



**HAL**  
open science

# A Multi-Mode Teleoperation Framework for Humanoid Loco-Manipulation: An Application for the iCub Robot

Luigi Penco, Nicola Scianca, Valerio Modugno, Leonardo Lanari, Giuseppe Oriolo, Serena Ivaldi

## ► To cite this version:

Luigi Penco, Nicola Scianca, Valerio Modugno, Leonardo Lanari, Giuseppe Oriolo, et al.. A Multi-Mode Teleoperation Framework for Humanoid Loco-Manipulation: An Application for the iCub Robot. IEEE Robotics and Automation Magazine, 2019, 26 (4), pp.73-82. 10.1109/MRA.2019.2941245 . hal-02291907

**HAL Id: hal-02291907**

**<https://inria.hal.science/hal-02291907>**

Submitted on 19 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Multi-Mode Teleoperation Framework for Humanoid Loco-Manipulation

Luigi Penco, Nicola Scianca, Valerio Modugno, Leonardo Lanari, Giuseppe Oriolo, Serena Ivaldi

## I. INTRODUCTION

Over the years, there have been many improvements on safety standards and conditions at work, but there are still many situations and environments where human lives are at risk, such as in search and rescue, construction sites and chemical plants. We envision a world where robots can act as physical avatars and effectively replace humans in those hazardous scenarios by means of teleoperation.

Despite the many successful cases of teleoperating mobile robots and manipulators, even in space, teleoperating humanoid robots is still very challenging. While humanoid robots are designed with the ambition to mimic the human body capabilities, differences in kinematics (e.g., structure and joint limits) and dynamics (e.g., mass distribution, inertia) are still significant. Another crucial issue is the necessity of ensuring the dynamic balance of the robot while trying to imitate the human motion. This is not straightforward during locomotion tasks, in which the dynamics is highly involved.

A possible solution is therefore to use two levels of teleoperation: a low-level for manipulation, realized via whole-body teleoperation (Fig. 1), and a high-level for locomotion, based on the generation of reference velocities that are then tracked by the robot. We believe that this combination of different modes of teleoperation will considerably ease the burden of controlling humanoid robots, ultimately increasing their adaptability to complex situations which cannot be handled satisfactorily by fully autonomous systems.

### A. Workplace risks

Studies [1], [2], [3] indicate that hundreds of thousands of workers die each year worldwide, with a staggering cost (around 4% of global GDP) due to the time loss, worker compensation, interruption of production, and medical expenses. Not surprisingly, data show that some work activities remain inherently dangerous even with strict work regulations in place. A study [3] from 2012 identifies cancer, respiratory disease, and accidents as the major causes of work-related deaths that could be prevented through workplace automation. In particular, a survey of the Bureau of Labor Statistics [4] shows that the most common accidents in the US are fatal falls, collisions with objects and equipment and injuries in confined spaces. The number of these casualties

The authors are with the Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Sapienza Università di Roma, Rome, Italy. E-mail: {scianca,modugno,lanari,oriolo}@diag.uniroma1.it; and INRIA Grand-est, Nancy, France. E-mail: {luigi.penco,serena.ivaldi}@inria.fr. This work is supported by the H2020 project AnDy.

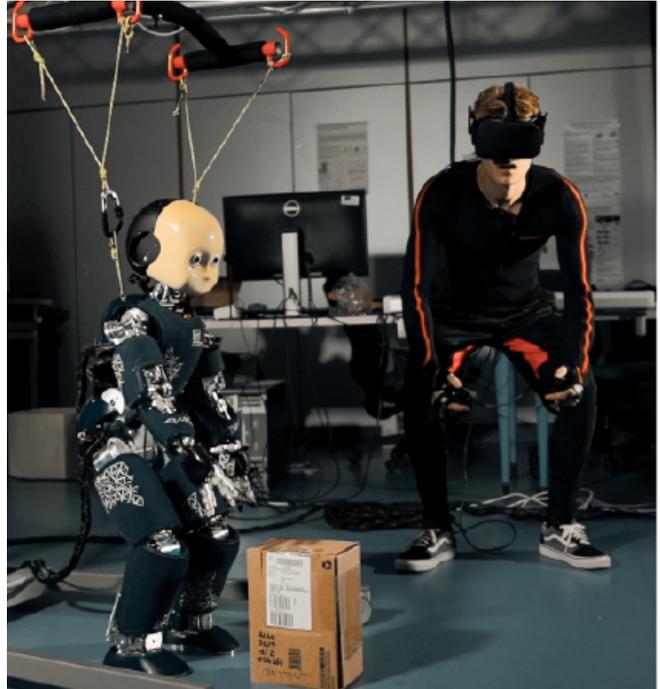


Fig. 1: A human operator controlling a humanoid during a pick-up task.

could be reduced by avoiding the physical presence of the operator.

For example, removing asbestos roof tiles is an operation that could be performed by teleoperated robots. Currently this task is carried out by humans in a context that is extremely dangerous for their health, not only because they have to move on roofs but also because they are exposed to asbestos particles (Fig. 2, top). In the oil and gas industry, workers are often required to enter confined spaces for inspection and maintenance, exposing themselves to hazards such as toxic vapors, not to mention the difficulties of evacuation in case of accidents (Fig. 2, bottom).

### B. Humanoids at work

In the last decade, many research projects have investigated the use of humanoids for reducing risks and fatigue of human workers. The Department of Defense of the USA launched in 2012 the DARPA Robotics Challenge (DRC), a prize competition for promoting and testing humanoid application in the context of search and rescue. Its aim was to develop semi-autonomous ground robots that could perform “complex tasks in dangerous, degraded, human-engineered

