

Using vision and haptic sensing for human-humanoid haptic joint actions

Don Joven Agravante, Andrea Cherubini, Abderrahmane Kheddar

► **To cite this version:**

Don Joven Agravante, Andrea Cherubini, Abderrahmane Kheddar. Using vision and haptic sensing for human-humanoid haptic joint actions. CISRAM: Cybernetics and Intelligent Systems Robotics, Automation and Mechatronics, Nov 2013, Manila, Philippines. 2013, <<http://www.cis-ram.org/2013/>>. <lirmm-00908439>

HAL Id: lirmm-00908439

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00908439>

Submitted on 22 Nov 2013

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.

Using vision and haptic sensing for human-humanoid joint actions

Don Joven Agravante¹, Andrea Cherubini¹ and Abderrahmane Kheddar^{1,2}

Abstract—Human-humanoid haptic joint actions are collaborative tasks requiring a sustained haptic interaction between both parties. As such, most research in this field has concentrated on how to use solely the robot’s haptic sensing to extract the human partners’ intentions. With this information, interaction controllers are designed. In this paper, the addition of visual sensing is investigated and a suitable framework is developed to accomplish this. This is then tested on examples of haptic joint actions namely collaboratively carrying a table. Additionally a visual task is implemented on top of this. In one case, the aim is to keep the table level taking into account gravity. In another case, a freely moving ball is balanced to keep it from falling off the table. The results of the experiments show that the framework is able to utilize both information sources properly to accomplish the task.

Index Terms—Physical Human-Robot Interaction, Human and humanoid skills/cognition/interaction, Human-Robot Collaboration

I. INTRODUCTION

Humanoid robots provide many advantages when working together with humans to perform various tasks. This is because humans have an extensive experience in physically collaborating with each other. Humanoids take advantage of this with their human-like range of motion and sensing capabilities. This reduces a human collaborator’s need to learn how to interact with the robot. However, many challenges are still present in the various research areas that study physical human-robot collaboration. The particular area of interest in this paper is the integration of vision and force information to enable “human-robot haptic joint actions”. These are collaborative tasks requiring both parties to physically interact with each other. It has two important aspects:

- 1) both robot and human are doing jointly the same task,
- 2) a haptic interaction exists.

An example of such a task is joint carrying/transportation of large objects [1]–[4]. This is illustrated in Fig. 1 along with three important information sources: prior task knowledge, vision and force. In this example, the haptic interaction exists through the object - the haptic channel. This means that a force/torque applied on one end of the object is felt by the partner on the other end. Because of this, previous research has focused primarily on regulating interaction forces for safety. The use of vision for the robot has not

been investigated before in this context and this is the main focus of this paper. Finally, the prior task knowledge can be used as a guideline on how vision and/or force information should be used for the task.

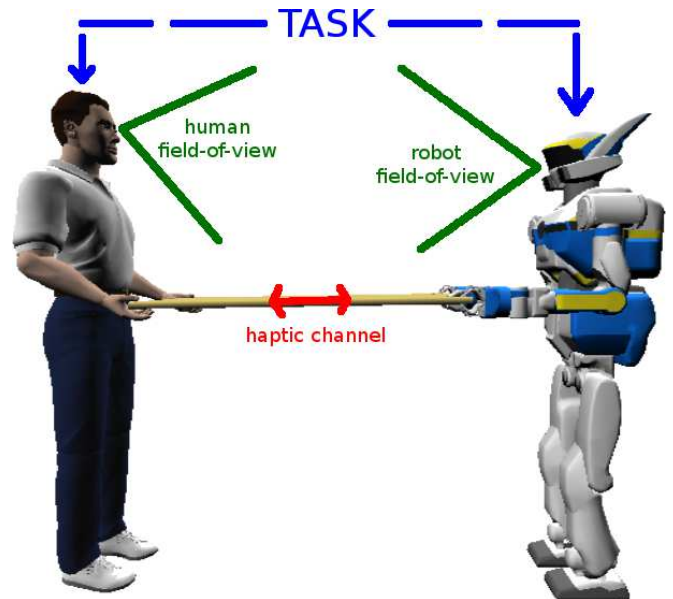


Fig. 1. A diagram of human-robot haptic joint actions. In this general case, both human and robot have a complete knowledge of the task (represented by the blue arrows). Furthermore, each uses both vision (green) and haptic (red) information to achieve this task.

