Assistance Adaptation during Human-Robot Haptic Joint Manipulation Task

Julie Dumora*, Franck Geffard*, Catherine Bidard*, Nikos A. Aspragathos† and Philippe Fraisse‡
*CEA, LIST, Interactive Robotics Laboratory, Gif-sur-Yvette, France, Email: firstname.lastname@cea.fr
†Mechanical Engineering and Aeronautics Department, University of Patras, Greece, Email: asprag@mech.upatras.gr
‡LIRMM, University of Montpellier 2/CNRS, France, Email: fraisse@lirmm.fr

***Abstract*—This work focus on human robot haptic joint collaboration for achieving large object manipulation tasks. We present our current lines of research addressing the topics of assistive control, haptic measures interpretation, human intent recognition and decision making parameters tuning in the context of large object manipulation.**

I. Introduction

In a physical human robot collaboration, each partner has complementary capacities: cognitive for the human *vs* physical for the robot. A huge challenge is the combination of these complementary strengths to create a synergy between them. The robot should physically assist the human partner who is able to decide on what to do and to instantaneously plan complex scenarios. To that end, the robot has to detect which motion assistance is needed by the human operator at each instant.

In the next Section, we present our physical human robot collaboration framework for large object manipulation. In Section III, we introduce our method for automatically tuning the intention detection parameters. Finally, we draw some conclusions in Section IV.

II. Physical Human Robot Collaboration Framework for Large Object Manipulation

In our context, the robot has no prior information about the task and the environment. Our method consists in endowing the robot with a library of assistances for performing standard collaborative motions. According to the haptic cues naturally transmitted by the human partner, the robot selects on-line the suitable assistance for the current intended collaborative motion. Thus, the operator can benefit from a sequence of suitable assistances generated on-line without prior knowledge about the task.

A. Assistive Control Framework: a Mechanical Analogy

One of the main goal is to **reduce operator effort**. However, in physical human robot interaction, the **safety** is fundamental. Moreover, the **user's feeling** is obviously a very important criteria.

According to these requirements, the assistance for each collaborative motion is carried out with a virtual motion guide, implemented according to the virtual mechanism concept [6]. This concept is based on a mechanical analogy. The robot end effector is virtually coupled with a virtual mechanism (VM) end effector, through a spring damper (Fig. 1):

$$\mathrm{F} = \mathrm{K}(\mathrm{X}_{vm} - \mathrm{X}) + \mathrm{B}(\mathrm{V}_{vm} - \mathrm{V})$$

with X_{vm} and X respectively the cartesian position of the virtual mechanism and of the robot, V_{vm} and V the cartesian velocity of the virtual mechanism and of the robot, B and K are the cartesian coupling gains between the robot and the virtual mechanism, respectively the damping and stiffness gains, F the wrench applied by the spring damper on the robot.

The VM geometric model and its Jacobian (L_{vm} and J_{vm}) define free motions, which are controlled in force, and constraints, which are controlled in position (Fig. 2). This assistive control ensures the passivity criteria, which guaranties the stability of the robot in interaction with any other passive system [2]: safe interactions with the environment and the human partner are ensured. Furthermore, a psychology study showed that human operators feel more comfortable and safe with a passive robot [8]. This result can be interpreted by the ability of the human to predict the passive robot behaviour while this is more difficult with a robot that continuously learns during tasks, or that has to estimate parameters for its trajectory generation. Moreover, previous works have shown that motion guidance can significantly reduce the manoeuvring effort of the human operator as well as the completion time of the task [1].

B. Human Motion Intent Interpretation through Haptic Measures: a Mechanical Based Approach

When the current assistance corresponds to the operator's motion intent (agreement direction), the collaborative motion

