Intercontinental, Multimodal, Wide-area Tele-Cooperation using a Humanoid Robot

Paul Evrard, Olivier Stasse, Nicolas Mansard, Abderrahmane Kheddar *

Thomas Schauß, Carolina Weber, Angelika Peer, Martin Buss **

Abstract—This paper is the continuation of our previous work in intercontinental, collaborative teleoperation with a humanoid robot. Our new achievement consists in an extension of the former single-arm bilateral teleoperation setting to include bimanual manipulation and walking. A coupling scheme for simultaneous manipulation and locomotion is developed. Furthermore, a task-based control framework, including a force-based control for the arms as well as a walking pattern generation, is presented to realize stable whole-body motions of the highly redundant humanoid robot. Experiments have been performed to assess the proposed control scheme. They bring to light additional scientific challenges that remain in order to reach a smooth and natural telepresent collaboration.

I. INTRODUCTION

Teleoperation consists in remotely controlling a robotic system, the teleoperator, to perform tasks at a distance. In doing so, a multi-modal human-system interface provides sensory feedback to the human operator and allows him/her to interact with the remote environment by mapping her/his actions to the teleoperator.

Teleoperating fixed based robotic arms is a well investigated problem in robotics; even mobile robots have been operated at a distance by means of joysticks [1] or foot pedals [2] or by using a mobile haptic interface [3]. In this paper, we are interested in realizing a teleoperation system that allows for remotely controlling a humanoid which is performing a joint physical collaborative task with another, onsite, human. So far, humanoid robots have been teleoperated using supervisory control to generate whole body motion and bilateral coupling to control the hand and foot positions [4], or using advanced virtual reality technology [5].

As this paper presents the continuation of our previous work [6], we focus hereafter on the problems that are brought to light when the collaborative teleoperation task is conducted while walking, extending the workspace to a much wider range. We consider an operator using a bimanual haptic interface mounted on an omnidirectional, nonholonomic mobile base (MHI) as a master system coupled to a humanoid teleoperator (TOP) as slave system, see Fig. 1.

Various bilateral coupling schemes can be thought of when coupling these two devices. For instance, one can consider the mobile platform at master side and the humanoid at teleoperator side as independent ‘tools’, which interact with the human operator and the manipulated object. To avoid discrepancies between the postures of the operator and the teleoperated robot, which might decrease the feeling of presence and task performance, a posture reflecting scheme can be implemented that maps the tracked posture of the operator and best-copies it on the distant humanoid platform.

In this paper, we describe the lighter implementation where the posture of the operator is not tracked. Our contribution regarding the system described in [6] is the possibility to perform wide area motions by moving the mobile platform and thereby making the humanoid platform walk. The main difficulty we faced is that the types of motion for each system are different: the master platform moves through a continuous contact interaction (wheels) while the humanoid robot’s motion is made through discrete alternate contacts. The impact of this different characteristics of motions is highlighted in the experimental results.

We report hereafter on the goals achieved so far in the realization of a collaborative teleoperation task while walking. We further discuss what remains to be improved to reach a highly sophisticated, smooth, and natural teleoperation.

II. COUPLING SCHEME

When integrating locomotion into the existing teleoperation system, various coupling schemes are possible to
position the mobile platform and the humanoid robot. A possible control design, shown in Fig. 2, is to only reflect the interaction between the TOP and its environment. This can be achieved by classical bilateral coupling between the end-effectors of the TOP and the MHI. The pose of the TOP and of the MHI are controlled independently to allow wide-range motion of the end-effectors. On master site, the MHI is positioned using a local manipulability-optimizing controller, and at teleoperator site the TOP is controlled to generate a stable gait so as to keep the end-effectors at a given position with respect to the body of the robot.

This scheme allows an almost complete separation of the constraints related to each platform (e.g. dynamic balance of the walking humanoid robot), since only the end-effectors are coupled. However, as the postures of the platform are controlled independently, discrepancies between the attitude and positioning of each platform relative to the local user will certainly occur. This is not critical at the master site, since the operator is totally immersed, but it can be problematic at the teleoperator site. If the humanoid posture and the operator posture are dissimilar, the remote operator might misunderstand the intentions of his/her partner. Moreover, it is important for the master operator to have visual feedback that conforms to his/her actual posture in order to take appropriate actions via the teleoperator platform. Moreover, as the locomotion of the distant humanoid robot is decoupled from the locomotion of the operator, the feeling of immersion is impaired. Finally, since the locomotion of the TOP results from the motion of its end-effectors, it is necessary to apply forces to reposition the teleoperator.