Aside from just being able to do the task, the desire is to make the robot proactive in helping the human [4]–[6]. A proactive behavior aims to lighten the load on the human, as opposed to being a passive follower –only reacting to human motion without any anticipation. A variety of ways are being researched to program proactive behaviors in joint actions. For example in [5] a minimum jerk model of human motion is used to have a good guess of the human’s intention. In [6], machine learning methods are used to obtain a proactive behavior. Another method is studying humans doing the task. An example of this is presented in [4], where it was found that a human dyad moves in constant velocity phases when doing joint transportation of large objects (as in Fig. 1). These findings were then applied to a controller to make the humanoid robot more proactive [4].

As previously mentioned, most works in this field use data from force/torque sensors only to do the task and make the robot pro-active. Although it is impressive what can be achieved using force data alone (i.e. a “blind” robot), vision is clearly required for some tasks and could possibly help

¹CNRS-UM2 LIRMM UMR 5506, Interactive Digital Human group, 161 Rue Ada, 34392 Montpellier, France {firstnames.lastname}@lirmm.fr.

²CNRS-AIST, JRL (Joint Robotics Laboratory), UMI 3218/CRT, Intelligent Systems Research Institute, AIST Central 2, Umezono 1-1-1, Tsukuba, 305-8568, Japan.

to make the robot proactive. For example in collaboratively carrying a table, force data alone cannot be used to see if an object on top of the table is about to fall down. It is clear that visual information is largely complementary to force information (analogous to the human senses of sight and touch). Combining these might enable a humanoid to perform more complicated tasks, similar to a human [7]. Although the benefits are great, there are not many established methods integrating vision and force control.

To start detailing this paper’s contribution, the concept of human-robot haptic joint actions is first elaborated in Section II. Following this, a short review on the state of the art in combining vision and force is presented in Section III. The proposed framework to be used is then detailed in Section IV. This is then applied to two case studies in Section V. The results of these are shown in Section VI. Finally Section VII ends with a brief summary and plans for future work.

II. HUMAN-ROBOT HAPTIC JOINT ACTIONS

Haptic joint actions are mainly characterized by a collaborative task and a haptic interaction throughout this task. This applies to several collaborative tasks that humans do together. For example: helping someone to their feet, the process of handing off objects, supporting someone while walking, etc. Another common example is carrying a large object together. This is illustrated in Fig. 1. This work investigates the addition of visual information to such tasks (shown in Fig. 1 by the field-of-view). To better show our contribution with respect to previous works, a few examples of these are shown in Fig. 2.

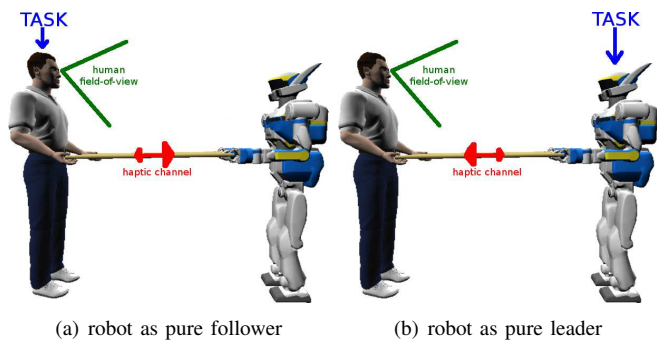


Fig. 2. Particular cases of Human-Robot Haptic Joint Actions

The most commonly studied scenario of human-robot collaboration is making the robot a pure follower as in Fig. 2(a). This illustrates that the task is only known to the human leader a priori. Through the haptic channel, the human communicates his intentions to the robot. With a properly designed controller that takes into account interaction forces, the robot can follow the human’s lead [4]. This is represented by the bigger arrow in the haptic channel from human→robot. Furthermore, the human is able to obtain visual and haptic data as to how good/bad the robot follows the task.

