on human operators' workload and teleoperation performance," *Int. J. of Human-Computer Interaction*, vol. 35, no. 19, pp. 1820–1830, 2019.
- [147] S. Radmard, A. J. Moon, and E. A. Croft, "Interface design and usability analysis for a robotic telepresence platform," in *IEEE Int. Symp. on Robot and Human Interactive Communication*, Sept 2015.
- [148] J. Jankowski and A. Grabowski, "Usability evaluation of vr interface for mobile robot teleoperation," *Int. J. of Human-Computer Interaction*, vol. 31, no. 12, pp. 882–889, 2015.
- [149] P. G. De Barros, R. W. Lindeman, and M. O. Ward, "Enhancing robot teleoperator situation awareness and performance using vibro-tactile and graphical feedback," in *IEEE Symp. on 3D User Interfaces*, Mar 2011.
- [150] V. Villani, L. Sabatini *et al.*, "A natural infrastructure-less human-robot interaction system," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1640–1647, 2017.
- [151] E. Paulos and J. Canny, "Social tele-embodiment: Understanding presence," *Autonomous Robots*, vol. 11, no. 1, pp. 87–95, 2001.
- [152] K. Nagatani, S. Kiribayashi *et al.*, "Emergency response to the nuclear accident at the Fukushima Daiichi Nuclear Power Plants using mobile rescue robots," *J. of Field Robotics*, vol. 30, no. 1, pp. 44–63, 2013.
- [153] J. Casper and R. R. Murphy, "Human-robot interactions during the robot-assisted urban search and rescue response at the world trade center," *IEEE Trans. on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 33, no. 3, pp. 367–385, 2003.
- [154] A. Kheddar, S. Caron *et al.*, "Humanoid robots in aircraft manufacturing: The Airbus use cases," *IEEE Robotics & Automation Mag.*, vol. 26, no. 4, pp. 30–45, 2019.
- [155] K. Yokoi, K. Nakashima *et al.*, "A tele-operated humanoid robot drives a backhoe in the open air," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Oct 2003.
- [156] Z. Li, P. Moran *et al.*, "Development of a tele-nursing mobile manipulator for remote care-giving in quarantine areas," in *IEEE Int. Conf. on Robotics and Automation*, May-Jun 2017.
- [157] A. Flores-Abad, O. Ma *et al.*, "A review of space robotics technologies for on-orbit servicing," *Progress in Aerospace Sciences*, vol. 68, pp. 1–26, 2014.
- [158] J.-H. Ryu, J. Artigas, and C. Preusche, "A passive bilateral control scheme for a teleoperator with time-varying communication delay," *Mechatronics*, vol. 20, no. 7, pp. 812–823, 2010.
- [159] N. Y. Lii, D. Leidner *et al.*, "Toward scalable intuitive telecommand of robots for space deployment with METERON SUPVIS Justin," 2017.
- [160] J. Broekens, M. Heerink *et al.*, "Assistive social robots in elderly care: a review," *Gerontechnology*, vol. 8, no. 2, 2009.



**Kourosh Darvish** Kourosh Darvish is a postdoctoral researcher at the Computer Science and Robotics Institute of the University of Toronto (UofT) and a member of the Vector Institute. Before joining UofT in 2022, he was a postdoctoral researcher at the Italian Institute of Technology (IIT). In 2019, he completed his PhD in Bioengineering and Robotics from the University of Genoa, Italy. His research focuses on shared autonomy, teleoperation, human-robot collaboration, and humanoid robotics.



**Jerry Pratt** Jerry Pratt is the CTO of Figure, a startup creating general purpose humanoid robots for commercial applications. Before joining Figure, Jerry was the Co-founder of Boardwalk Robotics, as well as a Principal Investigator at IHMC, where he co-lead a research group focused on walking and running robots and exoskeletons. Jerry was the team lead for Team IHMC in the DARPA Robotics Challenge (DRC). In 2015 IHMC won second place in the DRC finals.



