Towards Human-Robot object exchange lessons learned

Anthony Remazeilles[*], Miguel Prada[*], Ainara Garzo[*], Irati Rasines[*], Marco Controzzi[†], Christian Cipriani[†], Ilaria Strazzulla[†], Carlo Peccia[†], Joaquin Canseco[‡], David Cabañeros[‡], Victor Fernandez-Carbajales[‡], Ana Belen Rodriguez[‡], Alan M Wing[§], Elia Gatti[§], Mark Burgin[¶] and Geoff Pegman[¶]
[*]Health Division of Tecnalia Research & Innovation, Donostia, Spain
[†]The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy
[‡]Treelogic, Llanera, Spain, [§]University of Birmingham, UK, [¶]RU Robots Ltd, Manchester, UK

***Abstract*—This article gives an overview of the work conducted in the European project CogLaboration for improving human robot interaction through object exchange that has been iteratively used for around a thousand of interactions. A perception layer using Kinect cameras tracks the object and the human partner's hand and triggers the main robot motion phases. A dedicated object exchange database contains not only the object grasping poses, but also expected hand postures and object orientations to adjust respectively the delivery and grasping strategies. The control of the 7-DoFs LWR arm is designed using the DMP framework. It allows the handling of transport constraints, the online detection of any potential arm kinematics violation and the run-time requesting of a new motion pattern to alleviate this risk. The robot anthropomorphic hand has been equipped with an exteroceptive sensory system (tactile and force) for triggering the handover phases. Comparison of Human-Robot exchange and benchmarking data obtained from Human-Human object transfer points to areas for potential improvement.**

I. Introduction

Recent research results in robotics now make conceivable the close collaboration of humans and robots. In that context, object exchange, a key capability for creating such efficient collaboration, has been an area of increasing interest in recent years. In [15] a multi-criteria cost-map is defined to deduce the most appropriate exchange site. In [9], various motion profiles are compared to define the one most appreciated by humans. In most of these works, the critical handover procedure was not explicitly considered. [1] analyzed by experimentation whether humans preferred the robot to grasp the object with a reactive or fixed grasp mode, but the exchange site was once more defined off-line and not adjusted on line. In [2], the authors focused on the handover procedure through the analysis of the force perceived in the robotic hand that remained static, while humans can do so even when their arm is still moving.

We propose a more reactive exchange procedure in which the exchange location is adapted according to the perceived human motion, and the handover phase is detected even if the robot is moving. This article provides a general overview of the CogLaboration results, from the overall architecture (II), the arm control mechanism (III), the anthropomorphic hand extension (IV), the perception layer (V) and the dedicated exchange database (VI), to the human observation effort done for benchmarking the robotic system proposed (VII).

II. Overall architecture

The CogLaboration prototype has been implemented under the ROS framework. Hierarchical state machines (smach package) orchestrate the operations by triggering the different components involved in the object exchange and easing the information transmission between nodes.

Figure 1 describes the meta operations involved in the robot to human interaction. First information is requested from the knowledgebase to adjust the motion pattern for the next handover (like the object (or hand) to end-effector pose for grasping (delivering) the object, potential transport constraint, suitable grasping mode like cylindrical, tri-digital, etc.). The perception layer (here the human hand tracker) is then started. Once it detects the human motion, the robot automatically moves towards the learned exchange site. During the motion, the handover (grasping or delivery) mode is assessed. The preselected one is connected to an expected human behavior. If the human behavior observed is related to another handover mode, the robot strategy is adapted accordingly. The criteria used are the object orientation and the human hand posture, respectively in the human to robot (HR) and robot to human (RH) exchanges. The human motion monitoring detects also if the human moves towards an exchange site different from the learned one. If so, the robot switches the goal location to the real one observed. The contact monitoring is also launched, through sensors embedded in the arm and in the hand. This will trigger the hand opening and the robot stops. Note that contact detection can be done even if the robot is moving.