A possible solution for these problems, illustrated in Fig. 3, is to add a posture reflecting scheme on top of the task-space coupling scheme. In doing so, the humanoid robot would use its redundancy to best-copy the posture and relative position of the master operator under platform-specific and environmental constraints such as dynamic balance and self collision avoidance. The MHI at the master site is still controlled independently, since the operator does not perceive its posture. Such a scheme would probably result in a better understanding between the partners since information about the operator’s intentions would be conveyed through the humanoid posture. It also allows decoupling of locomotion and manipulation, and would result in a better feeling of presence, especially due to congruence between visual and proprioceptive feedback. This scheme requires the posture of the operator to be tracked and raises challenges, since the gait and balance of the robot must be generated while mimicking the posture of the operator.

For the time being, this last solution has not been implemented. The remainder of this paper will describe a lighter implementation where only the end-effectors are coupled. It requires fewer communication channels, needs no tracking of the operator’s position and incorporates locomotion into the existing system presented in [6] by only modifying the local control loops. An evaluation of this scheme will determine how the feeling of presence is impaired when the posture of the humanoid robot and the operator do not match and whether enhancing this scheme with a posture reflecting coupling is necessary to achieve a desired degree of immersion and task performance.

### III. Local Control at the Remote Side

On the remote side, a HRP-2 humanoid robot with 30 actuated degrees of freedom (DOF) is used to mirror the actions of the master operator. It is capable of bipedal locomotion, and is equipped with force sensors at the wrists and at the ankles. To take advantage of the high redundancy provided by this platform, we have chosen to use a task-based control approach [7], [8]. Such approaches simplify the definition and the implementation of a robotic behavior, especially in the case of highly redundant robots like humanoids. The approach provides whole-body control and can directly comprehend a walking pattern generator.

In the next paragraph, we will shortly recall the overall task-based control architecture. As the control is based on the kinematic inverse, care has to be taken while accounting for forces due to manipulation. A complete solution to obtain a transparent compliance of the arms of the robot is thus proposed in Section III-B. Finally, we recall the algorithm used for robot walking, and present the scheme used to link the whole-body behavior to the step decision (Section III-C).

### A. Stack of Tasks

A task is defined by the triplet \((e, J, \dot{e}^*)\), where \(e\) defines the space of the measure (typically, \(e\) could be an error between a sensor measure \(s\) and its reference value \(s^*\): \(e = s - s^*\)), \(J = \frac{\partial e}{\partial \theta}\) is the Jacobian of the task, and \(\dot{e}^*\) is the vector field defining the desired behavior of the system in the task space (typically, if requiring an exponential decrease of the error, the vector field is defined to be \(\dot{e}^* = -\lambda e\),

![Fig. 2. Coordination scheme considering coupling of end-effectors only](image1)

![Fig. 3. Coordination scheme considering coupling of bodies and end-effectors](image2)
with \( \lambda \) the gain that tunes the robot velocity). Using these
definitions, the model of evolution of the task is given by
\[
\dot{e} = J\dot{q} \tag{1}
\]

where \( \dot{q} \), the joint velocity of the robot, is considered as the
input of the system. To obtain the reference \( \dot{e} \), the generic solution is given by
\[
\dot{q} = J^\#\dot{e}^* + Pz \tag{2}
\]

where \( J^\# \) is a Pseudoinverse of \( J \) \cite{9}, \( z \) is any arbitrarily
chosen vector and \( P \) is the projector onto the null space of \( J \) that guarantees that, for every \( z \), the main task \( \dot{e}^* \) is not modified. The vector \( z \) can then be used to realize
a secondary task, and, by recurrence, any set of tasks hierarchically organized.

The stack of tasks (SoT) is a structure that orders the tasks that are currently active. Only the tasks in the stack are taken into account in the control law. The task at the bottom level has priority over all the others, and the priority decreases as the stack level increases. Any new task added in the stack does not disturb the tasks already in the stack. The control law is computed from the stack, using the redundancy formalism introduced in \cite{10}. Let \((e_1, J_1) \ldots (e_n, J_n)\) be \( n \) tasks. The control law computed from these \( n \) tasks ensures that the task \( e_i \) does not disturb the task \( e_j \) if \( i > j \).

