



HAL
open science

Exploiting visual servoing and centroidal momentum for whole-body motion control of humanoid robots in absence of contacts and gravity

Enrico Mingo Hoffman, Antonio Paolillo

► **To cite this version:**

Enrico Mingo Hoffman, Antonio Paolillo. Exploiting visual servoing and centroidal momentum for whole-body motion control of humanoid robots in absence of contacts and gravity. 2021 IEEE International Conference on Robotics and Automation (ICRA), May 2021, Xi'an, China. pp.2979-2985, 10.1109/ICRA48506.2021.9560739 . hal-03166288

HAL Id: hal-03166288

<https://hal.science/hal-03166288>

Submitted on 11 Mar 2021

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.

Exploiting visual servoing and centroidal momentum for whole-body motion control of humanoid robots in absence of contacts and gravity

Enrico Mingo Hoffman¹ and Antonio Paolillo²

Abstract—The big potential of humanoid robots is not restricted to the ground, but these versatile machines can be successfully employed in unconventional scenarios, e.g. space, where contacts are not always present. In these situations, the robot’s limbs can be used to assist or even generate the angular motion of the floating base, as a consequence of the centroidal momentum conservation. In this paper, we propose to combine, in the same whole-body motion control, visual servoing and centroidal momentum conservation. The former dictates a rotation to the floating humanoid to achieve a task in the Cartesian space; the latter is exploited to realize the desired rotation by moving the robot’s articulations. Simulations in a space scenario are carried out using COMAN, a humanoid robot developed at the Istituto Italiano di Tecnologia.

I. INTRODUCTION

Humanoid robots have the great advantage to be general-purpose platforms resembling human morphology. With no need of customization, they can operate in domestic [1] or industrial scenarios [2] and manipulate devices (e.g., a car [3]) tailored to humans. Many impressive humanoid platforms and dedicated software suites have been proposed, reducing the gap between research development and real-world deployment [4]. The big potential of these machines and their domain of applicability are not restricted to the ground, but humanoids can be successfully employed underwater [5], in the air [6] or space [7]. These scenarios opened new opportunities, posing challenges in different perspectives, from mechanical and sensory equipment design to actuation methods, to control paradigms. To facilitate their mobility in these domains, robots use propellers or jet engines. Nonetheless, the articulations could be exploited to generate or assist the Floating-Base (FB) motion, especially in absence of contacts. In fact, the nonholonomy of the angular momentum conservation allows to control the robot orientation in absence of contacts by moving its articulations, like a cat able to land on its feet after falling from an upside posture [8]. The same effect is also experienced in microgravity conditions by astronauts, who can change their body orientation with cyclic motions of arms and legs (see Fig. 1). A humanoid robot capable to perform this kind of behaviors would have obvious advantages for achieving tasks in the air or space. Furthermore, the articulations motion could be exploited for the sake of energy efficiency. In fact, most rotations may be achieved by controlling the motion of

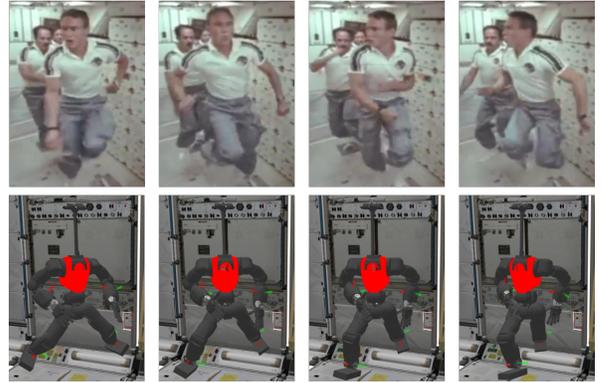


Fig. 1. In absence of contacts, articulated systems can rotate by moving the articulations, as effect of the nonholonomic nature of the angular momentum conservation. This phenomenon is experienced by NASA astronauts in microgravity conditions and replicated with humanoid robots in this work.

the limbs, i.e., using the robot battery. This strategy would save the precious fuel needed to run propeller or jet engines, thus extending the operational time.

As a particular case study, let us consider a space scenario, where there is absence of gravity and contacts. In space, Intra/Extravehicular Activities (IEVAs) are required for the successful execution of a mission. During IEVAs, vision represents a valuable source of information. Indeed, an autonomous system can naturally observe its surrounding to find visual references for the motion. For example, a floating humanoid might need to align with a tool to be further manipulated, for rescue or maintenance purposes. In practice, a Visual Servoing (VS) task [9] can be defined to rotate the robot towards the tool; the Centroidal Momentum Conservation (CMC) of the robot can be exploited to move the limbs and realize the desired rotation. Therefore, CMC represents a way to fulfill VS tasks with a floating humanoid robot, in absence of gravity and contacts.

Inspired by these considerations, we aim to extend the whole-body motion control of humanoids with the ability of changing their body orientation, e.g. tracking proper visual features, using the limbs motion in scenarios where contacts may not occur. Building on the current literature (detailed in Sect. II), we investigate the possibility of integrating VS and CMC under the same control framework exploiting Quadratic Programming (QP) optimization (recalled in Sect. III). Our approach, presented in Sect. IV, consists in designing a proper stack of tasks and constraints that allows a humanoid robot to control its body orientation in absence of gravity and contacts. Section V reports on the simulation results, while Sect. VI concludes the paper and discusses future work.

¹Humanoids & Human Centred Mechatronics (HHCMM) Lab., Istituto Italiano di Tecnologia (IIT), Genova, Italy enrico.mingo@iit.it

²Dalle Molle Institute for Artificial Intelligence (IDSIA), USI-SUPSI, Lugano, Switzerland antonio.paolillo@idsia.ch

Both the authors equally contributed to this work.