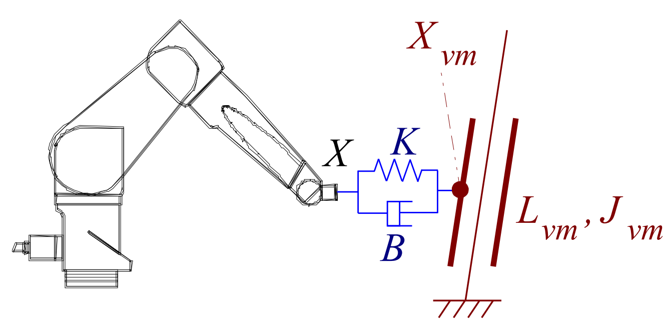

Fig. 1. Principle of virtual mechanism: mechanical analogy

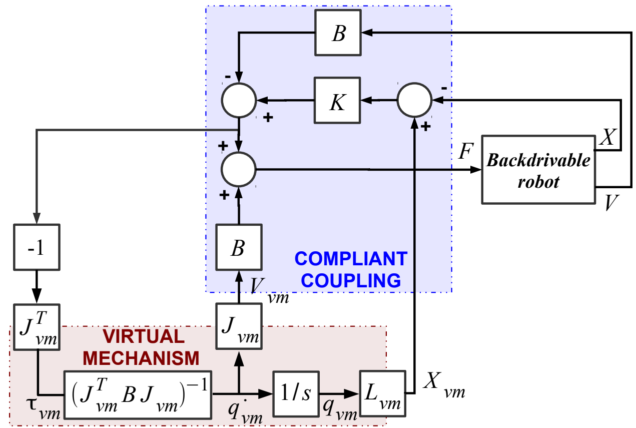

Fig. 2. Control scheme with virtual mechanism. *q and $\dot{q}$ are respectively the joint position and velocity. τ and F are respectively the joint torque and wrench applied by the spring damper on the robot. The vm subscript denotes the virtual mechanism.*

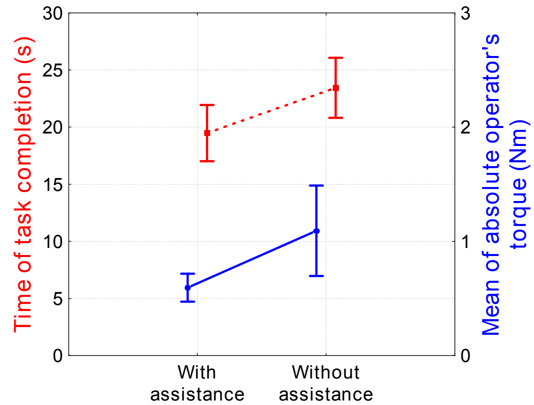

Fig. 3. Performance indices with and without assistance. Vertical bars denote $+/-$ standard errors.

can be performed with low effort since the motion guide is well-suited. When the human intends another collaborative motion, new directions of motions are solicited. The provided guide is not well-suited anymore since it constrains directions in which the human wants to move (disagreement direction). Therefore, wrench and twist patterns are used to detect the human partner intention of motion in order to select the most suitable guide. The idea is to describe the intended motion by a set of constraints and free direction of motion in the same way that the robot behaviour is described by the virtual mechanism. Through these models, the disagreement and agreement directions reveal the wrench and twist components that should be measured. The relationships between human intent, current assistance and haptic measures have been intuitively inferred on a case study and experimentally validated [3].

C. Benefits of the Human Robot Interaction Framework

We implemented a basic algorithm of intention detection to switch from the current assistance to the new suitable one. This algorithm is based on the relationships between haptic measures and human intent (See II. B). Thresholds were empirically tuned to determine significant haptic measures, which are useful for the intention detection. A user study highlights the improvement of the human-robot interaction with this method compared with a follower robot, for achieving a task involving different rotation/translation motions of a long object [4]. Both the torque applied by the operator and the completion time of task are reduced with the proposed method (Fig. 3).

In the next section, we present a method to automatically tune the intention detection parameters.

III. Tuning Human Intention Detection Parameters

A. Goal

In order to provide a suitable assistance, the goal is to infer the current operator's intention among the m overall possible intentions $\mathrm{C} = \{\mathrm{C}_1, ..., \mathrm{C}_m\}$ given the n current haptic measures $\mathrm{F} = \{\mathrm{F}_1 = f^1, ..., \mathrm{F}_n = f^n\}$. f^i is the measure of the *i*-th component F_i of the haptic data. The haptic data refers to wrench components as well as the displacement at the operator's grasping point.

B. Method

The features that are useful for the operator's motion intent inference are measured in directions constrained by the VM. This reason drives us to carry out a static analysis of these measures. The method is based on the Naive Bayes Classifier. This supervised learning algorithm relies on the Bayes' theorem and the class independent assumption.