Another situation is the exact opposite of the robot being a pure follower: it can be the leader of the task. This is illustrated in Fig. 2(b). An example of such a scenario is presented in [3] where a joystick is used to give the robot direct information on the task. Although a 2nd human provides this task information via the joystick, in the context of collaborative carrying, the robot is the leader. Without knowing the task of the robot, the human partner then tries to help the robot carry the object to a desired location. It is fair to assume that the human uses both the sense of sight and touch to achieve this task. Apart from the two illustrated examples, other possibilities exist. For example, a combined leader-follower strategy has also been developed [2], [8]. The concept for this strategy is that a sharing of roles can be possible where one is both leader and follower (with varying degrees).

Finally, it can be noticed that in these previous works the robot is solely reliant on haptic information. No visual data is used in the task as illustrated in Fig. 2 by the lack of the robot field-of-view. The aim of the work here is to move towards the general case of Fig. 1, particularly adding vision as another information channel. In order to do this, the main question is how does one combine vision and force information in the context of haptic joint actions. The context presents significant challenges specifically by having a human in the loop.

III. STATE OF THE ART ON COMBINING VISION AND FORCE FOR MANIPULATION

To our knowledge, there has been no previous work on combining vision and force in the context of human-robot haptic joint actions. However, previous works on combining vision and force for object manipulation tasks can be found in the literature. A brief review of these methods is given here as context for the choice of the control framework detailed in the next section.

The different approaches to combining vision and force information for manipulation tasks can be classified into three general categories [9]. These are: traded, hybrid and shared. The simplest strategy is termed as “traded” since the final control output is traded between a purely force-based controller and a purely vision-based controller. The switching between controllers depends on the task execution. A common example of this strategy starts by doing a “guarded move” visual servoing phase. A contact event causes the guarded move phase to stop. After this, the task is completed by force control.

The second category is termed as “hybrid” methods. These are hybrid in the sense that the vision-based controller and force-based controller act at the same time but in different spaces. This requires prior specification of a “task-frame” [10], [11]. The task frame is a Cartesian reference frame that can be divided into vision-controlled and force-controlled directions. Doing this decouples vision and force into orthogonal spaces. After this, the controllers can be designed separately and work independently in their own predefined space.

The third and final category is termed as “shared” methods. These are so-called since there is no separation in time (the case of traded control) or in space (the case of hybrid control). The control responsibility is shared by both vision and force throughout the operation. By doing this, all available information can be used [9]. An example of a shared method is described in [12], where force feedback is used to correct the visual servo control trajectory.

In the context of human-robot haptic joint actions, a shared method is needed. With the human in the loop, safety is always a top priority. Therefore, the control must always be compliant in all degrees of freedom (DOFs). This can be achieved by always making use of the force information in all DOFs. Therefore, in order to make use of visual data, a shared control strategy must be used. A good candidate for this is the impedance control framework [13]. Impedance control allows a manipulator to be compliant by defining a virtual mechanical impedance. Vision can be used in this framework to provide a reference trajectory that is tracked in the absence of external forces [14], [15]. This approach fits the criteria that the robot is always compliant even when moving according to the visual information.

IV. OVERVIEW OF THE ROBOT CONTROL FRAMEWORK

As mentioned in the preceding section, a vision-based controller coupled to an impedance controller is the core concept that is used in this work. However, there are a few more modules that need to be defined to make everything work on a humanoid robot. These are briefly detailed in this section for the sake of completion. As a guide, Fig. 3 shows the overall workflow in a scenario where the goal of the human and robot is to transport the plate while preventing the ball from falling off. This same scenario was tested for the framework here and is discussed more in the next sections as one of the case studies.

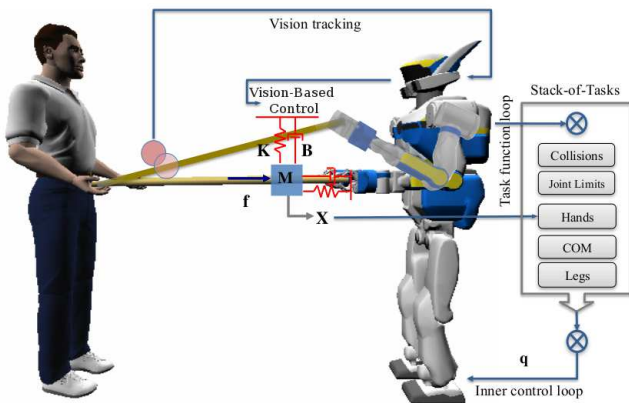


Fig. 3. An illustration of the overall workflow of the algorithms running in the robot to enable it to collaborate with the human using both vision and force information.