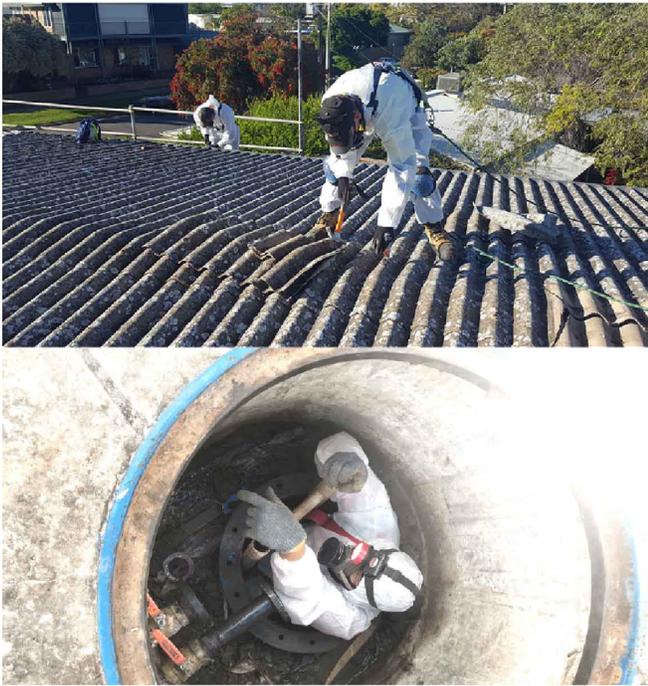


Fig. 2: Examples of hazardous working environments where robotic avatars could replace humans. Top: Asbestos tiles removal operation. Bottom: Maintenance operation in a confined space for oil and gas.

environments”. The DRC finals also represented a midterm evaluation for WALK-MAN [5], an EU Horizon 2020 project aiming to develop a humanoid platform for autonomous or teleoperated intervention in buildings.

COMANOID [6] has been another EU Horizon 2020 project focused on using humanoids to relieve human workers from low added-value tasks that are tiring or dangerous. The project revolved around the idea that humanoid robots are more suited to operate in narrow and cluttered environments typical of maintenance and manufacturing contexts, where more conventional robotic platforms like wheeled mobile manipulators would not be able to perform.

Yet another Horizon 2020 project, AnDy [7] addressed human-robot collaboration in industry. Among its objectives there was the design of collaborative policies for humanoids and cobots to anticipate and assist the human worker. In this project, teleoperation successfully demonstrated to be a very intuitive way to show collaborative policies to humanoids.

### C. Contribution

In all the above projects, humanoids are preferred over more conventional robotic platforms because their structure is a better fit for environments and tasks that are designed for and performed by human workers. This operative versatility makes humanoids suitable for working activities that require a variety of complex movements, such as inspection, maintenance, and interaction with human operators.

Unfortunately, flexibility and adaptability come at the cost of an increased complexity. Planning and controlling

tasks while maintaining balance is a challenging endeavour. Teleoperation can ease the control complexity and facilitate the interaction with the environment. Indeed, in spite of recent progress in robot cognition due to machine learning techniques, fully autonomous solutions are not yet viable. Our viewpoint is that the intuition and intelligence of human operators can be leveraged to make humanoids perform complex tasks, provided that suitable control interfaces and teleoperation modes are designed.

In this work, we present a teleoperation framework for executing loco-manipulation tasks with a humanoid. The proposed control architecture provides two different modes of teleoperation:

- a high-level teleoperation setting in which the operator uses a joystick to send reference commands to the robot such as direction and velocity of motion, without dealing with its actual execution;
- a low-level teleoperation setting in which the operator generates whole-body movements for the robot by means of a motion capture suit (*motion retargeting*).

In both cases the human operator receives visual feedback through a virtual reality (VR) headset connected to the cameras of the robot avatar. In this work, we demonstrate the framework for the teleoperation of the humanoid robot iCub.

The paper is organized as follows. In Sect. II we survey some relevant literature. In Sect. III we discuss first the proposed framework from a general viewpoint; then, in Sect. III-A, we describe the low-level mode based on retargeting, while Sect. III-B discusses the high-level mode that uses MPC for gait generation. Experiments are presented in Sect. IV. Conclusions and future work are outlined in Section V.

## II. RELATED WORK

The idea of teleoperating robots with VR was first proposed by Tachi [8], [9]. The retargeting of the upper-body joints, which is important for manipulation, has often been done independently from the motion generation of the legs, crucial for balancing and locomotion. In [10] for example, the authors employ the mobile manipulator Justin to retarget upper-body motions with haptic feedback at the hands, without considering leg motions (Fig. 3).

Kim *et al* [11] were among the first to extend the robot teleoperation to walking motions. In [11], upper-body motions and walking are separately retargeted onto the humanoid robot Mahru by using a wearable motion capture system. Whenever the operator walks, this triggers a human-independent walking by the robot. For the retargeting of upper-body motions only the arms are involved. Instead, in [12], Hu *et al.* focus on teleoperating exclusively the walking for the humanoid TORO, but considering also the human footsteps and configuration of the leg joints in the retargeting.

In [13], the authors teleoperate the iCub robot in an immersive scenario using a VR headset and a walking platform (Fig. 3). The robot starts and stops walking whenever the