Learning Step: A learning step is performed **once** in order to tune the intention detection parameters. According to a dataset, the method estimates the posterior probability of belonging to each intention by estimating the likelihood parameters:

$$\mathrm{P}(\mathrm{C}_k | \bigcap_{i=1}^{n} \mathrm{F}_i = f^i) = \frac{\prod_{i=1}^{n} \mathrm{P}(\mathrm{F}_i = f^i | \mathrm{C}_k)}{\eta}$$

$$\text{with} \quad \eta = \frac{1}{m.\mathrm{P}(\bigcap_{i=1}^{n} \mathrm{F}_i = f^i)}$$

with η the normalizing factor, which includes the prior and evidence that are independent of the class. The m prior probabilities are assumed equipossible since no information about the task are given: $\mathrm{P}(\mathrm{C}_k) = \frac{1}{m}$. Each haptic component F_i given the intention C_k is assumed to follow a normal distribution:

$$\mathrm{P}(\mathrm{F}_i = f^i | \mathrm{C} = \mathrm{C}_k) = \mathcal{N}(f^i, \mu_{ik}, \sigma^2_{ik})$$

The learning problem consists in estimating a set of the gaussian parameters for each haptic component given the intention $\mathrm{S}_{\mathrm{F}_i|\mathrm{C}_k} = \{\mu_{ik}, \sigma^2_{ik}\}$, with μ_{ik} the mean and σ_{ik} the standard deviation. A set of these parameters sets $\mathrm{S} = \{\bigcap_{k=1}^{m} \bigcap_{i=1}^{n} \mathrm{S}_{\mathrm{F}_i|\mathrm{C}_k}\}$

is estimated from a training dataset of D demonstrations for each intention. We use the maximum likelihood estimator:

$$\hat{\mu}_{ik} = \frac{1}{W.D}\sum_{d=1}^{D}\sum_{w=1}^{W} f^i_{w,d}, \quad \hat{\sigma}^2_{ik} = \frac{1}{W.D}\sum_{d=1}^{D}\sum_{w=1}^{W}(f^i_{w,d} - \hat{\mu}_{ik})^2$$

The current virtual mechanism prevents from demonstrating the switch from a motion to a next one. However, the user might demonstrate his motion intent by forcing against the virtual guide as long as he might be able to apply effort. Therefore, unlike straightforward supervised classification, no signal length is suitably defined by the demonstrations. We need to determine the optimal number of training signal data points W that will be used for the parameters estimation in order to get the best intention detection performance. To that end, one set of parameters $\mathrm{S} = \mathrm{S}_W$ is computed from data points within a window of size W, with W incrementally increased from 1 to the whole signal size (Fig. 4). Then, a cross validation step is carried out in order to select the best parameters set. The Matthew Correlation Coefficient (MCC) [7] is used for assessing the multi-class classification [5]. This metric returns a value between -1 and $+1$ and can be intuitively interpreted: -1 represents a total disagreement between prediction and observation, 0 a random prediction and $+1$ a perfect prediction. For each set S_W, the MCC mean μ_{MCC} in a chosen confidence level interval $\mathrm{I}_\gamma = [\gamma_{min}; \gamma_{max}]$ is computed:

$$\mu_{\text{MCC}} = \frac{1}{N_\gamma}\sum_{i=1}^{N_\gamma}\text{MCC}_i$$

with N_γ the number of MCC computed in the interval I_γ. The chosen set of parameters S^* is the one with the highest MCC

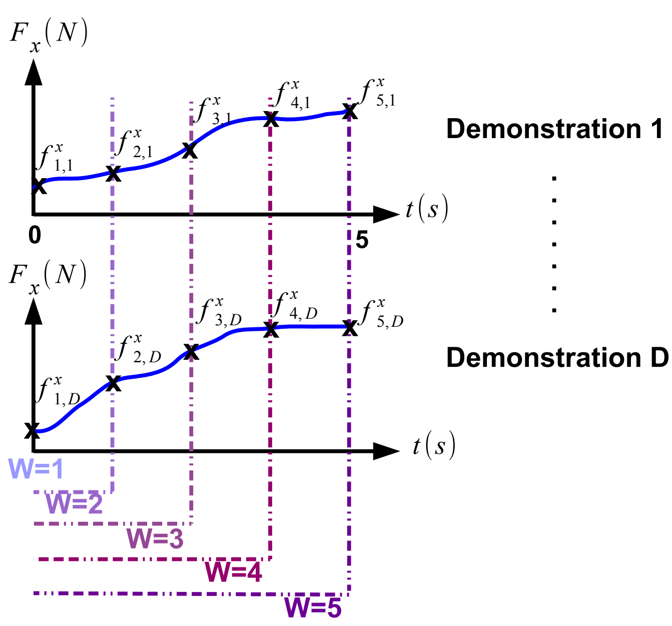

Fig. 4. Illustration of windows for D demonstrations of a training dataset.

mean:

$$\mathrm{S}^* = \arg\max_{\mathrm{S}_W}(\mu_{\text{MCC}})$$

Prediction step: Once the learning step has been carried out, the assistance can be adapted on-line, according to the current haptic measures.