A recursive computation of the joint velocity is proposed in \cite{10}:
\[
\dot{q}_i = \dot{q}_{i-1} + (J_i P^A_{i-1})^\#(e_i - J_i\dot{q}_{i-1}), \quad i = 1..n \tag{3}
\]

where \( P^A_i \) is the projector onto the null-space of the aug-
mented Jacobian \( J^A_i = (J_1, \ldots, J_i) \). \( J_i P^A_{i-1} \) is the limited
Jacobian of the task \( i \) and \( \dot{q}_0 = 0 \). The robot joint velocity
realizing all the tasks in the stack is \( \dot{q} = \dot{q}_n \). A complete
description of the Stack of Tasks is given in \cite{11}. The Stack of Tasks is also extensible for humanoid (free-floating under-
actuated) robots \cite{12}.

B. Force-based control

Using the foregoing control scheme, it is straightforward
to realize many tasks in free space. However, for force-
based control and tasks involving contact, the redundancy
formalism does not directly apply. A solution to this problem is the dynamic inverse as proposed in \cite{8}, \cite{13} (rather
than using the kinematic inverse \cite{2}). However, this solution requires the robot to be torque controlled, which is not
the case for our robot. We rather propose to modify the previous
control scheme in order to obtain a similar behavior.

The space in which the control is designed is the opera-
tional space (6D position) of the contact point denoted by \( r \). A common scheme for position-controlled robots is an admittance controller, that defines the reference through the
differential equation of a virtual mass-damping system:
\[
M\ddot{f} + B\dot{f} = f \tag{4}
\]

where \( r \) is a reference position and orientation in the task
space, \( M \) and \( B \) are arbitrary mass and damping matrices of the
equivalent virtual system, and \( f \) are forces and torques
exerted on the equivalent virtual point. For this setup, \( f \) is
defined as the sum of the real forces measured at the contact
point and the reference forces sent by the master operator.

In \cite{2}, the matrix \( J^\# \) can be chosen as any Pseudo-inverse.
We now choose to apply a weighted Pseudoinverse \cite{14},
defined when \( J \) is full-row-rank by
\[
J^\#W = W^{-1}J^T(JW^{-1}J^T)^{-1} \tag{5}
\]

with \( W \) an arbitrary invertible matrix that represents the
weights on the joints. Taking the derivative of \cite{2} with \( z = 0 \)
and introducing \cite{5} yields:
\[
\ddot{q} = W^{-1}J^T(JW^{-1}J^T)^{-1}M^{-1}(f - B\dot{r}) \tag{6}
\]

Selecting the weight \( W \) as the inertia matrix \( A \) of the robot
and the virtual mass \( M \) as \( A = (JA^{-1}J^T)^{-1} \), the apparent
inertia of the end-effector of the robot, we obtain:
\[
A\ddot{q} = J^Tf - B_\phi \dot{q} \tag{7}
\]

where \( B_\phi = J^TB \) is the friction factor of the whole
body structure. This last equation corresponds to a simplified
version of the dynamic equation of the whole-body compliant
robot with gravity compensation, with forces \( f \) acting on \( r \) and a friction \( B_\phi \) that may be tuned by selecting \( B \) to
stabilize the system.

Consequently, if the control parameters are chosen as
described above, the obtained control is equivalent to the
real dynamics of the robot. The control represents a general-
ization of \cite{2}, which means that both force-based contact
tasks and position-based free space tasks can be realized
within this control structure. No specific values have to be
chosen or tuned, except for the damping gain \( B \), that has
been experimentally verified to be very robust.

This solution generalizes the use of the kinematics
inverse for collaborative task. We now recall quickly how to
incorporate the walking pattern generation into the overall
framework.

C. Pattern Generator

Given the position of the next footprints, the pattern
generator computes the robot’s Center of Mass (CoM) and
foot trajectories that ensure that the robot will not fall. The
stability criterion used for controlling the robot not to fa ll
is dynamically stable.

The pattern generator used in the context of this experi-
ment is based on a simplified model and the preview control
method proposed in \cite{15}. The particularity of the algorithm is two-folded: it uses a preview window to compute the control
law from a simplified Linearized Inverted Pendulum Model
(LIPM), and then compensates for the difference between
the LIPM and the multi-body model of the robot. The core
part of this algorithm is to solve a quadratic programming
(QP) problem, where the cost function is to minimize the
jerk of the LIPM and to follow an ideal ZMP trajectory.
From this ZMP trajectory defined by the foot prints given in
input, the solution of the QP gives a CoM trajectory which is
dynamically stable.
An appealing solution to apply this algorithm for walking gait generation on the humanoid platform would be to recompute the CoM and feet trajectories for the whole preview window at each control cycle. This is currently impossible as the computation cost to solve the QP problem is too high to fit the control loop running on the teleoperator. Hence the pattern generator is used as a planner which generates a stable motion within a fixed time (currently 3.2s). Consequently, it is impossible to input any footprint reference within this time window.