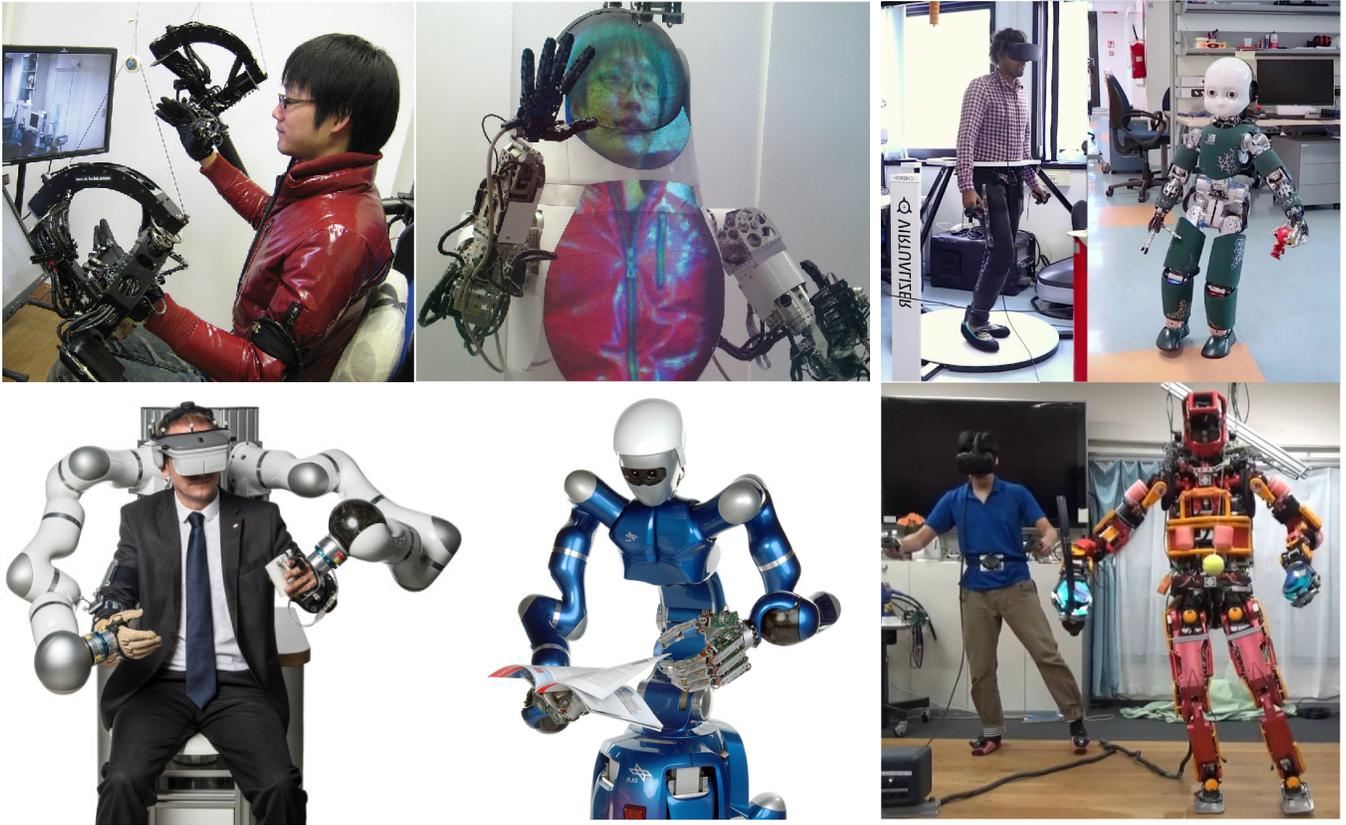


Fig. 3: State-of-the-art in humanoid teleoperation. Top-left: TELESAR II [9], a mutual telexistence system; Bottom-left: Immersive teleoperation of the Justin robot with haptic feedback at the hands [10]; Top-right: Telexistence of the iCub robot with a VR walking platform [13]; Bottom-right: Dynamical whole-body teleoperation of the JAXON robot [14].

operator does, but the retargetable double-support motions are only limited arm movements.

Motion retargeting can be performed also at whole-body level. Ishiguro *et al* [14] conducted some experiments re-targeting highly dynamic upper-body and leg motions onto the humanoid robot JAXON (Fig 3). Although suitable for executing challenging movements, such as kicking or hitting a tennis ball with a racket, their technique cannot be used for motions such as jumping, running, or walking on rough terrain.

The main challenge of teleoperating highly dynamical motions is to ensure smooth and stable motions in real-time while guaranteeing the robot’s balance. Inverse Dynamics approaches would be ideal for dealing with the changing dynamics of the robot during teleoperation, however they are computationally expensive and prone to numerical ill-conditioning. For this reason, the classic approaches, including the works cited above, rely on Inverse Kinematics.

Similar issues must be addressed in robotic walking, where a widespread approach to generate robust dynamic motions exploits Model Predictive Control (MPC) on reduced models of the robotic system. For gait generation the most common strategy relies on the concept of Zero Moment Point (ZMP), i.e., the point with respect to which the horizontal momenta of the ground reaction forces are zero. Dynamic equilibrium

is guaranteed by keeping the ZMP at all times within the robot support polygon, i.e., the convex hull of the contact points.

Many successful techniques for generating stable gaits are based on a simplified linear dynamic model [15] relating the ZMP to the Center of Mass (CoM), derived by neglecting any rotational contribution around the CoM which is also assumed to be at constant height. This model is called the Linear Inverted Pendulum (LIP) or the Cart-Table (CT), depending on whether the ZMP is treated as an input or an output.

In the fundamental work [16] the CT model is used to design a linear quadratic controller with preview. Constraints were added in [17] leading to an MPC formulation and also allowing the automatic choice of the footsteps [18]. To cope with the unstable nature of the LIP, an explicit stability constraint ensuring that the CoM trajectory is bounded w.r.t. the ZMP has been introduced in the MPC design in [19]. Extensions to walk-to locomotion [20] or uneven ground [21] have also been proposed.

### III. MULTI-MODE TELEOPERATION FRAMEWORK

Our framework is illustrated in Fig. 4. The human operator can choose between two different teleoperation modes using the buttons of the VR controller: a *low-level* mode for full real-time control of the robot via motion retargeting and