A. Visual Tracking Algorithm

To start the whole control framework, visual information needs to be processed. In the example depicted in Fig. 3, this

algorithm gives data about the ball on top of the table. The raw data is a combination of an RGB image and a depth map, obtained from an ASUS Xtion sensor mounted as the “eyes” of the humanoid. This raw data needs to be processed into an estimate of the pose or position to be controlled by the vision-based controller. This is done by tracking a salient visual feature throughout the image sequence and extracting the needed pose or position information from this. However, there is no generic visual tracking algorithm that can work for any and all cases. Furthermore, some knowledge of the task is needed to know what is the important visual information that is needed. This is the reason why the case studies presented here use trivial objects - a cube of known size and color and a ball of known color. However visual tracking is a fairly well-developed field in computer vision and a wide range of algorithms have been developed as reported in this extensive survey [16]. A future plan is to improve in this area by making use of more state-of-the-art algorithms but this is not the focus of this paper. But it should be noted that the main constraints for the vision algorithm are robustness and speed since it needs to be used with a controller. Therefore, visual trackers such as [17] that are designed with this in mind are preferred.

B. Vision-Based Control

After the pertinent information has been extracted from the images, this data needs to be used in a controller that does the desired task. Visual servoing is a well-studied research field that uses the visual information directly into a control law [18]. However, as with the visual tracking algorithm, a visual servoing controller needs to be specifically designed depending on the task. Hence a prior knowledge of the task is also needed. In the two case studies to be shown, the tasks are different and as such different vision-based controllers are created. The details of these are presented in the next section.

C. Impedance Control

Impedance control is a general framework that regulates the contact interaction of the robot and its environment according to a defined virtual mechanical impedance [13]. In the context of human-robot collaboration, this means that interaction forces between the human and robot are regulated at all times. A general formulation of impedance control is:

$$\mathbf{f} = \mathbf{M}(\ddot{\mathbf{X}}_d - \ddot{\mathbf{X}}) + \mathbf{B}(\dot{\mathbf{X}}_d - \dot{\mathbf{X}}) + \mathbf{K}(\mathbf{X}_d - \mathbf{X}) \quad (1)$$

where \mathbf{f} is the wrench (force-torque screw) composed of the force and torque vectors. The vectors \mathbf{X}_d , $\dot{\mathbf{X}}_d$ and $\ddot{\mathbf{X}}_d$ are a desired pose and its first and second derivative. Correspondingly, vectors \mathbf{X} , $\dot{\mathbf{X}}$ and $\ddot{\mathbf{X}}$ represent an actual pose and its first and second derivative. Finally, matrices \mathbf{M} , \mathbf{B} and \mathbf{K} are the inertia, damping and stiffness parameters that define the desired virtual mass-spring-damper system [13].

The general formulation in Eq. (1) can account for “different” control methods. Firstly, it can either be “impedance controlled” by controlling forces/torques (\mathbf{f}) or “admittance controlled” by controlling pose (\mathbf{X}). We use the latter,

since the force-torque sensors on the HRP-2 wrists provide the signal to be used for \mathbf{f} and it is inherently position-controlled. Secondly, a variety of specific methods can be defined, depending on the values of \mathbf{M} , \mathbf{B} , and \mathbf{K} . Here, all three matrices \mathbf{M} , \mathbf{B} , \mathbf{K} are utilized. These are tuned from experiments to achieve the desired compliant behavior.

To completely define the impedance controller, the desired pose and its derivatives \mathbf{X}_d , $\dot{\mathbf{X}}_d$ and $\ddot{\mathbf{X}}_d$ are needed. These are just the output of the vision-based controller as shown in Fig. 3. Therefore in the absence of external forces, the vision-based controller achieves its control objectives.

D. Stack-of-Tasks

The Stack-of-Tasks is a generalized inverse kinematics abstraction layer [19]. As its name implies, the main advantage it gives is the hierarchical organization of different tasks to be executed. This allows efficient integration and abstraction of the different humanoid robot tasks. In the context of collaborative object carrying depicted in Fig. 3, the vision-based control and impedance control are used to servo the humanoid hands. While this is done, other tasks are also in the stack, namely: collision and joint limit avoidance, a task to maintain a good posture and control the robot's center of mass (COM) and finally a task to control the legs for walking [20]. The Stack-of-Tasks framework is able to directly control all the humanoid robot's joints considering all the tasks and their priority. Since critical tasks have a higher priority, their execution is ensured [19].