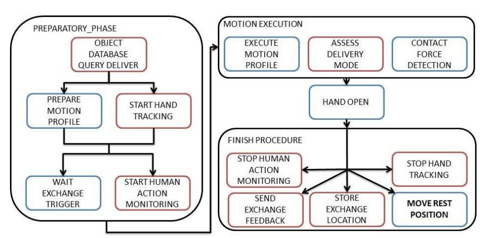

Fig. 1. Illustration of the *Exchange RH* CogLaboration State Machine

Finally, the robot moves back to its rest position, launched processes are inactivated and relevant information is stored to adjust next exchanges (such as the observed exchange location so that the robot can move directly towards it, and the handover mode so that the robot can select it with priority later on).

III. Arm controller

The control of the arm moving towards the exchange site is designed using Dynamic Movement Primitives (DMP). We proposed an extension of the basic formalism for better tuning the transition (when and how fast) between feedforward and feedback phases [11]. We experimented with it [12], using as reference motion pattern the Cartesian trajectory of a human hand captured during behavioral analysis (see sec. VII). The goal location was set to the hand or object being tracked.

We saw two major limitations of such an implementation. The first issue is that if the human behavior deviates from the reference pattern, the robot may require high velocities and accelerations to compensate for the error which may make the robot reach its limits and enforce the saturation of the velocities applied. The second drawback is that if the human desires to exchange the object at another location, the feedforward phase would delay the reaction to the feedback phase, which is likely to make the first limitation occur.

The kinematic constraint violation has been alleviated in three ways. Firstly, we designed the controller with a two-level architecture. The high level component prepares the motion pattern that the low-level component will apply and monitor. The pattern executor is able to foresee any potential velocity limit violation by applying a few iterations of the DMP scheme in advance. If this occurs, it requires the high-level controller to provide online a new motion pattern. Secondly we propose to use the reflexxes library to generate the motion pattern. The advantages of this library are two-folds: (i) it can generate a trajectory pattern in real-time, and (ii) it is able to satisfy initial and final velocity and acceleration requests. This is convenient for smoothly online switching the motion pattern if the current one is considered inappropriate. Thirdly, we decided to use the joint space as main control space, to make sure the trajectory generated by the pattern generator will respect the robot constraint. Nevertheless, the Cartesian space is kept for objects with transport constraints (like a cup that needs to be maintained vertical). In such case only the Cartesian position is handled by the DMP, while the hand orientation is adjusted to the transport constraint.

The second limitation, related to the lack of adaptation of the feedforward phase, is mitigated by adding an exchange site learner: after each exchange, the location reached by the human partner (obtained by the object or hand tracking depending on the exchange direction) is recorded, and the next exchange will be using a default motion pattern computed by the pattern generator towards that location. This way, if the human requests exchanges in a similar location in the future, the robot will directly move towards that site.

IV. Sensorized anthropomorphic hand

The robot hand is an improved version of the IH2-Azzurra Hand commercialized by Prensilia Srl, Italy [3]: it consists of five underactuated digits (two joints per digit, for a total of 11 DoFs) driven by five motors which actuates the flexion-extension of the thumb, index, middle, ring-little as a pair and the ab-adduction of the thumb (see Fig. 2-(a)). The robot hand is self-contained (i.e. all functional components are housed within the size of the hand itself) and weighs 520g.

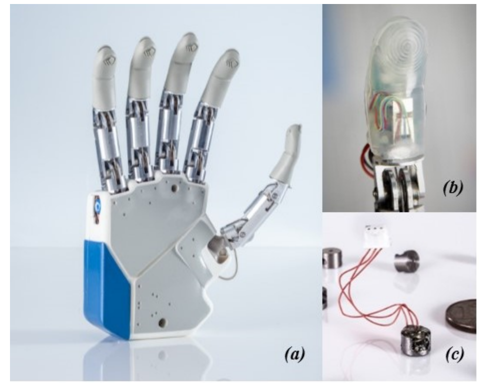

Fig. 2. (a) IH2 Azzurra hand, (b) 2-axis force sensor embedded within the compliant fingertip, (c) miniaturized tendon tension force sensor.