**Luigi Penco** Luigi Penco is a Senior Research Associate at the Florida Institute for Human and Machine Cognition (IHMC). He earned his bachelor's degree in Electronics Engineering from Roma Tre University in 2015 and a master's in Artificial Intelligence and Robotics from La Sapienza University of Rome in 2018. In 2022 he received a PhD in robotics from Université de Lorraine, while conducting his doctoral studies at Inria Nancy Grand-Est. His research focuses on humanoid robotics, with a particular interest in teleoperation and machine learning techniques used to improve the control and skills of robots.

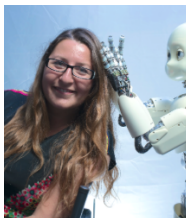


**Eiichi Yoshida** Eiichi Yoshida received M.E. and Ph.D. degrees on Precision Machinery Engineering from Graduate School of Engineering, the University of Tokyo in 1996. He then joined former Mechanical Engineering Laboratory, later in 2001 reorganized as National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan. He served as Co-Director of AIST-CNRS JRL (Joint Robotics Laboratory) at LAAS- CNRS, Toulouse, France, from 2004 to 2008, and at AIST, Tsukuba, Japan from 2009 to 2021. Since 2022, he is Professor of

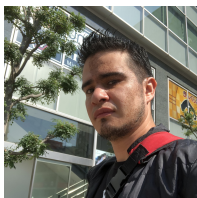
Tokyo University of Science, at Department of Applied Electronics, Faculty of Advanced Engineering. He is IEEE Fellow, and member of RSJ, SICE and JSME and is currently serving as Senior Editor of IEEE Transactions on Robotics. His research interests include robot task and motion planning, human modeling, humanoid robotics and advanced logistics technology.



**Joao Ramos** Joao Ramos is an Assistant Professor at the University of Illinois at Urbana-Champaign (UIUC) and the director of the RoboDesign Lab. He previously worked as a Postdoctoral Associate at the Biomimetic Robotics Laboratory at MIT before joining UIUC in 2019. He received a PhD from the Department of Mechanical Engineering at MIT in 2018. He is the recipient of the 2021 NSF CARRER Award. His research focuses on the design and control of dynamic humanoid robots, human-machine interfaces for whole-body teleoperation, legged locomotion dynamics, and actuation systems.



**Serena Ivaldi** Serena Ivaldi is a research scientist at Inria Nancy Grand-Est, France. She obtained her Ph.D. in Humanoid Technologies in 2011 at the Italian Institute of Technology and University of Genoa, Italy, and the French Habilitation to Direct Research (HDR) at the University of Lorraine, France, in 2022. Her research is focused on humanoid robotics and human-robot interaction.



**Rafael Cisneros** Rafael Cisneros (Member, IEEE) received the B.Eng. degree in Electronics and Computers from the University of the Americas—Puebla (UDLA-P), Puebla, Mexico, in 2006, the M.Sc. degree in Automatic Control from the Center of Research and Advanced Studies, National Polytechnic Institute (CINVESTAV-IPN), Mexico City, Mexico, in 2009, and the Ph.D. degree in Intelligent Interaction Technologies from the University of Tsukuba, Tsukuba, Japan, in 2015. Since then, he has been with the National Institute of Advanced Industrial

Science and Technology (AIST), Tsukuba, Japan, from 2015 to 2018 as a Postdoc, and since 2018, as a Researcher. He is currently a member of CNRS-AIST JRL (Joint Robotics Laboratory), IRL, AIST. His research interests include torque control, whole-body multi-contact motion control of humanoid robots, multibody collision dynamics, teleoperation, and tactile feedback.



**Daniele Pucci** Daniele received the bachelor and master degrees in Control Engineering with highest honors from "Sapienza", University of Rome, in 2007 and 2009, respectively. In 2013, he earned the PhD title with a thesis prepared at INRIA Sophia Antipolis, France, with the supervision of Tarek Hamel, Salvatore Monaco, and Claude Samson. From 2013 to 2017, he has been a postdoc at the Istituto Italiano di Tecnologia (IIT). From August 2017 to August 2021, he has been the head of the Dynamic Interaction Control lab. Since September

2021, Daniele is the PI leading the Artificial and Mechanical Intelligence research line at IIT, a team composed of about forty members that combines AI and Mechanics to devise the next generation of the iCub humanoid robot.