V. CASE STUDIES

To test the framework described in Section IV, two simple cases of human-robot haptic joint actions are created that clearly benefit from visual information. Because the task of collaborative "table-carrying" has been well-studied within the group, this is used as the base task. An object is then placed on top of the table and the additional task is concerned with this object and the table tilt angles (ϕ_x, ϕ_y). A simplified side-view of the task in Fig. 4 shows ϕ_y and its relation to the height difference z_r and the table length l_t . Furthermore, three important reference frames are drawn in this image to facilitate the explanations that follow - the control frame $\{cf\}$, a local reference frame on the robot $\{l\}$ and the table frame $\{t\}$. The control for the robot can be done just by defining the pose ${}^l\mathbf{T}_{cf}$. This is justified by assuming a rigid grasp during the whole task. This means that the pose of the right and left hands: ${}^{cf}\mathbf{T}_{rh}$ and ${}^{cf}\mathbf{T}_{lh}$ are constant throughout the task and generating the 2-handed control is just a change of frame.

A. Stationary Object - keeping the plane level

In this scenario, a green cube is placed on top of the table as a representative "stationary object". To visually track this object and get its pose, the model based tracker of the ViSP software package [17] is used. The additional goal here is to keep this object upright with respect to the gravity field. This implies that the table plane must be kept level. More formally, the objective can be thought of as $\phi_x = 0, \phi_y = 0$.

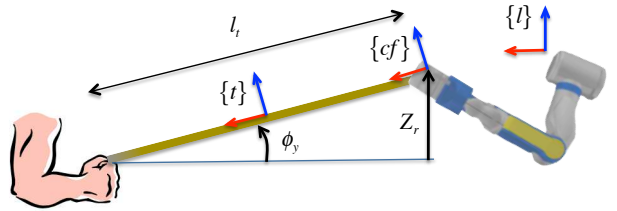


Fig. 4. A simplified "thin beam" model used to control the table tilt through the height

The case of ϕ_x is trivial in this regard (for as long as no conflicting haptic intention is applied by the human). The case of ϕ_y is more interesting as it depends directly on the human as seen in Fig. 4. From the figure, it can be seen that there are 2 ways to affect ϕ_y . First is applying a rotation and hence a torque on the controlled end. Since the table is rigid, the human needs to react to this appropriately by changing his z . This can be uncomfortable for the human, hence the other option is used here which is the exact opposite - the robot changes z_r . It should also be noted that a rotational compliance in both partners is needed for such a control strategy. With this, the model of the task is then:

$$l_t \sin \phi_y = z_r. \quad (2)$$

The final control law is obtained by using the method to derive visual servoing controllers described in [18]. By differentiating the model and setting an exponential decrease of the error $\dot{e} = \dot{\phi}_y = -\lambda \phi_y$ the control law of Eq. (3) is obtained where λ is a gain parameter.

$$\dot{z}_r = -\lambda \hat{\phi}_y \cos \hat{\phi}_y \quad (3)$$

To use this in the impedance control framework, numerical integration is done such that $z_r(t) = z_r(t - dt) + \dot{z}_r(t)dt$ where dt is the time step of the controller. A piece-wise constant velocity is assumed making the acceleration null.

B. Moving Object - keeping a ball from falling off

In this scenario, a ball is placed on top of the table. Any disturbance will tend to make the ball fall off the table. In this case, a specific visual tracking method was developed for the ball using well-established computer vision algorithms. Using both rgb and depth information, the ball's position relative to a pre-defined goal location on the table is tracked visually. With this information, the controller can then be defined.