The robot hand has been provided with bio-inspired compliant fingers [4] to increase the grasp capabilities and to ensure a safe interaction with the human partner. Taking inspiration from the multi-layered structure of the human finger, the artificial fingertips developed comprise of a rigid core covered by two layers of polymeric materials with different degrees of stiffness and topped by a hard nail. This specific design is crucial to obtain mechanical features and appearance similar to the human fingertips.

The embedded control of the hand is arranged in a hierarchical architecture consisting of five Low Level Motion Controllers (LLMC) and one High Level Hand Controller (HLHC). Each motor is directly actuated and controlled with a LLMC that achieves position, force, and current control. All the LLMC are controlled by the HLHC that regulates the overall hand operations through high level functions (like automatic grasps or automatic object release for a fluent handover [5]) and acts as interface with the external world.

Both control levels are crucially dependent on the artificial sensory system of the hand (especially for cooperative actions). For this reason we extended the sensory system through an exteroceptive and a proprioceptive sensory subsystem. The first monitors and measures the interaction between the grasped object and the robot hand, and between the object and the environment (collision with an obstacle, contact with the human partner, and so on). The second provides useful information about hand kinematic and internal forces produced in the hand transmission. The exteroceptive sensory system consists of different sensors functionally emulating touch sensors in the human skin using different low-cost and reliable technologies, in particular: i) four tendon tension force sensors (see Fig. 2-(c), housed within the actuation unit of the hand and monitoring the tensile force stressing the tendons that drive the fingers), ii) three fingertip 2-axis force sensors (see Fig. 2-(b), mounted on the Thumb, Index and Middle fingers, monitoring normal and tangential forces acting at the tip of the finger), iii) two fingertip touch sensors (mounted on the Ring and Little fingers, informing through a digital output about the contact status), and iv) one palm touch sensor. All sensors developed (excluding the tension force and the palm touch sensors for practical reasons) are embedded in the compliant fingertip developed [4].

V. Perception

The perception layer relies on two XBox 360 Kinect sensors and several software modules developed using ROS and PCL frameworks. The object recognition component is capable of handling both particular instance and general object categories. This system is based on multi-view model training and multi-descriptor classification. A dataset of more than 300 point clouds has been captured for evaluating the system, and an accuracy of near 91% has been observed.

The object pose is estimated in two steps. The object location is found using several segmentation and hand detection strategies. The orientation is then computed using RANSAC IA [14], and refined with the Iterative Closest Point [8] approach. For the challenging task of estimating the object pose when held by the human, an orientation sensor developed in the project improves the RANSAC estimation.

The human motion tracking system locates the human hand and tracks it during the whole exchange process. It is developed using multi-color space skin segmentation, skeleton tracking and Kalman filtering prediction approaches [10].

Finally the perception layer monitors online the human handover strategy in order to adjust the robot strategy accordingly. In HR exchanges, the tracked object orientation is used to check if the grasping mode initially selected remains relevant, as stored within the knowledge base. In RH exchange, the human hand posture is recognized through Gaussian Mixture Models to adjust the robot approach strategy to the human expectation [13].

VI. Object exchange knowledgebase

The object exchange database models the knowledge acquired from human-human (HH) exchange experiments. The database interoperability and extensibility is obtained using a semantic-ontological [16] approach, complemented with a set of ad-hoc utilities for easing the knowledge inference [17], query and management tasks. Contrary to traditional relational databases, the semantic-ontological modeling of the objects taxonomy strongly eases the modification and extension of the knowledgebase. The ontology model is based on the object affordances concept [7], described as the sum of the properties of a situation, including agents, environment and objects, especially those that describe how they can be used to act.

In addition to the object visual properties needed for the Perception layer, the knowledgebase contains a set of features characterizing the handover conditions. Each object is related to a set of stable grasp poses and postures as provided by the robot hand and based on the Cutkosky taxonomy [6], that are indexed with respect to an expected object orientation. The delivery mode is also considered, defining for each grasp mode a list of possible delivery strategies (pose wrt to the hand, approach direction, etc.) indexed to different expected human partner hand postures. This extends the initial conception of a grasping database to a fully-featured exchange knowledgebase.