In practice, we used the pattern generator software described in [16]. To account for the preview window, four footprints had to be provided to the algorithm. A new footprint is then added each time one step was executed by the robot. The output of the pattern generator are reference trajectories for the feet and for the CoM, that can be taken into account as top-priority tasks in the SoT.

D. Dynamic balance and collaborative tasks

The challenge when controlling a walking humanoid platform to perform a physical collaboration task lies in the fact that interaction forces induced by the task create sudden perturbations on the ZMP of the robot. In order to keep dynamic balance, the footprints would have to be modified within the preview window. The technical difficulty to cope with this, is hence to be able to compute a new reference for the CoM and the feet trajectories with modified inputs in the preview window, and/or with a new initial state.

In this work, we have experimentally evaluated whether the humanoid platform could keep its balance while performing a collaborative task. It appears that the commercial stabilizer provided with the control software and the admittance control implemented for the end-effectors limit to some extent the interaction forces between the teleoperator and the onsite operator. Thus, perturbations on the ZMP of the robot are small enough for the stability criterion to be satisfied without changing the footprints in the preview window.

To decide on the next step to be provided to the pattern generator, we considered the distance of the current position to a reference position of the hands given in the waist coordinate system. Steps are planned to regulate this error to a reference transmitted from the master.

E. Definition of the tasks

To summarize, seven tasks have been considered in the experiments.

1) e_walk tracks the desired trajectories of the legs provided by the pattern generator. Since this task controls the absolute position and orientation of both feet of the robot it uses 12 DOF of the robot. In future work, only the position of the free flying foot will be controlled, and the support foot will be considered as fully constrained by the contact with the ground. This will relax 6 DOF of the robot, thus allowing a wider range of motions.

2) e_com regulates the position of the CoM on the trajectory given by the pattern generator.

3+4) e_xr and e_xl regulates the altitude of both right and left hands of the robot to a desired fixed height, in order to simplify the task of the master operator when the robot is walking. Position of the end-effectors r_xr and r_xl is controlled by taking into account only the z component of the vector. These tasks are removed when the robot is not walking.

5) e_fr and e_fl are the force-based control of both right and left hands of the robot. They are defined as in [4]. The reference force f is then the sum of real measured forces and reference forces sent by the master.

6) e_head controls the orientation of the head according to the reference position given by the master operator.

7) e_grip is finally added to control specifically the gripper aperture to the reference transmitted from the master.

IV. LOCAL CONTROL AT THE MASTER SIDE

As haptic interfaces two ViSHaRD7 are used on operator site [17]. The two 7-DOF arms have a human-sized workspace and high force output capabilities. The haptic interfaces can hereby perform movements in a workspace similar to one of the human operator at a fast speed. The interfaces were especially designed to be mounted on an omnidirectional, nonholonomic mobile base [18]. The mobile base can perform arbitrarily large movements at a slower speed. The combination is capable of displaying fast movements and high forces in an arbitrarily large workspace.

The haptic interfaces are controlled using position-based admittance controllers. A virtual admittance is used at the local site to render a compliance with inertia M, stiffness K, and damping B. Thereby \( \dot{x}_\Delta \) is the offset between the desired end-effector position \( x_d \) of one arm and the end-effector position \( x_e \) received from the slave. Consequently, the input to the inner position control loop is determined as

\[
\dot{x}_d = x_\Delta + (x_S - x_B),
\]

where \( x_B \) is the position of the mobile platform relative to its starting position.

The base is repositioned so as to always maximize the minimum manipulability of the two end-effectors. For the given kinematic structure ideal manipulability is given if both end-effectors are located on circles with fixed radius \( r_{opt} \) around the shoulder joints. In addition, the first rotational joint of each interface should be controlled in such a way, that the minimum distance to their joint limits is maximized, i.e. the two joint angles must be equal.

The resulting configuration is symmetric and the corresponding platform position can be obtained by simple geometric calculations: the mobile platform must be aligned parallel to the connecting line from \( x_f \) to \( x_R \) and its center point must lie on the perpendicular bisector of the connecting line (see Fig. 2). The distance \( n_{opt} \) can be computed using

\[
n_{opt} = \sqrt{r_{opt}^2 - \left( \frac{||d|| - d_P}{2} \right)^2} + n_P
\]
Fig. 4. Geometric solution for optimal positioning of the mobile platform

where \( d = x_R - x_L \) is the vector connecting the two end-effectors in world coordinates, \( r_{\text{opt}} \) is the radius of the circle with maximum manipulability, \( d_P \) is the distance of the first rotational joints of the haptic interfaces, and \( n_P \) is the distance of the platform center from the connecting line between the first rotational joints of the haptic interfaces.