II. STATE OF THE ART

Robots are employed in space to help and potentially substitute astronauts [10]. Usually, they consist of a humanoid upper body with an articulated link to keep the contact with the station for safety sake, which also allows the control of the torso attitude [11]. A floating full humanoid [6], instead, could reach a higher level of freedom of the FB motion, leaving to a rope the role of safety. Here, CMC can be used to govern the FB attitude, assisted by VS using visual references to decide when to start and stop moving.

Control of Centroidal Momentum (CM) has been introduced for stabilizing the balance of humanoid robots under external disturbances [12], and it is often used in the optimization framework for humanoids. For example, in [13], CM is encapsulated into a motion controller using conic optimization to compute the inverse dynamics. A QP-based controller at velocity level includes the CM minimization to perform physio-therapeutic juggling with a humanoid in [14]. Concerning VS [9], [15], it is worth mentioning that it can be integrated in the optimization-based whole-body motion control of humanoids, as done within a QP in [16], and in [17] for estimating the configuration of an object to be manipulated. In [18], VS is integrated in a differential dynamic programming framework, which takes into account also the robot centroidal dynamics.

The benefit obtained by the integration of VS and momentum conservation has been shown by only few works. In [19], an image-based controller to perform the guidance of a free-floating robot manipulator is presented. In [20], the controller of a satellite equipped with a dual arm robot considers VS as the primary task and the angular momentum conservation of the base as a secondary task, to reduce the attitude disturbances. However this choice of priority may result in non-consistent motions of the base.

Even if these approaches show the potential of combining VS and CMC, to the best of the authors knowledge, there are no works integrating them in the same whole-body QP-based prioritized control of floating humanoids. Based on these considerations, the main contributions of this work are:

- a comprehensive formulation of a stack of tasks and constraints, exploiting VS and CMC, for any QP-based controller of humanoid robots at velocity level (the same results can be easily extended to the acceleration level),
- the achievement of visual tasks with floating humanoids in situations where no external forces occur.

III. BACKGROUND

Let us consider a humanoid robot with n Degrees of Freedom (DoFs). Its configuration can be expressed by:

$$\mathbf{q} = \begin{bmatrix} \mathbf{q}_b \\ \mathbf{q}_a \end{bmatrix} \in \mathbb{R}^{d+n} \quad (1)$$

where $\mathbf{q}_a \in \mathbb{R}^n$ contains the joints position of the articulations, while $\mathbf{q}_b \in \mathbb{R}^d$ is a representation the FB pose, obtained by using, e.g., a virtual chain of three prismatic and three revolute joints ($d = 6$) or a position and a quaternion

vector ($d = 7$). If we model the FB pose with the virtual chain, the configuration space velocities are

$$\dot{\mathbf{q}} = \begin{bmatrix} \dot{\mathbf{q}}_b \\ \dot{\mathbf{q}}_a \end{bmatrix} \in \mathbb{R}^{6+n}. \quad (2)$$

A. Whole-body motion control

The Cartesian velocity ${}^w\mathbf{v}_e \in \mathbb{R}^6$ of an end-effector frame \mathcal{F}_e w.r.t. a reference \mathcal{F}_w is related to $\dot{\mathbf{q}}$ through the relation

$${}^w\mathbf{v}_e = {}^w\mathbf{J}_{w,e} \dot{\mathbf{q}}, \quad (3)$$

where ${}^w\mathbf{J}_{w,e} \in \mathbb{R}^{6 \times (6+n)}$ is the Jacobian¹ of the frame \mathcal{F}_e w.r.t. \mathcal{F}_w expressed in \mathcal{F}_w . One can consider the problem of computing desired $\dot{\mathbf{q}}$ from a desired ${}^w\mathbf{v}_e$ by inverting (3) using the well-known pseudo-inverse techniques. The high level of redundancy, associated to humanoid robots, has the advantage to permit the achievement of multiple tasks at the same time. The complexity of simultaneously handling multiple tasks has been addressed formulating the inversion of (3) as a QP optimization, with the possibility to take into account linear constraints:

$$\begin{aligned} \dot{\mathbf{q}}^* = \arg \min_{\dot{\mathbf{q}}} \quad & \| \mathbf{J} \dot{\mathbf{q}} - \mathbf{v}_r \|_{\mathbf{W}}^2 + \epsilon \| \dot{\mathbf{q}} \|^2 \\ \text{s.t.} \quad & \underline{\mathbf{u}} \leq \mathbf{C} \dot{\mathbf{q}} \leq \bar{\mathbf{u}}, \end{aligned} \quad (4)$$