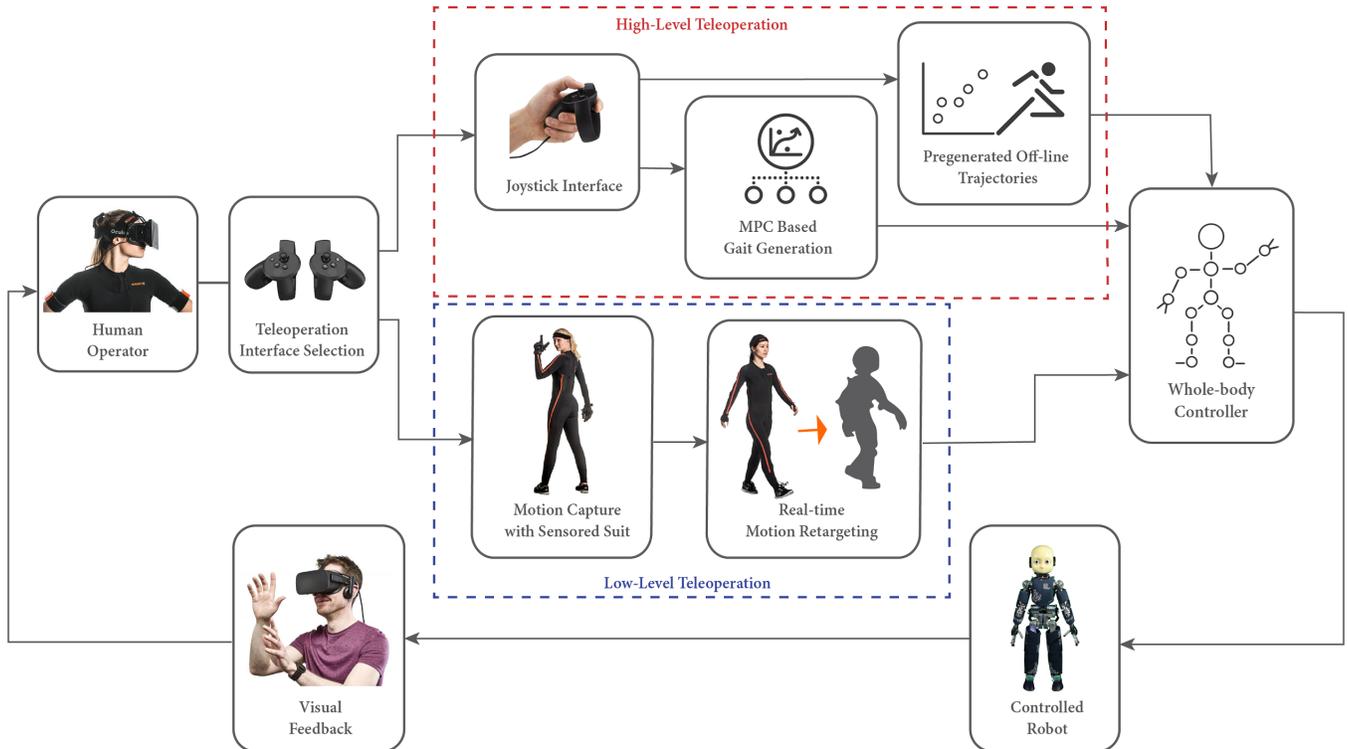


Fig. 4: Overview of the proposed robot teleoperation framework. The user can choose between two operational modes, i.e., high- or low-level teleoperation.

a *high-level* mode to command walking or pre-optimized task trajectories. Both modalities share the same whole-body controller for computing in real-time the commands to be sent to the robot.

In the first operational mode, the human posture is tracked by a motion tracking system, in our case the Xsens motion capture suit. The data are then mapped to feasible corresponding joint values for the robot. To achieve dynamical balance, the references are corrected by a stabilizer and then fed to the whole-body robot controller. The latter is formalized as a multi-task Quadratic Programming (QP) controller, where the task references are the desired robot posture and its stabilized CoM.

The second operational mode is characterized by a higher degree of autonomy of the robot, and it is triggered by the operator using the joystick. For walking, the operator gives a direction and velocity references for the humanoid gait through the analog sticks. These are translated into a timed sequence of footsteps, swinging foot and CoM trajectories through an MPC-based control scheme. Alternatively, the operator uses the joystick buttons to select the one among different pre-defined task trajectories. The corresponding trajectories, which have been pre-optimized off-line, are sent as reference to the robot controller and then simply reproduced.

The reason for the two distinct modes is that retargeting is essentially kinematic and it is not effective for on-line teleoperation of dynamic motions such as walking or stepping. In our experience, retargeting of walking is not a viable

solution for many reasons: first, the stride of the robot is typically shorter than the operator's; second, the robot foot trajectory is often optimized for balance and impacts, while the trajectory retargeted from the human is not compliant with these requirements; third, since humanoids cannot walk as fast as humans, in a retargeting context the operator would be forced to walk in an unnatural way, ultimately leading to inefficient robot locomotion. For these reasons, it is better to rely on MPC-based gait generation and avoid retargeting altogether in this phase. Similar considerations can be made for many motion primitives or task trajectories that impact or leverage the robot dynamics, such as stepping or serving in tennis: in our view, this is the kind of motions that should be pre-optimized off-line as they are specific to the robot dynamics.

The following subsections describe in detail the two teleoperation modes.

#### A. Low-level teleoperation mode

The motion retargeting in the low-level teleoperation mode is built on our previous work [22]. Joint positions are measured and grouped into subcategories: head, torso, left arm, right arm, left leg and right leg. Also the ground projection of the CoM, the height of the waist, the orientation of the head and the position of the feet are controlled.

In the joint retargeting module, the Xsens skeleton degrees of freedom are assigned to the corresponding ones of the iCub robot as shown in Fig. 5. Then, we consider the joint angle variations of the human with respect to the starting

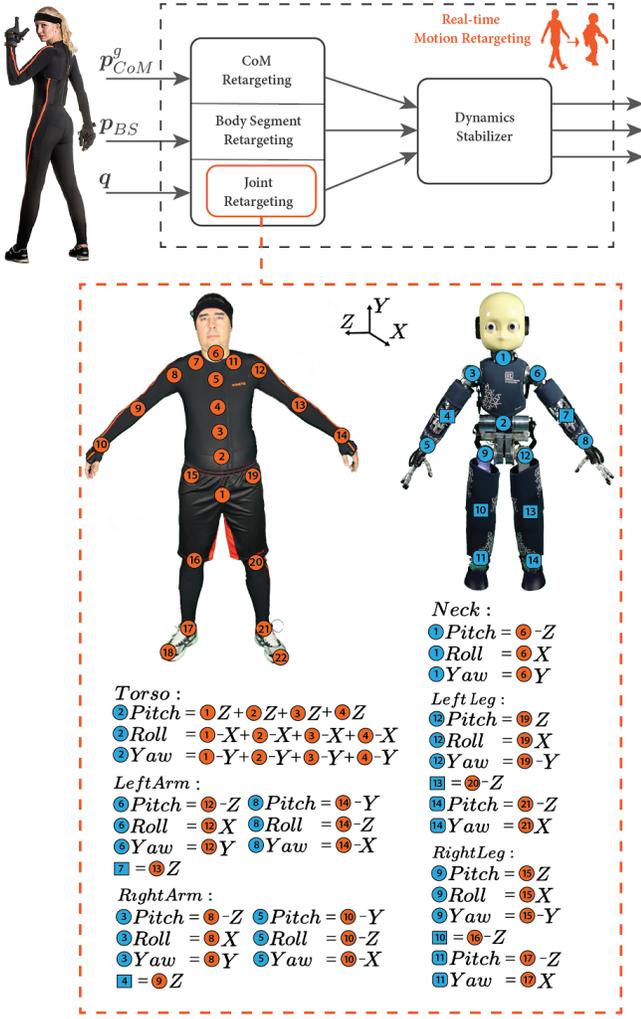


Fig. 5: Pipeline of the whole-body retargeting approach.

posture to compute the corresponding instantaneous values of the robot joint angles

$$\mathbf{q}_R^k = \mathbf{q}_R^0 + (\mathbf{q}_H^k - \mathbf{q}_H^0)$$

where  $\mathbf{q}$  is the joint positions vector, the superscripts 0 and  $k$  refer to measurements at initial time and at time  $k$ , and the subscripts  $H$  and  $R$  indicate measurements on human and robot, respectively.