The scenario described here is the classical "ball-on-plate" example problem in control theory. This is well-known to be marginally stable with just a proportional control law but it is also known that a proportional-derivative (PD) control can be enough which is what is tested in the work presented here. Furthermore, the same argument about controlling ϕ_y from the stationary object case can be made here. Hence the control is done in z_r and ϕ_x . With these considerations, the control laws are shown as follows:

$$\begin{cases} z_r = K_{p,x} (x_d - x) - K_{d,x} \dot{x} \\ \phi_x = K_{p,y} (y_d - y) - K_{d,y} \dot{y}, \end{cases} \quad (4)$$

In Eq. (4), (x, y) pertain to the ball position, (x_d, y_d) its desired location and (\dot{x}, \dot{y}) the error derivatives considering a constant desired location. K represent the control gains which are experimentally tuned.

To use this result in the impedance control framework, numerical differentiation is done to obtain the velocity terms. However, the acceleration is again considered null to prevent noise from differentiation.

VI. RESULTS

The case studies of Section V were tested experimentally on the HRP-2 humanoid robot. These were also integrated with previous work in the group [3], [4] that does the collaborative transportation task. Snapshots of the experiments for the stationary object case is shown in Fig. 5 and the experiments for the moving ball are shown in Fig. 6. In these experiments, the walking gait of the humanoid produces an external disturbance. But even with this, the designed task is achieved showing that the approach presented works.

A. Stationary Object - keeping the plane level

To verify the vision controller design, a step response of the error (ϕ_y) was obtained in experiments. This is shown in Fig. 7. The top plot shows that the error decreases exponentially to 0. This is the expected behavior from setting ($\dot{e} = -\lambda e$). This result proves that the implemented system works (although with some noise and noticeable latency of the vision algorithm).

B. Moving Object - keeping a ball from falling off

To show the performance for this case, the resulting ball trajectory while walking is plotted in Fig. 8. This also corresponds to the snapshots of Fig. 6 (bottom). This data is from the visual tracker used and as such is an estimate of the actual result. However, it should be noted that in this experiment, the ball did not fall while walking with the human even when changing directions (which is the difficulty for the faster ball). The results here are from the slower yellow ball, and the desired ball position is set to $(0.15, 0)$, which is 15cm closer to the human than the table centroid. The red border signifies a rough approximate of the table boundaries. This result show that although the ball moves a lot, it doesn't fall off the table during the transportation task.

VII. CONCLUSION AND FUTURE WORK

In this paper, a framework to combine vision and haptic information in human-robot haptic joint actions is presented. The core idea of the framework is using vision-based controllers to define a desired virtual trajectory for the impedance controller which regulates interaction forces. Two case studies are presented here with different vision-based controllers and both are shown to be able to achieve the task. These experiments verify the chosen approach. This is also just preliminary steps for the planned integration of visual information into collaborative tasks. A next area of research is in increasing the proactive behavior of the system by detecting visual cues from the partner.

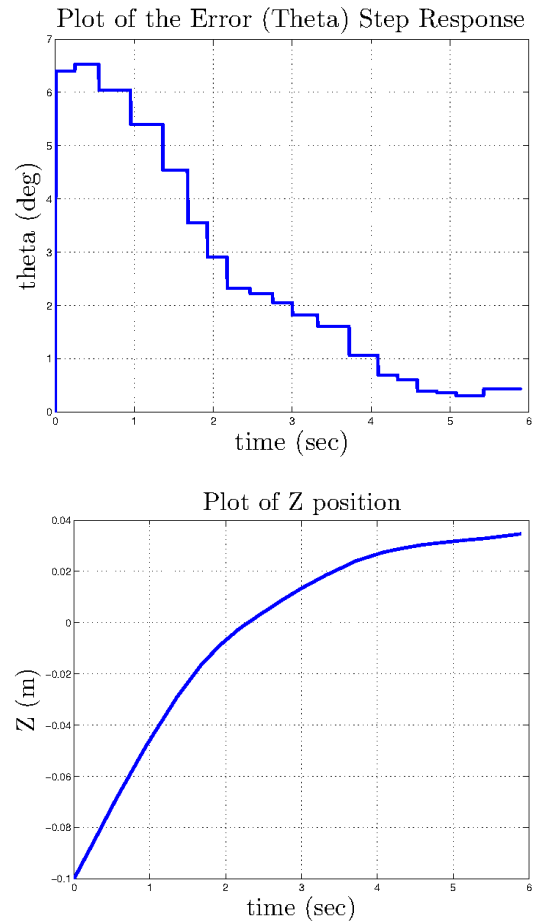


Fig. 7. Step response of the vision controller. Top: Plot of the error (ϕ_y). Bottom: Plot of $l_{z_{cf}}$

VIII. ACKNOWLEDGEMENT

This work is supported in part by the FP7 IP RoboHow.Cog project (www.robohow.eu). FP7-ICT-2011-7 Contract No 288533.