where $\mathbf{J} \in \mathbb{R}^{m \times (6+n)}$ and $\mathbf{v}_r \in \mathbb{R}^m$ are the task Jacobian and the reference velocity, respectively. For Cartesian tasks, \mathbf{v}_r is normally defined as ${}^w\mathbf{v}_{e,d} + \lambda({}^w\mathbf{p}_{e,d} - {}^w\mathbf{p}_e)$, where ${}^w\mathbf{p}_e$ is the current pose of the end-effector involved in the task whereas ${}^w\mathbf{p}_{e,d}$ and ${}^w\mathbf{v}_{e,d}$ are the desired pose and velocity of the end-effector, respectively²; $\lambda \geq 0$ is a control gain used to tune the convergence rate of the task. It is also possible to set postural tasks, in which case the Jacobian is an identity matrix, and the reference velocity is $\dot{\mathbf{q}}_d + \lambda(\mathbf{q}_d - \mathbf{q})$. $\mathbf{W} \in \mathbb{R}^{m \times m}$ is a positive-definite weight matrix setting the relative *soft* priority between multiple tasks. In the cost function of (4), ϵ is a positive regularisation term which minimize the joint velocities, particularly usefull in the neighborhood of a singularity. Employing QP has the advantage to easily set linear equality and inequality constraints over the optimization variable $\dot{\mathbf{q}}$. In (4), $\underline{\mathbf{u}}$ and $\bar{\mathbf{u}} \in \mathbb{R}^c$ are the lower and upper bounds of the constraints while $\mathbf{C} \in \mathbb{R}^{c \times (6+n)}$ is the constraint matrix. It is also possible to consider *hard* priorities between tasks solving a cascade of QPs of the same form of (4), where each QP carries the optimality of the previous solved tasks in form of an additional set of equality constraints. This technique is known as *inequality Hierarchical QP* (iHQP) [21]. For position-controlled robots, the solution $\dot{\mathbf{q}}^*$ is integrated and sent in open-loop to the low-level PD joint positions controller.

Several tasks and constraints can be concisely expressed by using the so-called *Math of Tasks* (MoT) formalism [22], which defines a stack as a set \mathcal{S} of tasks \mathcal{T}_i and constraints \mathcal{C}_i . Each \mathcal{T}_i contains information on the corresponding task

¹Here and in what follows, for the sake of brevity of the notation, we do not express the dependence of the matrices on the configuration vector \mathbf{q} .

²The sign ‘-’ denotes the operator to properly define the error between Cartesian poses, obeying the rules of the chosen orientation representation.

(e.g., gain, weight, control frame) as well as \mathcal{C}_i does for the constraint (e.g., the bounds of the joint limits). The *augment* operator ‘+’ is used to denote soft priority between tasks (e.g., $\mathcal{T}_1 + \mathcal{T}_2$) and concatenate constraints ($\mathcal{C}_1 + \mathcal{C}_2$). Hard priority between two tasks writes as $\mathcal{T}_1/\mathcal{T}_2$, meaning that task 2 is solved respecting the optimality of task 1. Constraints are considered using the *insert* operator ‘ \ll ’. As an example, $\mathcal{S} = (((\mathcal{T}_1 + \mathcal{T}_2)/\mathcal{T}_3) \ll (\mathcal{C}_1 + \mathcal{C}_2))$ says that \mathcal{S} has 3 tasks and 2 constraints; \mathcal{T}_3 keeps optimal \mathcal{T}_1 and \mathcal{T}_2 , between which there is a soft priority regime.

B. Visual servoing

VS is a mature technique used to control robots with the information extracted from camera images [9]. It computes velocity commands to zero the error between f visual features \mathbf{s} and their desired values \mathbf{s}_d . Assuming a fixed visual target, VS is built on the relationship relating the motion of the visual feature to the $\dot{\mathbf{q}}$ vector [15]:

$$\dot{\mathbf{s}} = \mathbf{J}_s \dot{\mathbf{q}} \quad (5)$$

where $\mathbf{J}_s = \mathbf{L}^c \mathbf{V}_h^h \mathbf{J}_{w,h} \in \mathbb{R}^{f \times (n+6)}$ is the feature Jacobian, being: $\mathbf{L} \in \mathbb{R}^{f \times 6}$ the interaction matrix, relating the time derivative of the features to the camera velocity in the camera frame \mathcal{F}_c ; ${}^c\mathbf{V}_h \in \mathbb{R}^{6 \times 6}$ the constant velocity twist transformation from \mathcal{F}_c to the control frame \mathcal{F}_h ; ${}^h\mathbf{J}_{w,h} \in \mathbb{R}^{6 \times (6+n)}$ the Jacobian of \mathcal{F}_h , which is rigidly connected to \mathcal{F}_c ³. As in the kinematic case (Sect. III-A), one can define the reference velocity of the visual features as composed of a feed-forward and an error term: $\dot{\mathbf{s}}_r = \dot{\mathbf{s}}_d + \lambda(\mathbf{s}_d - \mathbf{s})$.

C. Centroidal momentum in absence of contacts and gravity

The centroidal momentum $\mathbf{h} \in \mathbb{R}^6$ of the robot is the resultant momentum acting on its CoM, related to $\dot{\mathbf{q}}$ through the following equation [23]:

$$\mathbf{h} = \mathbf{A} \dot{\mathbf{q}}, \quad (6)$$

where $\mathbf{A} \in \mathbb{R}^{6 \times (6+n)}$ is the so-called centroidal momentum matrix that can be efficiently computed according to [24]. For the Newton’s equations of motion, we can write:

$$\dot{\mathbf{h}} = \mathbf{f} - \mathbf{g}, \quad (7)$$

where $\dot{\mathbf{h}} \in \mathbb{R}^6$ is the variation of the robot centroidal momentum, \mathbf{f} and $\mathbf{g} \in \mathbb{R}^6$ are the resultant wrench at the robot’s Center of Mass (CoM) due to external forces and the gravity, respectively. In absence of contacts and gravity, (7) reduces to $\dot{\mathbf{h}} = \mathbf{0}$, implying that \mathbf{h} is constant. If the system starts from rest, it is $\mathbf{h} = \mathbf{0}$ which, substituted in (6), gives

$$\mathbf{A} \dot{\mathbf{q}} = \mathbf{0}. \quad (8)$$

This equation expresses the robot CMC in absence of contacts and gravity. As remarked in [8], its linear part is a holonomic constraint stating that the CoM does not move in absence of external forces. The angular part is, instead, a nonholonomic constraint indicating that a certain body orientation can be reached as a result of the articulations motion, as empirically observed, e.g., in Fig. 1.