The same approach is used to retarget the relative Cartesian position  $\mathbf{p}_{BS}$  of a body segment with respect to a base link, with the difference that the variation of the human positions has to be properly scaled by the human-robot limb length ratio, as explained in [22].

To track the human CoM we use normalized offsets, from which we then reconstruct the robot CoM ground position. We consider the ground projection of the human CoM  $\mathbf{p}_{CoM}^g$ . Its position with respect to an arbitrary foot (let us say the left) is projected onto the line connecting the two feet. The result is then normalized to obtain an offset  $o \in [0, 1]$

$$o = \frac{(\mathbf{p}_{CoM}^g - \mathbf{p}_{lFoot}^g) \cdot (\mathbf{p}_{rFoot}^g - \mathbf{p}_{lFoot}^g)}{\|\mathbf{p}_{rFoot}^g - \mathbf{p}_{lFoot}^g\|^2}$$

where  $\mathbf{p}_{lFoot}^g$  and  $\mathbf{p}_{rFoot}^g$  are the ground projections of the left and right foot respectively. When the human is in a symmetric pose, the offset  $o$  has a value around 0.5 and when the human stands on a single foot, it is either 0 (left foot) or 1 (right foot). The robot CoM ground projection is then reconstructed on the line connecting its feet by means of this offset value. To retarget also changes of the CoM that are not on the line connecting the feet, we can apply the same concept while considering the maximum backward and forward CoM displacement in the orthogonal direction of the line connecting the feet as done in [22].

The resultant retargeted motion it is not guaranteed to be dynamically balanced and different stabilizers can be used to correct it (Fig. 5). In our teleoperation approach we want a dynamically balanced CoM trajectory and we adopt the LIP model to properly modify the reference trajectory.

We previously recalled that balancing moments leads to the definition of the ZMP. Dynamic balance is enforced by keeping this point at all times within the robot support polygon. By neglecting rotational terms and assuming a constant height  $h_{CoM}$  for the CoM, the moment balance equation of the robot leads to the LIP,

$$x_{ZMP} = x_{CoM} - \frac{1}{\eta^2} \ddot{x}_{CoM} \quad (1)$$

where  $\eta = \sqrt{g/h_{CoM}}$ , with  $g$  the gravitational constant, while  $x_{CoM}$  and  $x_{ZMP}$  are respectively the CoM and ZMP positions along  $x$  (similarly for  $y$ ).

By employing the dynamic equation of the LIP it is possible to set up a QP optimization problem that provides at each control iteration a correction of the desired CoM that satisfies the balance condition on the humanoid

$$\min_{x_{ZMP}} (\dot{x}_{CoM}^{ref} - \dot{x}_{CoM})^2$$

$$\text{subject to: } \dot{x}_{CoM} = \dot{x}_{CoM}^m + \frac{\delta g}{h_{CoM}^m} (x_{CoM} - x_{ZMP})$$

$$x_{ZMP}^{min} < x_{ZMP} < x_{ZMP}^{max}$$

where  $\dot{x}_{CoM}^{ref}$  is the reference human retargeted CoM velocity,  $\delta$  is the sampling time,  $\dot{x}_{CoM}^m, h_{CoM}^m$  are respectively the last CoM and the last CoM height measured from the robot,  $x_{ZMP}^{min}$  and  $x_{ZMP}^{max}$  are the lower and upper bound of the support polygon of the robot and the first constraint is derived from (1) using the Euler approximation.

The stabilized CoM reference, Cartesian tasks and postural tasks are set as reference tasks in the multi-task QP controller: at each time step, it solves a linearly constrained QP optimization problem to minimize the given cost function characterizing the motion tracking, subject to the system constraints such as the joint and torque limits. More detail about the QP controller can be found in [23].

### B. High-level Teleoperation Mode

As illustrated in Fig. 4, the humanoid desired motion is either defined offline (pre-generated task trajectories or action primitives [24]) or online (MPC-based gait generator).

Pre-generated trajectories are triggered by the operator depending on the situation. For example, in the final phase

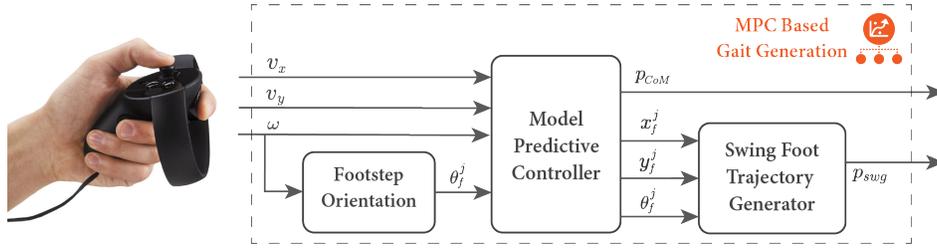


Fig. 6: Pipeline of the MPC-based gait generation module.

of the reported experiment the robot moves its feet apart autonomously, independently of the human's lower limbs motion, for picking up the box more effectively.

Other tasks or repetitive movements like standing up, can be recorded off-line and replicated for a quick and precise execution. We also recommend the use of this option to make the robot perform motions that are not ergonomic for the operator or can be uncomfortable to perform. Note also that the execution of pre-defined motions may be more convenient in the presence of signal degradation or delay.

The online gait generation is based on the MPC scheme proposed in [19], [25] and is summarized in Fig. 6. The joystick provides reference velocities  $(v_x, v_y)$  and  $\omega$  for respectively the sagittal, coronal and angular motions. Footstep orientations are computed in a separate stage and used as known parameters in the next module. This is useful to guarantee linearity of the constraints in the following MPC formulation.

To generate the footstep orientations a first QP problem

$$\begin{aligned}
 \text{(QP1):} \quad & \min_{\Theta_f^k} \sum_{j=1}^F (\theta_f^j - \theta_f^{j-1} - \omega T_s)^2 \\
 \text{subject to} \quad & |\theta_f^j - \theta_f^{j-1}| \leq \theta_{\max}
 \end{aligned}$$

is solved for all the  $F$  predicted footsteps that fall within the prediction horizon of the MPC problem. These are denoted as  $\Theta_f^k = (\theta_f^1, \dots, \theta_f^F)$ , while  $T_s$  is the step duration and  $\theta_{\max}$  the maximum allowed rotation between two consecutive steps.