REFERENCES

- [1] K. Yokoyama, H. Handa, T. Isozumi, Y. Fukase, K. Kaneko, F. Kanehiro, Y. Kawai, F. Tomita, and H. Hirukawa, "Cooperative works by a human and a humanoid robot," in *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*, vol. 3, pp. 2985–2991, IEEE, 2003.
- [2] P. Evrard and A. Kheddar, "Homotopy switching model for dyad haptic interaction in physical collaborative tasks," in *EuroHaptics conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*, pp. 45–50, IEEE, 2009.
- [3] A. Bussy, P. Gergondet, A. Kheddar, F. Keith, and A. Crosnier, "Proactive behavior of a humanoid robot in a haptic transportation task with a human partner," in *RO-MAN, 2012 IEEE*, pp. 962–967, IEEE, 2012.
- [4] A. Bussy, A. Kheddar, A. Crosnier, and F. Keith, "Human-humanoid haptic joint object transportation case study," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 3633–3638, IEEE, 2012.
- [5] Y. Maeda, T. Hara, and T. Arai, "Human-robot cooperative manipulation with motion estimation," in *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on*, vol. 4, pp. 2240–2245, IEEE, 2001.

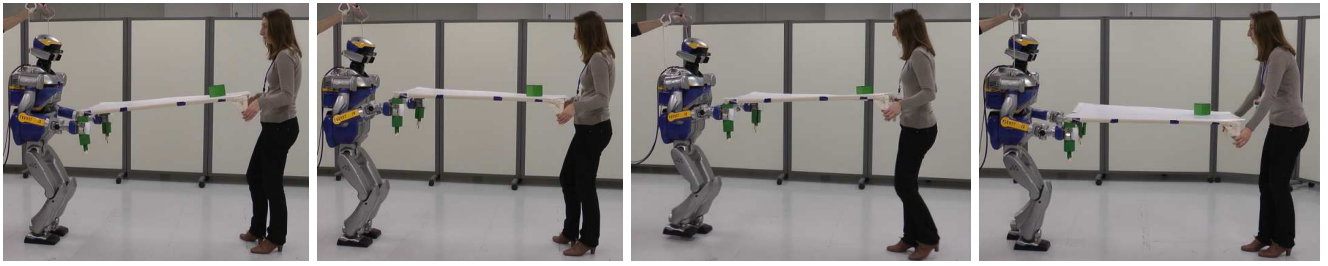


Fig. 5. Snapshots of an experiment for the stationary object (green cube)

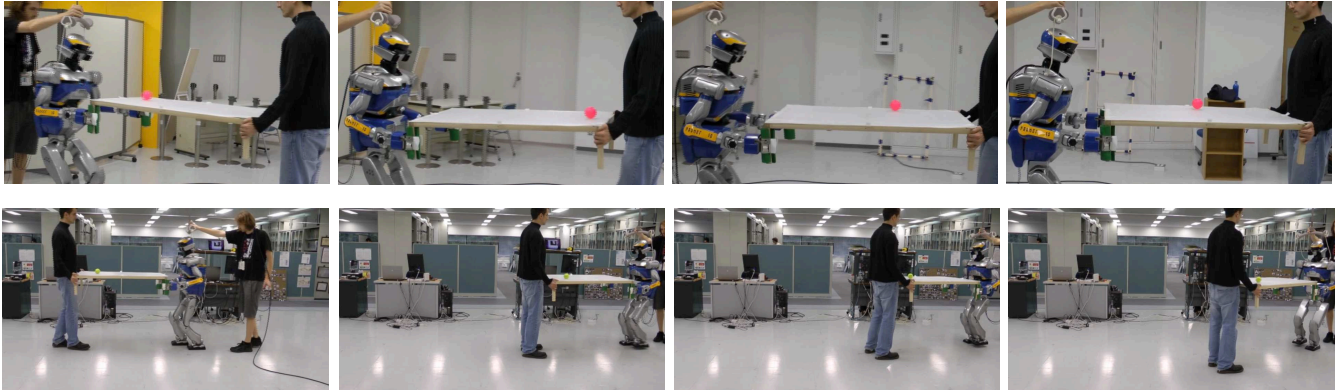


Fig. 6. Snapshots of two experiments where the human-humanoid dyad transports a table with a freely-moving ball on top (a fast moving ball in the top sequence, and a slow moving ball in the bottom one).