³Notice that if \mathcal{F}_c coincides with \mathcal{F}_h then ${}^c\mathbf{V}_h$ is an identity matrix.

IV. VISUAL SERVOING OF FLOATING HUMANOID ROBOTS

The formal similarity of (3), (5) and (8) allows the natural integration of these components in the same QP (e.g., as done in [16] to integrate VS in a QP-based motion control). In this section, we present a stack for a whole-body motion controller, to achieve visual tasks with a humanoid floating in open space, i.e., in absence of gravity and contacts. Our stack exploits the advantage of having VS and CMC in the same QP. More in details, following the MoT formalism, we design our stack as follows:

$$\mathcal{S}_{\text{VS-CMC}} = \mathcal{T}_{\text{VS}} \ll (\mathcal{C}_{\text{CMC}} + \mathcal{C}_{\text{JOINTS}}) \quad (9)$$

where:

- \mathcal{T}_{VS} is the VS task (presented in Sect. III-B), used to move the robot configuration in order to point the on-board camera towards a chosen visual pattern.
- \mathcal{C}_{CMC} is the CMC constraint as in (8).
- $\mathcal{C}_{\text{JOINTS}}$ limits the robot joint positions and velocities in their respective bounds.

In (9), CMC provides a way to actuate the VS commands, coherently with the absence of contacts. By constraining the rate of change of the centroidal momentum to zero, the FB motion ($\dot{\mathbf{q}}_b$) required to achieve the VS task is obtained as a result of the articulation free motion ($\dot{\mathbf{q}}_a$) in the air. Essentially, the CMC constraint acts as the contacts constraint in a classical whole-body VS scheme with contacts:

$$\mathcal{S}_{\text{VS-CONT}} = \mathcal{T}_{\text{VS}} \ll (\mathcal{C}_{\text{CONT}} + \mathcal{C}_{\text{JOINTS}}) \quad (10)$$

where $\mathcal{C}_{\text{CONT}}$ is the contacts constraint, implemented as $\mathbf{J}_c \dot{\mathbf{q}} = \mathbf{0}$, being \mathbf{J}_c the Jacobian of the contact points. As the $\mathcal{C}_{\text{CONT}}$ constrains the motion of the system to be coherent with the contacts, \mathcal{C}_{CMC} has the same role without the contacts. In fact, without these constraints, the QP would compute physically inconsistent actuation commands for the virtual joints of FB.

The stack, as it is in (9), could generate a behavior of the robot articulations that might end up in inconvenient configurations for further operations with the humanoid. For example, to perform a rotation around the vertical axis, the solution given by (9) would mostly employ the torso vertical joint. This strategy would drive the robot in an inconvenient situation, with arms and legs far apart each other. Instead, it might be more reasonable to wave arms and legs (as in Fig. 1) or, more in general, keep a desired configuration as reference during the motion. With this regard, in order to enforce a more reasonable whole-body motion, one can simply introduce a postural task \mathcal{T}_P in the null-space of (9) to track an *ergonomic* posture during the operations:

$$\mathcal{S}_{\text{VS-P-CMC}} = (\mathcal{T}_{\text{VS}}/\mathcal{T}_P) \ll (\mathcal{C}_{\text{CMC}} + \mathcal{C}_{\text{JOINTS}}), \quad (11)$$

where the ergonomic posture is a design choice, taken according to the particular task to be executed.

The proposed stack contains the essential components to achieve VS tasks in the free space with a humanoid robot, exploiting the CMC constraint. Other components, such as auxiliary constraints or secondary tasks, can be easily added

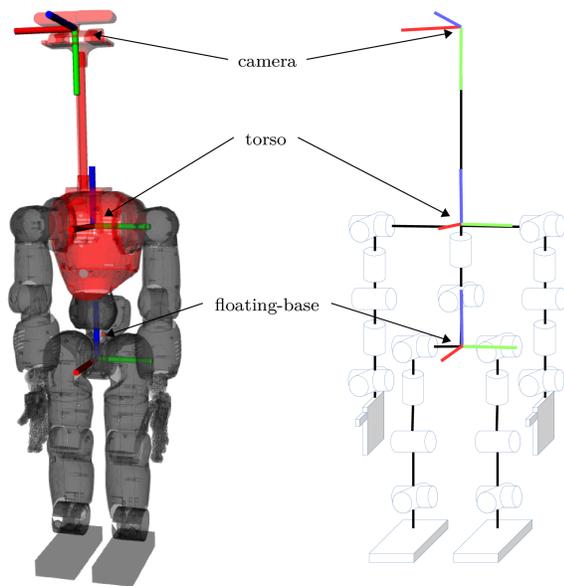


Fig. 2. The humanoid robot COMAN, used in the simulations validating our approach, and the schematic of its kinematics structure.

to the stack, e.g., to avoid undesired motion of the camera. In Sect. V, we show, supported by simulation results, the role played by the main components of our stack.