A second module based on MPC generates the CoM trajectory and the positions of the footsteps, while their orientations are inherited from the first module. At each time instant  $t_k$ , the MPC solves a QP problem on a prediction horizon  $[t_k, t_{k+C}]$  based on a prediction model and constraints. The sampling interval has duration  $\delta$ . We use as prediction model the LIP (1) with an additional integrator on the input (dynamic extension), so that the ZMP velocities  $(\dot{x}_{ZMP}, \dot{y}_{ZMP})$  are used as control input. The decision variables of the QP problem are therefore  $\dot{x}_{ZMP}^k, \dots, \dot{x}_{ZMP}^{k+C-1}, \dot{y}_{ZMP}^k, \dots, \dot{y}_{ZMP}^{k+C-1}$  and the footstep positions  $\dot{x}_f^1, \dots, \dot{x}_f^F, \dot{y}_f^1, \dots, \dot{y}_f^F$ , which are collected in the vector  $U^k$  of decision variables. The ZMP velocities are assumed to be constant in each sampling interval  $[t_i, t_{i+1}]$  resulting in a piece-wise linear ZMP.

The constraints enforced in the second QP are:

- a *balance* constraint, which ensures that the ZMP is at all times inside the support polygon;
- a *kinematic* constraint, guaranteeing that the footsteps are placed in a region which is kinematically feasible, and avoid self-collisions;
- a *stability* constraint, which makes sure that the generated CoM trajectory does not diverge with respect to the ZMP.

The balance or ZMP constraint at a generic instant of the single support is expressed as

$$R_j^T \begin{pmatrix} x_{ZMP}^{k+i} - x_f^j \\ y_{ZMP}^{k+i} - y_f^j \end{pmatrix} \leq \frac{1}{2} \begin{pmatrix} d_{z,x} \\ d_{z,y} \end{pmatrix} \quad (2)$$

where  $d_{z,x}$  and  $d_{z,y}$  denote the size of the rectangular support polygon while  $(x_{ZMP}^{k+i}, y_{ZMP}^{k+i})$  is the ZMP position at the  $i$ -th prediction instant.  $R_j^T$  is the rotation matrix associated with the orientation of the  $j$ -th predicted footstep within the prediction horizon, which is computed in the Footstep Orientation block of Fig. 6. The ZMP constraint is enforced at each instant of the prediction horizon, i.e., for  $i = 1, \dots, C$ .

The kinematic constraint ensures that the footsteps are placed consistently with the robot capabilities. The constraint is

$$R_{j-1}^T \begin{pmatrix} x_f^j - x_f^{j-1} \\ y_f^j - y_f^{j-1} \end{pmatrix} \leq \pm \begin{pmatrix} 0 \\ \ell \end{pmatrix} + \frac{1}{2} \begin{pmatrix} d_{k,x} \\ d_{k,y} \end{pmatrix} \quad (3)$$

where  $d_{k,x}$  and  $d_{k,y}$  represent the size of a rectangular region which is kinematically feasible and avoids self-collisions, and  $\ell$  the position of its center w.r.t. the previous footstep. The  $\pm$  sign regularly alternates discriminating between left and right footsteps.

To understand the necessity of the stability constraint, note that the LIP model (1), and hence the prediction model, has an unstable eigenvalue; therefore, the generic CoM trajectory will in general diverge w.r.t. the ZMP. There exists however a *stability condition*, relating the CoM initial state in  $t_k$  to the future ZMP, expressed as

$$x_{CoM}^k + \frac{\dot{x}_{CoM}^k}{\eta} = \eta \int_{t_k}^{\infty} e^{-\eta(\tau-t_k)} x_{ZMP}(\tau) d\tau \quad (4)$$

which ensures that the CoM trajectory remains bounded with respect to the ZMP. A similar expression holds along  $y$ .

The stability constraint of the MPC is obtained by computing (4) for a piece-wise linear ZMP trajectory. However,

note that the integral requires the future ZMP trajectory, which is available only up to the prediction horizon. The remaining part, after  $t_{k+C}$ , can be conjectured by using the available information (e.g., planned reference velocities) beyond the prediction horizon. One possibility is to use an *infinite replication* of the control inputs over the horizon, which is especially appropriate for regular gaits that exhibit a periodic behavior; the resulting stability constraint is

$$\sum_{i=0}^{C-1} e^{-i\eta\delta} \dot{x}_{ZMP}^{k+i} = \eta \frac{1 - e^{-C\eta\delta}}{1 - e^{-\eta\delta}} (x_{CoM}^k + \frac{\dot{x}_{CoM}^k}{\eta} - x_{ZMP}^k) \quad (5)$$

The second QP problem can finally be stated as

$$\begin{aligned} \text{(QP2):} \quad & \min_{U^k} \sum_{i=0}^{C-1} \left( (\dot{x}_{ZMP}^k)^2 + (\dot{y}_{ZMP}^k)^2 + \right. \\ & \beta_x (x_{CoM} - v_x \cos(\omega\delta i) + v_y \sin(\omega\delta i))^2 + \\ & \left. \beta_y (y_{CoM} - v_x \sin(\omega\delta i) - v_y \cos(\omega\delta i))^2 \right) \end{aligned}$$

subject to:

- ZMP constraints (2)
- kinematic constraints (3)
- stability constraint (5) for  $x$  and  $y$

In the cost function,  $\beta_x$  and  $\beta_y$  represent the weights associated with the velocity tracking terms.

Once both QP problems are solved, the first value of the ZMP derivative  $(\dot{x}_{ZMP}^k, \dot{y}_{ZMP}^k)$  is used to compute the CoM trajectory  $p_{CoM}$  through the prediction model, while the first predicted footstep  $(x_f^1, y_f^1, \theta_f^1)$  is used to generate, using a predefined polynomial shape, a swing foot trajectory  $p_{swg}$  ending at the predicted footstep. Both  $p_{CoM}$  and  $p_{swg}$  are finally tracked by the kinematic controller.