The robot provides a new assistance when it predicts a new operator intention C^* with a confidence degree of at least γ:

$$\mathrm{C}^* = \arg\max_{\mathrm{C}_k}(\mathrm{P}(\mathrm{C}_k | \bigcap_{i=1}^{n} \mathrm{F}_i = f^i)) \geq \gamma$$

As long as the confidence level is not overstepped, the robot remains in the current state since it is not enough confident of the operator's intention. Accounting for a confidence level in the decision rule instead of the simple MAP rule allows to be more robust.

C. Experiment

Experimental setup: We confine the experiment in the horizontal plane. The library of assistances is composed of four assistances for performing basic motions that an operator might achieve in a collaborative planar task. These motions are described hereafter:

- $\mathrm{T_Y}$ (Fig. 5. a.): pure lateral translation;
- $\mathrm{T_X}$ (Fig. 5. b.): front/back translation;
- $\mathrm{R_O}$ (Fig. 5. c.): rotation of the object around the operator's gripping point;
- $\mathrm{R_A}$ (Fig. 5. d.): rotation of the object around the robot's gripping point.

Starting with a current assistance for performing $\mathrm{R_A}$ motion, the goal is to detect the next intended motion C^* among $\mathrm{C} = \{\mathrm{T_Y}, \mathrm{T_X}, \mathrm{R_O}, \mathrm{R_A}\}$ (Fig. 5). The relevant features highlighted in relationships inferred in [3] are used: $\mathrm{F_X}$, $\mathrm{M_Z}$ and $\Delta\mathrm{X_Y}$. They are expressed at the operator's grasping point O in the frame depicted Fig. 5. The current assistance for performing $\mathrm{R_A}$ motion is carried out with a virtual mechanism described by a pivot link.

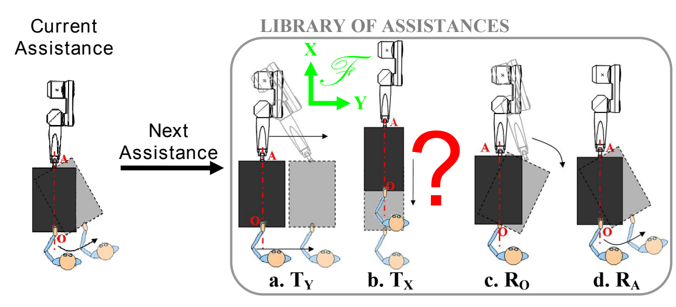

Fig. 5. Goal of the experiment: Which is the intended motion?

Thirty subjects intended to perform each collaborative motion five times. Although the virtual mechanism prevented them from performing $\mathrm{T_Y}$, $\mathrm{T_X}$ and $\mathrm{R_O}$, they had to intend to execute these motions. Each test lasted 5s. The force and torque at the user's handle as well as the position of each robot joint were recorded at 500Hz. The measurements of the robot joint positions have been used to measure the user's handle position, in order to obtain the displacement of the operator's hand. A Force/Torque sensor at the user's handle have been used to acquire the wrench measurement.

Data from 15 participants who have been randomly selected constitute the training dataset. Data from 8 participants who have been randomly selected among the remaining 15 participants are used for the cross validation dataset. A test is carried out with data of the remaining 7 participants.

Results: The different parameters sets $\mathrm{S_W}$ have been computed. As an example, Fig. 6 presents the estimated parameter $\hat{\mu}_{F_X|T_X}$ for each set $\mathrm{S_W}$. Note that the higher the window size is, the higher the mean of the force $\mathrm{F_X}$ is. Using a small window for the learning step allows to predict the intention before the operator applies too much effort. Fig. 7 presents

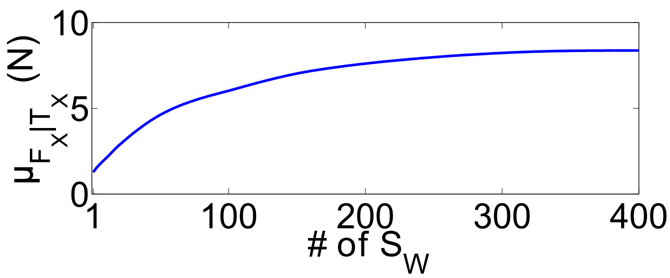

Fig. 6. Mean of $\mathrm{F_X}$ feature given the intention $\mathrm{T_X}$ for each set of parameters $\mathrm{S_W}$, *i.e* the evolution of $\hat{\mu}_{F_X|T_X}$ according to demonstration signals length used for the estimation.