V. SIMULATION RESULTS AND DISCUSSION

Our approach has been validated with Gazebo simulations of COMAN [25], a humanoid robot with 29 DoFs, equipped with a simulated Xtion sensor (see Fig. 2), used as a monocular camera streaming images of 640×480 pixels at 10 Hz. All the presented simulations reproduced a space scenario (we zeroed the gravity in Gazebo), where COMAN was placed in a space vehicle and had to align with a visual pattern without establishing contacts with the environment. As VS task, we considered the classical problem of tracking four point features. To ease the image processing, we chose four green dots as visual pattern, and used standard tools available in the OpenCV library [26] for their detection. In particular, we applied (i) a color thresholding to isolate green areas, (ii) erosion and dilation operators to remove salt-and-pepper noise, (iii) a blob detector to extract the points coordinates. Figure 3 shows the camera image at the beginning of the simulations, where the green dots and the red circles are the detected and desired visual features, respectively. The task was designed in a way that the robot had to perform a rotation around its vertical axis. The designed task was feasible, in the sense that the visual error could be zeroed at a reachable pose of the robot's on-board camera. The task could not be simply realized by a torso rotation, to which the on-board camera is rigidly connected, due to the limits of the torso yaw joint at ± 0.78 rad, but it requires also the FB motion.

The stack proposed in Sect. IV has been implemented within the OpenSoT [22] library, as part of the CartesI/O [27] framework. The VS task has been integrated leveraging the ViSP software suite [28], while the CMC constraint has been



Fig. 3. Image captured by the robot camera at the beginning of each simulation. The visual task consists in making the green dots match with the red circles, performing a rotation of the robot around the vertical axis.

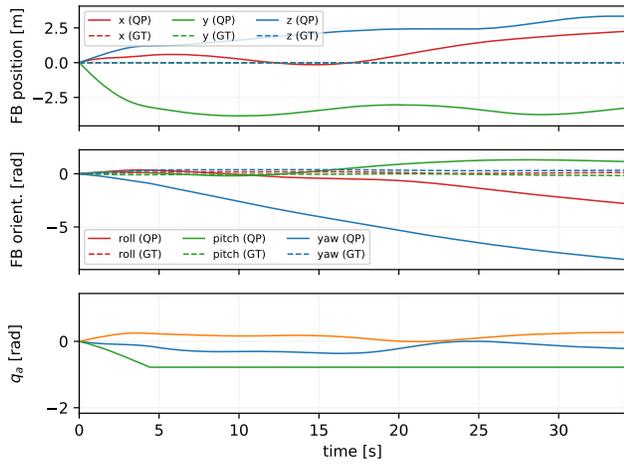
implemented using the RBDL library [29], used also for all the kinematics computation. As QP solver, we selected qpOASES [30], because it manages to respect the constraints with the required tolerance that we demanded. The objective function of the QP was set with a regularization term of 10^{-3} ; the control gain of the VS task was set to 0.001; regarding the postural task, we set the control gain equal to 0.1, the weights related to the virtual joints of the FB to zero, and the one related to the waist yaw joint to 100; all the others weights were set to 1.

In the first simulation, we show the impact of the absence of the CMC constraint to, firstly, achieve the desired task, and secondly, realize a coherent motion with the physics of the scenario. To this end, we show the performance reached by a stack simply composed as

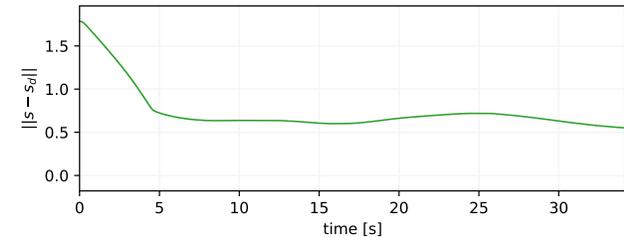
$$\mathcal{S}_{VS} = \mathcal{T}_{VS} \ll \mathcal{C}_{JOINTS} \quad (12)$$

basically obtained from (9) or (10) removing \mathcal{C}_{CMC} or \mathcal{C}_{CONT} , respectively. To realize the desired motion, the QP-based controller fully exploited the range of the torso joints, saturating, in particular, the yaw (see Fig. 4(a), bottom). Furthermore, to rotate the robot towards the desired pose, the QP solver computed direct commands for the FB virtual joints, which obviously could not be executed by the robot, as confirmed by the Ground Truth (GT), i.e. the FB pose given by Gazebo (Fig. 4(a), top and middle). For these reasons, the visual error could not be zeroed, as shown in Fig. 4(b), and the task remained unaccomplished.

In the second simulation, we addressed the same problem by inserting the CMC constraint, i.e., by using the stack \mathcal{S}_{VS-CMC} , as in (9). Thanks to this stack, and the presence of \mathcal{C}_{CMC} , the FB moved consistently with the physics of the simulation (QP commands and GT signals matched, see Fig. 5(a)), as a result of the motion of all the joints. The FB rotation induced by the motion of the limbs, as effect of CMC, allowed to efficiently zero the visual error, shown in Fig. 5(b). Figure 5(c) shows that CM, after an initial transient time, was kept close to zero, respecting (8). Furthermore, as



(a) Configuration of the robot



(b) Norm of the visual error

Fig. 4. First simulation: using a stack composed of only the VS task and joints constraint. In (a) only the articulation joints that move are plotted.

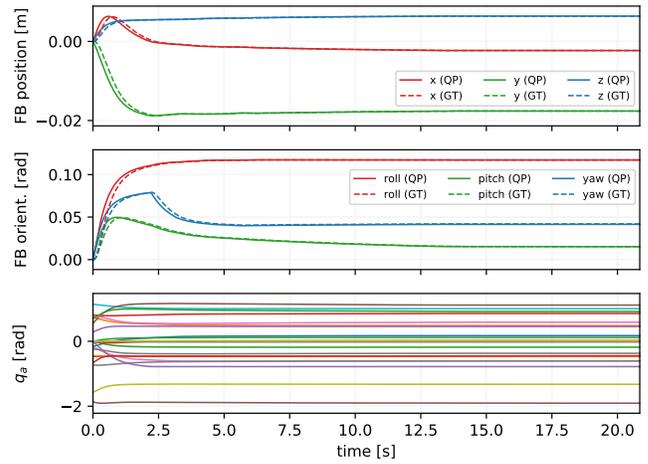
expected, even if little motion of the FB was possible, the CoM could not translate, due to the absence of contacts.