VII. Comparative experimentation

A. Human-human interactions

In order to evaluate the performance of the robotic system, a number of studies of human handover interaction were first undertaken. Each study involved two participants, one taking the role of assistant passing or receiving objects to and from the other in order to move objects from one table to another (see Fig. 3(left)). The aim was to characterize the efficacy of the handover action with regard to scenario changes defining task (eg car mechanic vs activity of daily living), posture (eg standing, seated, lying) assigned to the person being assisted and as a function of the objects being transferred and instructions to the participants to vary the style of the handover (eg normal vs slow). Measures of the handover then provided a benchmark for evaluating performance when the robot takes the role of the assistant.

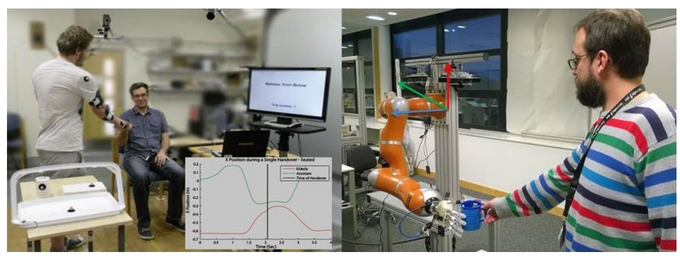

Fig. 3. Human Human (left) and Human Robot interactions (right)

Data collected during HH handover trials covered three aspects. The first comprised subjective quality assessment of the exchange defined with 9-point Likert-scale items such as I was satisfied with the interaction, the interaction was easy, etc. The second type of data collected involved motion capture recordings of the participants hand positions. The third focus was on motion of the objects, which were instrumented to indicate contact by each participant, object orientation, acceleration and jerk.

B. Human Robot interactions

In order to test the robotic handover system, two experiments were performed, one with adults of working age and the second with older adults over 55.

In general, the quality reports obtained by both populations of users were satisfactory. In fact, participants reported their interaction with the robot as a comfortable and safe experience. They also judged the handovers overall as easy and natural, with the robotic system behaving in a predictable way.

Quantitative results obtained by motion capture were compared between the two populations of interest and the data from the same scenario (activities of daily living) obtained from the studies of HH interaction. Comparison was performed on one spatial feature (handover correction: how much the user needed to move to the handover position in order to handle the object to the robot) and one temporal feature (movement peak velocity while reaching the handover position) (see Fig. 4). In general, results of adults and elderly did not differ. On the other hand, difference between HH and HR interaction was significant in both the tested features, although relatively small. This suggests that, despite the satisfaction of the user and the reported usability of the system from a subjective perspective, work has still to be done in order to achieve an interaction comparable to the human one.

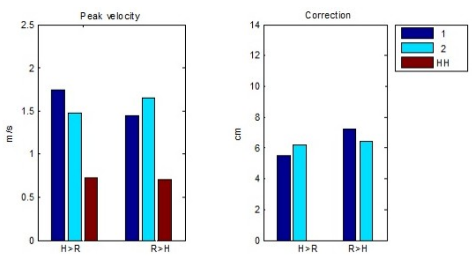

Fig. 4. Comparison of peak velocity and spatial correction for younger (1), elderly adults (2), and Human-Human transfer

Data from sensors embedded in the instrumented objects also supported this position. In particular, exchange time was approximately five times shorter in the HH trials compared to HR trials. Moreover, the variation in object orientation during the handover was two to three times larger in HR trials than in HH trials. One interpretation of these differences might indeed be an increase difficulty of the HR interaction compared with Human-Human interaction.

VIII. Conclusion

This article has given an overview of the work conducted in the CogLaboration project, in which we have proposed several technical improvements, in terms of hardware in particular in the design of tactile and force sensors for the anthropomorphic hand, and in terms of software, for obtaining reactive arm control, the appropriate visual perception of the scene, and a dedicated knowledge database for object exchange. A special feature of this work is that it considered human robot exchange in both directions.