the classification assessment obtained with each parameters sets S_W. The red bar points out the chosen parameters set S^* with the highest performance. The comparison between the mean $\mu_{\mathrm{F}_X|\mathrm{T}_X} = 6.63\mathrm{N}$ in the selected set and the mean in sets with a higher W (Fig. 6) shows the interest of selecting a demonstration signal length: better performances are obtained while applying lower efforts. Finally, Table I presents the

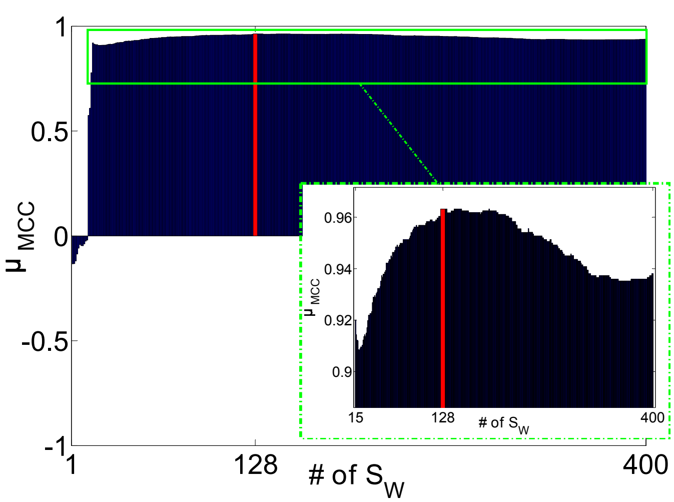

Fig. 7. Classification performance obtained with each set of parameters S_W.

motion intention classification results carried out on 140 trials with a confidence level of 90%. A 97% success rate has been reached with only 2 wrong detections and 2 decisions that have not been made by lake of confidence in the detection.

TABLE I
CONFUSION MATRIX OBTAINED WITH $\gamma = 0.9$

		Predicted intention			
		$\mathrm{T_X}$	$\mathrm{T_Y}$	$\mathrm{R_O}$	$\mathrm{R_A}$
Actual intention	$\mathrm{T_X}$	35	0	1	0
	$\mathrm{T_Y}$	0	35	0	2
	$\mathrm{R_O}$	0	0	33	0
	$\mathrm{R_A}$	0	0	1	33

IV. CONCLUSION

Three contributions have been presented in this paper. First, assuming that many tasks can be decomposed into a sequence of collaborative motions, our physical human robot collaboration framework for large object manipulation allows to perform a wide spectrum of tasks with a sequence of assistances generated on-line while the robot has no prior knowledge about the task. Moreover, the robotic partner fulfils all these important features:

- safety
- predictability
- human effort reduction
- movement redundancy (but should be enriched)
- interpretation of human disagreement and intention (unexpected variability in order to provide suitable assistance)
- task-independence

Finally, the method for tuning the intention detection parameters can be applied without demonstrating the switch between motions. As no signal length is suitably defined by the demonstrations, the proposed method selects the optimal number of demonstration data points that will be used to estimate the intention detection parameters in order to get the best detection performance. A penalty may be applied on the window width increase in order to have a trade-off between the applied effort (and time switching) and the improvement of the detection performance. The aim is to extend this method to tackle this problem as an early recognition problem in order to avoid the operator to apply too much efforts.

REFERENCES

[1] E.S. Boy, E. Burdet, C.L. Teo, and J.E. Colgate. Motion guidance experiments with scooter cobot. In *11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 63 – 69, March 2003.

[2] J.E. Colgate. *The Control of Dynamically interacting Systems*. PhD thesis, Massachusetts Institute of Technology, 1988.

[3] J. Dumora, F. Geffard, C. Bidard, T. Brouillet, and P. Fraisse. Experimental study on haptic communication

of a human in a shared human-robot collaborative task. In *IEEE/RSJ IROS*, 2012.

[4] J Dumora, F. Geffard, C. Bidard, and P. Fraisse. Towards a robotic partner for collaborative manipulation. In *HRI Workshop on Collaborative Manipulation*, March 2013.

[5] J. Gorodkin. Comparing two k-category assignments by a k-category correlation coefficient. *Computational Biology and Chemistry*, 28(56):367–374, December 2004.

[6] L.D. Joly, C. Andriot, and V. Hayward. Mechanical analogies in hybrid position/force control. In *IEEE ICRA*, volume 1, pages 835–840, April 1997.

[7] B.W Matthews. Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et biophysica acta*, 405(2):442–451, October 1975.

[8] A. Mortl, M. Lawitzky, A. Kucukyilmaz, M. Sezgin, C. Basdogan, and S. Hirche. The role of roles: Physical cooperation between humans and robots. *The International Journal of Robotics Research*, August 2012.