The main drawback of \mathcal{S}_{VS-CMC} is that there is no attention to the posture of the robot while executing the task. This means that the robot could end up in an inconvenient posture at the end of the task or cause self-collisions. In fact, with both \mathcal{S}_{VS} and \mathcal{S}_{VS-CMC} , the torso yaw joint saturated (see Fig. 6), and with \mathcal{S}_{VS-CMC} the robot hands collided with the hip, see Fig. 7 (top). To handle this kind of situations, in the third simulation, we provided the stack with a postural task, i.e. we considered $\mathcal{S}_{VS-P-CMC}$ as in (11), where the QP control enforces a desired posture to the robot motion. As desired posture, we chose the one that the robot had at the beginning of the simulations, visible in the first snapshots of Fig. 7. As a result, the postural task permits to achieve the VS task ending up in a posture where the robot camera, floating-base and feet are all aligned towards the visual pattern (Fig. 7, bottom), without saturating the torso yaw joint (Fig. 6). Furthermore, the robot managed to keep the arms and legs at distance from its body, thus avoiding self-collisions as well.

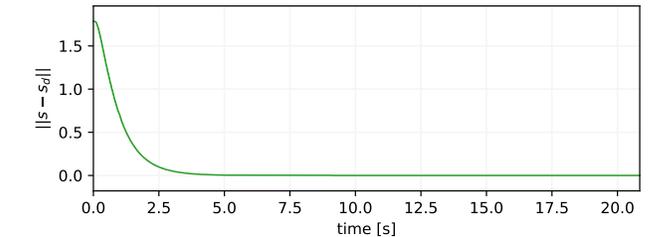
All the video showing the presented simulations, along with the on-board camera views, are included in the material accompanying this paper.

VI. CONCLUSIONS AND FUTURE WORK

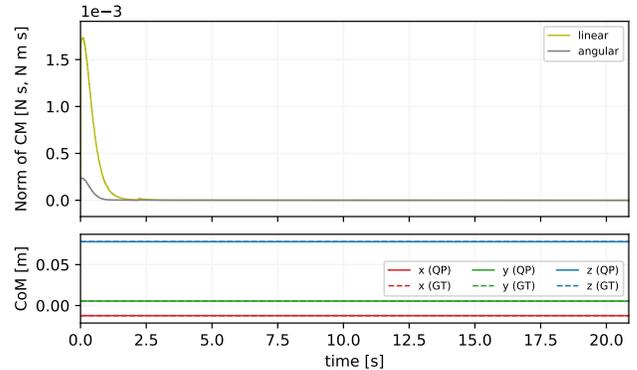
This paper presented a comprehensive formulation of a stack of tasks and constraints for a QP-based controller at velocity level, incorporating VS and CMC for a humanoid robot, in scenarios with absence of gravity and contacts. A



(a) Configuration of the robot



(b) Norm of the visual error



(c) Norm of the CM, and CoM

Fig. 5. Second simulation: using a stack composed of the VS task, CMC and joints constraints. In (a) only the articulation joints that move are plotted.

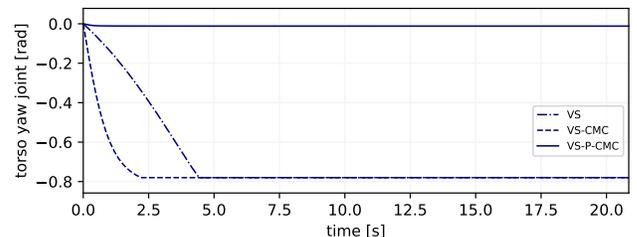


Fig. 6. Evolution of the torso yaw joint during the three presented experiments: \mathcal{S}_{VS} and \mathcal{S}_{VS-CMC} tends to saturate it while the use of the postural in $\mathcal{S}_{VS-P-CMC}$ keeps it almost still, far from the limit.

space simulation scenario, where the robot had to align with a visual pattern, was used to validate the proposed formulation.

We showed that, in some cases, the addition of a postural task can partially avoid self-collisions, requiring however