The dataset captured during human observation will be made accessible as a benchmarking tool for further work in this field. As the system experimentation has shown, the human subjects give an overall good rating to the robot behavior. Nevertheless, the quantitative comparison with human behavior shows that such formalism could be improved, in term of velocity and reactivity, to get closer to the timing observed with humans. The visual perception of the exchange advancement, in particular at the handover phase, remains also a challenging issue that would need to be tackled to get a more fluid interaction with humans.

We encourage interested readers to visit the project website, to get a deeper look at the different deliverables produced during the project.

Acknowledgments

This work was supported by the CogLaboration project within the 7^{th} Framework Programme of the European Union (FP7-28788), see www.coglaboration.eu. We would like also to thank M. J. Puerto for her valuable contribution, as well as all the people that took part in our handover studies.

References

[1] M. Cakmak, S. Srinivasa, M. Lee, J. Forlizzi, and S. Kiesler. Human preferences for robot-human hand-over configurations. In *IEEE/RSJ IROS*, pages 1986–1993, San Francisco, CA, 2011.

[2] P. Chan, C. Parker, M. Van der Loos, and E. Croft. A human-inspired object handover controller. *International Journal of Robotics Research*, 32(8), 2013.

[3] C. Cipriani, M. Controzzi, and M. C. Carrozza. The smart-hand transradial prosthesis. *Journal of neuroengineering and rehabilitation*, 8(1):29, 2011.

[4] M. Controzzi, M. D'Alonzo, C. Peccia, C. M. Oddo, M. C. Carrozza, and C. Cipriani. Bioinspired fingertip for anthropomorphic robotic hands. *Applied Bionics and Biomechanics*, 11 (1):25–38, 2014.

[5] M. Controzzi, I. Strazzulla, C. Peccia, A.M. Wing, and C. Cipriani. Improving the fluency of the robot to human handovers using a human inspired feed-forward release controller. *Inter. Workshop on Human-Friendly Robotics*, pages 1–2, Oct. 2014.

[6] M. R. Cutkosky. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Trans. on Robotics and Automation*, 5(3):269–279, 1989.

[7] JJ Gibson. The theory of affordances. *Hilldale USA*, 1977.

[8] C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey vision conference*, pages 147–151. Manchester, UK, 1988.

[9] M. Huber, H. Radrich, C. Wendt, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer. Evaluation of a novel biologically inspired trajectory generator in human-robot interaction. In *IEEE RO-MAN*, pages 639–644, Toyama, Japan, Sept. 2009.

[10] X. Li, K. Wang, W. Wang, and Y. Li. A multiple object tracking method using kalman filter. In *IEEE ICIA*, pages 1862–1866, 2010.

[11] M. Prada, A. Remazeilles, A. Koene, and S. Endo. Dynamic movement primitives for human-robot interaction: Comparison with human behavioral observation. In *IEEE/RSJ IROS*, pages 1168–1175, Tokyo, Japan, 2013.

[12] M. Prada, A. Remazeilles, A. Koene, and S. Endo. Implementation and experimental validation of dynamic movement primitives for object handover. In *IEEE/RSJ IROS*, pages 2146–2153, Chicago, USA, 2014.

[13] I Rasines, A. Remazeilles, and P. Iriondo. Feature selection for hand pose recognition in human-robot object exchange scenario. In *IEEE ETFA*, pages 1–8, Barcelona, Spain, 2014.

[14] R. Rusu, N. Blodow, and M. Beetz. Fast point feature histograms (fpfh) for 3d registration. In *IEEE ICRA*, pages 3212–3217, Kobe, Japan, 2009.

[15] E. Sisbot, L. Marin-Urias, R. Alami, and T. Simeon. A human aware mobile robot motion planner. *IEEE Trans. on Robotics*, 23(5), 2007.

[16] K. Varadarajan and M. Vincze. Ontological knowledge management framework for grasping and manipulation. In *IROS Workshop: Knowledge Representation for Autonomous Robots*, 2011.

[17] K. M. Varadarajan and M. Vincze. Knowledge representation and inference for grasp affordances. In *Computer Vision Systems*, pages 173–182. Springer, 2011.