#### IV. EXPERIMENTS

This section presents an illustrative experiment performed with a human operator and the iCub humanoid robot<sup>1</sup>. In our presented scenario shown in Fig. 7, the robot must walk and pick a box on the floor to hand it to the worker. The operator is equipped with a wearable motion capture Xsens MVN suit and the VR Oculus Rift system composed by a headset and a pair of joysticks. The suit provides real-time estimation of the posture, the headset gives visual feedback from the robot cameras, while the joysticks allow the operator to switch among the two different teleoperation modes and to guide the robot.

As a first step, the robot should autonomously walk to the box, guided by the operator. After selecting the high-level mode on the VR controller, the robot walks receiving velocity references commands through the joystick. Figure 8 shows the MPC-generated CoM trajectory (to be sent to the whole-body controller) together with the footsteps.

When the robot gets in front of the box, the operator stops it. Then, still in high-level mode, the operator prepares

<sup>1</sup>iCub is only 104 cm high and cannot lift heavy weights, but our methods are not iCub-specific and could be easily applied to adult-sized humanoids with heavier payloads.



Fig. 7: Top: robot walking in the high-level mode. Bottom: robot controlled in the low-level mode with a motion capture suit and a VR headset.

himself for the pickup by moving his feet apart, and selects a pre-defined motion for the humanoid to do the same movement independently.

Finally, the operator switches to the low-level mode by using the VR controller and performs a squat motion to pick up the box. Motion retargeting makes the robot follow in real time the movements of the operator, successfully handing the box over to the human. Balance is maintained throughout this phase thanks to the stabilization performed on the retargeted human references, as shown in Fig. 9. A video clip of this experiment is available online<sup>2</sup>.

#### V. CONCLUSIONS

In this work we proposed a multi-mode teleoperation framework for a humanoid robot on loco-manipulation tasks. The first mode is a low-level teleoperation of all the joints of the robot, while the other enables the execution of high-level commands and pre-designed motion primitives, which can be useful for locomotion or other specific tasks.

The use case presented in the experiment consists of walking to a box on the ground using the high-level mode, and then picking it up by switching to the low-level mode. It is possible to envision several scenarios in which the presented framework might be employed. For example, the operator might switch to the low-level mode to use control panels, open doors or recover items, all actions that could be necessary during exploration. The robot might also be equipped to execute specific actions necessary for maintenance operations (e.g., tightening screws, assembling parts, etc...) using specialized tools activated from the joystick. In the future, we will test these scenarios with adult-sized humanoid robots.

<sup>2</sup>[https://www.youtube.com/playlist?list=PLaViA12WLPmFdT-w2s0Qz7y1\\_e-JEWNsJ](https://www.youtube.com/playlist?list=PLaViA12WLPmFdT-w2s0Qz7y1_e-JEWNsJ)

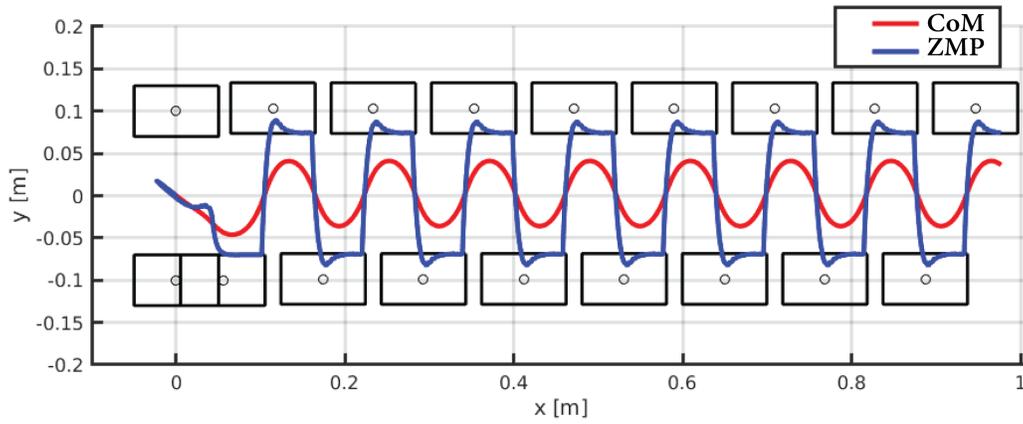


Fig. 8: Walking phase: CoM and ZMP trajectories along the MPC generated gait.

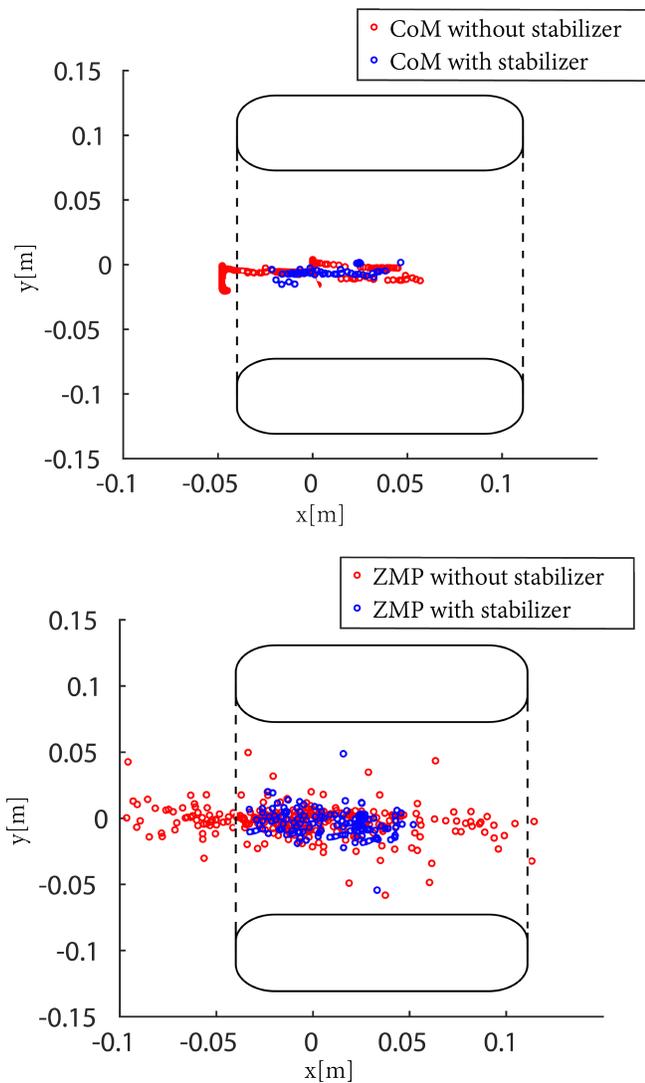


Fig. 9: Pickup phase: the feet of the robot are shown in black. Top: CoM reference position reconstructed from the human reference (red) and corresponding stabilized value (blue). Bottom: ZMP of the LIP model associated to the CoM reference without (red) and with (blue) stabilization.