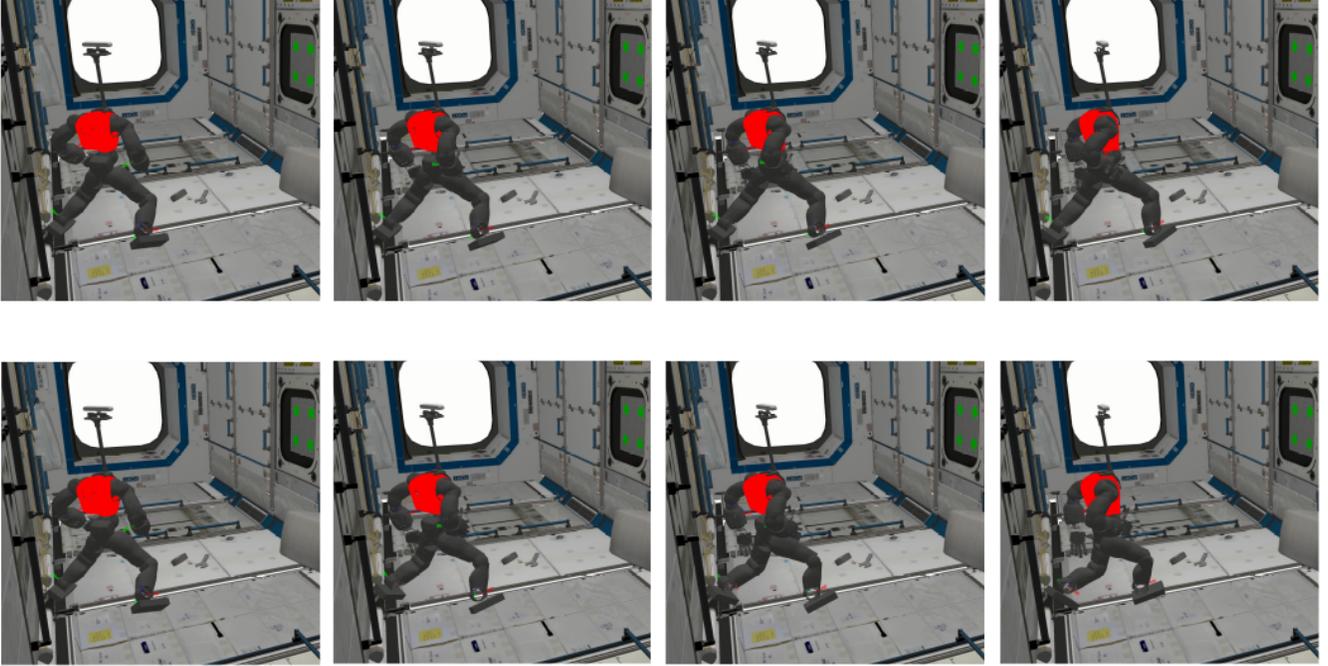


Fig. 7. Snapshots of the simulations using S_{VS-CMC} (top) and $S_{VS-P-CMC}$ (bottom). The postural task allows to accomplish the task without self collisions.

fine tuning of the tracked posture. In future work, we will tackle the possibility to explicitly avoid self-collisions, which often occurs when considering the CMC constraint. In fact, the QP-based formulation permits to easily add self-collision avoidance as constraint or task, as done in [31], which can be integrated and adapted in our scheme.

We limited our validation on a task laying on the transverse plane. To cover other different and wider motions without the risk of meeting singular configurations, the robot should perform complex whole-body cyclic motions that our reactive strategy, in principle, is not able to carry out. However, we observed that under particular working conditions (e.g., when rotating the torso on the transverse plane with a particular reference posture) the presence of the postural task and CMC constraint can surprisingly lead to simple periodic motions, as can be seen in the accompanying video. This behavior is unexpected, due to the instantaneous nature of the used control scheme, which does not take into account any future horizon. This motivates us to further analyse this possibility and investigate other alternative approaches, from optimal control strategies to reinforcement learning methods.

Further developments will be driven by the needs of energy-efficient control strategies of floating robots equipped with jet engines or propellers. In fact, the presented formulation could be used to perform rotation maneuvers saving the precious fuel of the engines increasing the operational time.

VII. ACKNOWLEDGMENT

The authors would like to acknowledge Dr. Nikos G. Tsagarakis and Dr. Alessandro Giusti for supporting the research carried in this manuscript. This work was granted by the European Union’s Horizon 2020 Research and Innovation Programs under Grant No. 779963 (EUROBENCH) and Grant No. 871743 (1-SWARM).

REFERENCES

- [1] K. Tanie and K. Yokoi, “Humanoid and its potential applications,” in *IEEE Int. Conf. on Industrial Technology*, vol. 1, pp. 1–6, 2003.
- [2] A. Kheddar, S. Caron, P. Gergondet, A. Comport, A. Tanguy, C. Ott, B. Henze, G. Mesesan, J. Engelsberger, M. A. Roa, P. Wieber, F. Chaumette, F. Spindler, G. Oriolo, L. Lanari, A. Escande, K. Chap-pellet, F. Kanehiro, and P. Rabat e, “Humanoid robots in aircraft manufacturing: The Airbus use cases,” *IEEE Robot. Autom. Mag.*, vol. 26, no. 4, pp. 30–45, 2019.
- [3] A. Paolillo, P. Gergondet, A. Cherubini, M. Vendittelli, and A. Kheddar, “Autonomous car driving by a humanoid robot,” *J. Field Robot.*, vol. 35, no. 2, pp. 169–186, 2018.
- [4] E. Mingo Hoffman, S. Caron, F. Ferro, L. Sentis, and N. G. Tsagarakis, “Developing humanoid robots for applications in real-world scenarios [from the guest editors],” *IEEE Robot. Autom. Mag.*, vol. 26, no. 4, pp. 17–19, 2019.
- [5] H. Stuart, S. Wang, O. Khatib, and M. R. Cutkosky, “The ocean one hands: An adaptive design for robust marine manipulation,” *Int. J. Robot. Res.*, vol. 36, no. 2, pp. 150–166, 2017.
- [6] D. Pucci, S. Traversaro, and F. Nori, “Momentum control of an underactuated flying humanoid robot,” *IEEE Robot. and Autom. Lett.*, vol. 3, no. 1, pp. 195–202, 2017.