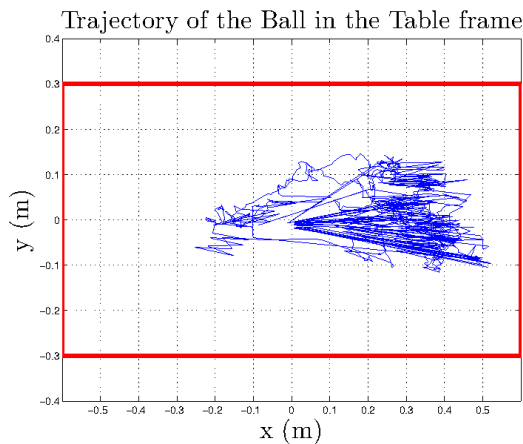


Fig. 8. Controlled ball trajectory (blue) during the experiment, computed from the visual data. The red border is a rough estimate of the table edges

[6] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pp. 4326–4332, IEEE, 2011.

[7] M. O. Ernst and H. H. Bühlhoff, "Merging the senses into a robust percept," *Trends in Cognitive Sciences*, vol. 8, pp. 162–169, April 2004.

[8] A. Kheddar, "Human-robot haptic joint actions is an equal control-sharing approach possible?," in *Human System Interactions (HSI), 2011 4th International Conference on*, pp. 268–273, IEEE, 2011.

[9] B. J. Nelson, J. D. Morrow, and P. K. Khosla, "Improved force control through visual servoing," in *American Control Conference, 1995. Proceedings of the*, vol. 1, pp. 380–386, IEEE, 1995.

[10] M. T. Mason, "Compliance and force control for computer controlled

manipulators," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 11, no. 6, pp. 418–432, 1981.

[11] J. Baeten, H. Bruyninckx, and J. De Schutter, "Integrated vision/force robotic servoing in the task frame formalism," *The International Journal of Robotics Research*, vol. 22, no. 10–11, pp. 941–954, 2003.

[12] M. Prats, P. Martinet, A. P. del Pobil, and S. Lee, "Vision force control in task-oriented grasping and manipulation," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pp. 1320–1325, IEEE, 2007.

[13] N. Hogan, "Impedance control - An approach to manipulation. I - Theory. II - Implementation. III - Applications," *ASME Transactions Journal of Dynamic Systems and Measurement Control B*, vol. 107, pp. 1–24, Mar. 1985.

[14] A. De Santis, V. Lippiello, B. Siciliano, and L. Villani, "Human-robot interaction control using force and vision," *Advances in Control Theory and Applications*, pp. 51–70, 2007.

[15] G. Morel, E. Malis, and S. Boudet, "Impedance based combination of visual and force control," in *Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on*, vol. 2, pp. 1743–1748, IEEE, 1998.

[16] V. Lepetit and P. Fua, *Monocular model-based 3D tracking of rigid objects*. Now Publishers Inc, 2005.

[17] A. I. Comport, E. Marchand, M. Pressigout, and F. Chaumette, "Real-time markerless tracking for augmented reality: the virtual visual servoing framework," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 12, no. 4, pp. 615–628, 2006.

[18] F. Chaumette and S. Hutchinson, "Visual servo control. i. basic approaches," *Robotics & Automation Magazine, IEEE*, vol. 13, no. 4, pp. 82–90, 2006.

[19] N. Mansard, O. Stasse, P. Evrard, and A. Kheddar, "A versatile generalized inverted kinematics implementation for collaborative working humanoid robots: The stack of tasks," in *Advanced Robotics, 2009. ICAR 2009. International Conference on*, pp. 1–6, IEEE, 2009.

[20] A. Herdt, N. Perrin, P.-B. Wieber, et al., "Walking without thinking about it," in *IEEE-RSJ International Conference on Intelligent Robots & Systems*, 2010.