A current limitation of this framework, that prevents us to address more complicated teleoperation scenarios, is the absence of haptic feedback. While safe physical interaction with the environment may be ensured by the robot low-level control, haptic feedback is still critical to enhance the remote control capabilities of the operator. Typically, haptic feedback is localized in the end-effector, where most of the interaction occurs; in the case of teleoperated humanoids, whole-body haptic feedback should be considered, possibly by means of wearable vibro-tactile devices.

## REFERENCES

- [1] OECD, *Occupational accidents in oecd countries, OECD Employment Outlook 1989, chapter 4*, 1989. [Online]. Available: <http://www.oecd.org/els/emp/3888265.pdf>
- [2] P. Hmlinen, J. Takala, and K. Saarela, "Global estimates of occupational accidents," *Safety Science*, vol. 44, pp. 137–156, 02 2006.
- [3] J. Takala, P. Hämäläinen, K. L. Saarela, L. Y. Yun, K. Manickam, T. W. Jin, P. Heng, C. Tjong, L. G. Kheng, S. Lim, and G. S. Lin, "Global estimates of the burden of injury and illness at work in 2012," *Journal of Occupational and Environmental Hygiene*, vol. 11, no. 5, pp. 326–337, 2014.
- [4] "Census of fatal occupational injuries summary, 2017." [Online]. Available: <https://www.bls.gov/news.release/cfoi.nr0.htm>
- [5] "WALK-MAN: Whole-body Adaptive Locomotion and Manipulation, EU Project." [Online]. Available: <https://www.walk-man.eu/>
- [6] "COMANOID: Multi-Contact Collaborative Humanoids in Aircraft Manufacturing, EU Project." [Online]. Available: <http://comanoid.cnrs.fr/>
- [7] S. Ivaldi, L. Fritzsche, J. Babič, F. Stulp, M. Damsgaard, B. Graitmann, G. Belluscì, and F. Nori, "Anticipatory models of human movements and dynamics: the roadmap of the AnDy project," in *DHM*, 2017.
- [8] S. Tachi, *Telexistence*. World Scientific 2nd edition, 2015.
- [9] M. Y. Saraiji, C. L. Fernando, M. Furukawa, K. Minamizawa, and S. Tachi, "Virtual telesar - designing and implementation of a modular based immersive virtual telexistence platform," in *2012 IEEE/SICE International Symposium on System Integration (SII)*, Dec 2012, pp. 595–598.
- [10] N. Y. Lii, D. S. Leidner, A. Schiele, P. Birkenkamp, R. Bayer, B. Pleintinger, A. Meissner, and A. Balzer, "Simulating an extraterrestrial environment for robotic space exploration: The meteron supervis-justin telerobotic experiment and the solex proving ground," 2015.
- [11] 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.
- [12] K. Hu, C. Ott, and D. Lee, "Online human walking imitation in task and joint space based on quadratic programming," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, May 2014, pp. 3458–3464.

- [13] M. Elobaid, Y. Hu, G. Romualdi, S. Dafarra, J. Babic, and D. Pucci, "Telexistence and teleoperation for walking humanoid robots," *Intel-liSys*, 2019.
- [14] Y. Ishiguro, K. Kojima, F. Sugai, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba, "High speed whole body dynamic motion experiment with real time master-slave humanoid robot system," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, May 2018, pp. 1–7.
- [15] S. Kajita, H. Hirukawa, K. Harada, and K. Yokoi, *Introduction to Humanoid Robotics*. Springer Publishing Company Inc., 2014.
- [16] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa, "Biped walking pattern generation by using preview control of zero-moment point," in *2003 IEEE Int. Conf. on Robotics and Automation*, 2003, pp. 1620–1626.
- [17] P.-B. Wieber, "Trajectory free linear model predictive control for stable walking in the presence of strong perturbations," in *2006 IEEE-RAS Int. Conf. on Humanoid Robots*, 2006, pp. 137–142.
- [18] A. Herdt, H. Diedam, P.-B. Wieber, D. Dimitrov, K. Mombaur, and M. Diehl, "Online walking motion generation with automatic footstep placement," *Advanced Robotics*, vol. 24, no. 5-6, pp. 719–737, 2010.
- [19] N. Scianca, M. Cognetti, D. De Simone, L. Lanari, and G. Oriolo, "Intrinsically stable MPC for humanoid gait generation," in *16th IEEE-RAS Int. Conf. on Humanoid Robots*, 2016, pp. 101–108.
- [20] A. Aboudonia, N. Scianca, D. De Simone, L. Lanari, and G. Oriolo, "Humanoid gait generation for walk-to locomotion using single-stage MPC," in *17th IEEE-RAS Int. Conf. on Humanoid Robots*, 2017, pp. 178–183.
- [21] A. Zamparelli, N. Scianca, L. Lanari, and G. Oriolo, "Humanoid gait generation on uneven ground using Intrinsically Stable MPC," in *12th IFAC Symposium on Robot Control*, 2018.
- [22] L. Penco, B. E. P. Clement, V. Modugno, E. M. Hoffman, G. Nava, D. Pucci, N. G. Tsagarakis, J.-B. Mouret, and S. Ivaldi, "Robust real-time whole-body motion retargeting from human to humanoid," *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*, pp. 425–432, 2018.
- [23] E. M. Hoffman, A. Rocchi, A. Laurenzi, and N. G. Tsagarakis, "Robot control for dummies: Insights and examples using OpenSoT," in *HUMANOIDS*, 2017, pp. 736–741.
- [24] J. Kober and J. Peters, "Learning motor primitives for robotics," in *2009 IEEE International Conference on Robotics and Automation*, May 2009, pp. 2112–2118.
- [25] N. Scianca, D. De Simone, L. Lanari, and G. Oriolo, "MPC for humanoid gait generation: Stability and feasibility," 2019. [Online]. Available: <https://arxiv.org/abs/1901.08505>