- [7] A. Omer, K. Hashimoto, H.-O. Lim, and A. Takanishi, "Study of bipedal robot walking motion in low gravity: investigation and analysis," *International Journal of Advanced Robotic Systems*, vol. 11, no. 9, p. 139, 2014.
- [8] P.-B. Wieber, "Holonomy and nonholonomy in the dynamics of articulated motion," in *Fast motions in biomechanics and robotics*, pp. 411–425, Springer, 2006.
- [9] F. Chaumette and S. Hutchinson, "Visual servo control. Part I: Basic approaches," *IEEE Robot. Autom. Mag.*, vol. 13, no. 4, pp. 82–90, 2006.
- [10] M. A. Diftler, J. S. Mehling, M. E. Abdallah, N. A. Radford, L. B. Bridgwater, A. M. Sanders, R. S. Askew, D. M. Linn, J. D. Yamokoski, F. A. Permenter, B. K. Hargrave, R. Platt, R. T. Savely, and R. O. Ambrose, "Robonaut 2 - the first humanoid robot in space," in *IEEE Int. Conf. on Robotics and Automation*, pp. 2178–2183, 2011.
- [11] F. Rehnmark, R. O. Ambrose, and M. Goza, "The challenges of extra-vehicular robotic locomotion aboard orbiting spacecraft," in *IEEE Int. Conf. on Robotics and Automation*, pp. 2110–2115, 2004.
- [12] B. Stephens, "Humanoid push recovery," in *IEEE Int. Conf. on Humanoid Robots*, pp. 589–595, 2007.
- [13] P. M. Wensing and D. E. Orin, "Generation of dynamic humanoid behaviors through task-space control with conic optimization," in *IEEE Int. Conf. on Robotics and Automation*, pp. 3103–3109, 2013.
- [14] P. Mohammadi, M. Malekzadeh, J. Kodl, A. Mukovskiy, D. L. Wigand, M. Giese, and J. J. Steil, "Real-time control of whole-body robot motion and trajectory generation for physiotherapeutic juggling in VR," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 270–277, 2018.
- [15] F. Chaumette and S. Hutchinson, "Visual servo control. Part II: Advanced approaches," *IEEE Robot. Autom. Mag.*, vol. 14, no. 1, pp. 109–118, 2007.
- [16] D. J. Agravante, G. Claudio, F. Spindler, and F. Chaumette, "Visual servoing in an optimization framework for the whole-body control of humanoid robots," *IEEE Robot. and Autom. Lett.*, vol. 2, no. 2, pp. 608–615, 2017.
- [17] A. Paolillo, K. Chappellet, A. Bolotnikova, and A. Kheddar, "Inter-linked visual tracking and robotic manipulation of articulated objects," *IEEE Robot. and Autom. Lett.*, vol. 3, no. 4, pp. 2746–2753, 2018.
- [18] K. Giraud-Esclasse, P. Fernbach, G. Buondonno, C. Mastalli, and O. Stasse, "Motion planning with multi-contact and visual servoing on humanoid robots," in *IEEE/SICE International Symposium on System Integration*, pp. 156–163, 2020.
- [19] J. P. Alepuz, M. R. Emami, and J. Pomares, "Direct image-based visual servoing of free-floating space manipulators," *Aerosp. Sci. Technol.*, vol. 55, pp. 1–9, 2016.
- [20] A. Abdul Hafez, P. Mithun, V. Anurag, S. Shah, and K. M. Krishna, "Reactionless visual servoing of a multi-arm space robot combined with other manipulation tasks," *Robot. Auton. Syst.*, vol. 91, pp. 1–10, 2017.
- [21] O. Kanoun, F. Lamiroux, and P.-B. Wieber, "Kinematic control of redundant manipulators: Generalizing the task-priority framework to inequality task," *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 785–792, 2011.
- [22] E. Mingo Hoffman, A. Rocchi, A. Laurenzi, and N. G. Tsagarakis, "Robot Control for Dummies: Insights and Examples using OpenSoT," in *IEEE Int. Conf. on Humanoid Robots*, pp. 736–741, 2017.
- [23] D. E. Orin, A. Goswami, and S.-H. Lee, "Centroidal dynamics of a humanoid robot," *Auton. Robot.*, vol. 35, no. 2-3, pp. 161–176, 2013.
- [24] P. M. Wensing and D. E. Orin, "Improved computation of the humanoid centroidal dynamics and application for whole-body control," *Int. J. Hum. Robot.*, vol. 13, no. 01, p. 1550039, 2016.
- [25] N. G. Tsagarakis, S. Morfey, G. M. Cerda, L. Zhibin, and D. G. Caldwell, "Compliant humanoid coman: Optimal joint stiffness tuning for modal frequency control," in *IEEE Int. Conf. on Robotics and Automation*, pp. 673–678, 2013.
- [26] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.
- [27] A. Laurenzi, E. M. Hoffman, L. Muratore, and N. G. Tsagarakis, "Cartesi/O: A ROS based real-time capable cartesian control framework," in *IEEE Int. Conf. on Robotics and Automation*, pp. 591–596, 2019.
- [28] E. Marchand, F. Spindler, and F. Chaumette, "ViSP for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Robot. Autom. Mag.*, vol. 12, no. 4, pp. 40–52, 2005.
- [29] M. L. Felis, "RBDL: an efficient rigid-body dynamics library using recursive algorithms," *Auton. Robot.*, vol. 41, pp. 495–511, 2017.
- [30] H. J. Ferreau, C. Kirches, A. Potschka, H. G. Bock, and M. Diehl, "qpOASES: a parametric active-set algorithm for quadratic programming," *Math. Program. Comput.*, vol. 6, no. 4, pp. 327–363, 2014.
- [31] C. Fang, A. Rocchi, E. Mingo Hoffman, N. G. Tsagarakis, and D. G. Caldwell, "Efficient self-collision avoidance based on focus of interest for humanoid robots," in *IEEE Int. Conf. on Humanoid Robots*, pp. 1060–1066, 2015.