A PD controller is then used to drive the mobile base to the desired position. As the current base position is taken into account in the virtual admittance the positions received from remote site are displayed in world-coordinates on the local site.

Here only a short overview of the control scheme has been given. For details on the control of the haptic interfaces, see [17]. The mobile haptic interface is presented in [18].

V. Experiments

A. Experimental setup

The setup used for the experiments is depicted in Fig 5 for a detailed description please refer to [6]. The human-system interface, located in Munich, Germany, consists of the mobile haptic interface, described in Sec. IV to allow haptic interaction and a head-mounted display to provide visual feedback to the operator. As teleoperator, located in Tsukuba, Japan, the humanoid robot HRP-2, as described in Sec. III is used. The two narrow-angle cameras of HRP-2 record the remote scene. For control, forces and torques from both arms are sent from master to slave, while end-effector positions and orientation of both arms are sent from slave to master. At a transmission rate of 50 Hz, the packet loss was negligible (< 1%) and the round-trip time between Germany and Japan was determined to be \( T = 276 \pm 5 \) ms.

B. Results and discussion

Experiments in which both, the human operator and the teleoperator performed some steps have been successfully performed. Fig. 6 depicts the results for an experiment in which the human operator in Munich applies most of the force necessary to initiate the movement and the human collaborator in Tsukuba loosely holds the end-effectors of the teleoperator. A very good position tracking between the respective end-effectors is achieved (solid lines), also while walking. The slight deviations that can be seen are caused by time-delay in the communication channel and a desired compliant behavior of the haptic interfaces on master side.

However, a large deviation of the body positions between the human operator and the teleoperator are observed (dashed lines). E.g. after 63 s the teleoperator starts moving backwards although the human operator is still moving forwards. Also, some low frequency oscillations of the teleoperator position that are caused by the current implementation of the walking pattern generator are evident. The human operator observes this as disturbing. Thus, a high degree of immersion and good task performance while walking were not achieved using this coupling scheme.

Fig. 7 depicts the results for the reverse experiment, i.e. the human operator in Munich loosely holds the end-effectors and is guided by the human collaborator in Tsukuba. Position tracking of the end-effectors is equally good. Again, a deviation of the body positions is apparent. However, as the human collaborator on site notices oscillations caused by the walking pattern generator more quickly and reacts immediately, no low frequency oscillations of the teleoperator’s body position can be observed.

Summarizing the experimental results it can be concluded that the goal of performing a teleoperation experiment, in which walking and haptic interaction occur simultaneously, has been achieved. Further improvements to this concept are however necessary to achieve a high degree of immersion and a good task performance.
Fig. 7. Position tracking of body and end-effectors with a dominant human collaborator at remote site. Data of the human operator at master site is depicted in black, data of the teleoperator in gray. The dashed lines represent the body position whereas the solid lines represent the right hand/end-effector.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented the ongoing work in realizing a collaborative teleoperation tasks. We described how the local and remote platforms have been coupled, and what are potential enhancements of this coupling scheme. The local controllers on each site have been presented, with focus on the humanoid robot control.

The experimental results show that the limitations of the walking gait generation of the teleoperated humanoid platform is the cause of big deviations between the postures of the operator and teleoperator. This might lead to a degradation of the immersion and task performance. However, augmenting the current scheme by reflecting the posture of the operator to the humanoid platform requires generating foot steps at a very high rate.

Hence, our future work will consist in improving the walking pattern generator and in the implementation of algorithms for the prediction of human operator motions. For the former problem, methods have been already proposed to deal with perturbations on the ZMP [19] or to reduce the computation time [20] so as to allow the recomputation in real-time of the CoM and ZMP trajectories when changing the foot print in the preview window. However, in [20] a range of situations are given where the ZMP is going outside the support polygon. We recently [21] proposed a new solver which would solve the QP program which is the core of the preview control method under the constraints of the support polygon. This new solver allows to compute a new command in 2 ms on the HRP-2 robot, and should be able to ensure that the solution is inside the support polygon. Its inclusion into the overall system is still a work in progress.

Another challenge will be to reflect the posture of the operator on the humanoid platform. As the teleoperator has not the same kinematics as the operator, and the body control of the teleoperator is subject to very hard constraints, such as keeping balance in presence of disturbing forces, care will have to be taken to find a compromise between posture reflection and constraint violations.

REFERENCES


