

Sensory-motor augmentation of the robot with shared human perception

This is the peer reviewed version of the following article:

Original:

Ishida, R., Meli, L., Tanaka, Y., Minamizawa, K., Prattichizzo, D. (2018). Sensory-motor augmentation of the robot with shared human perception. In Proc. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp.2596-2603). IEEE [10.1109/IROS.2018.8594496].

Availability:

This version is available <http://hdl.handle.net/11365/1066842> since 2019-05-02T17:38:44Z

Publisher:

IEEE

Published:

DOI:10.1109/IROS.2018.8594496

Terms of use:

Open Access

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. Works made available under a Creative Commons license can be used according to the terms and conditions of said license.

For all terms of use and more information see the publisher's website.

(Article begins on next page)

Sensory-motor augmentation of the robot with shared human perception

Ryuya Ishida¹, Leonardo Meli², Yoshihiro Tanaka^{1,3}, Kouta Minamizawa⁴, Domenico Prattichizzo^{2,5}

Abstract—Robots have replaced people in many manufacturing production lines but the information they gather from sensors might not be sufficient to autonomously accomplish dexterous manipulation operations. Symbiotic human-robot cooperation appears to be a more realistic near future in industrial scenarios. In this paper we present a configuration of human-robot collaboration in which the robot is sensory-augmented by means of a set of tactile signals coming from the human operator. The incorporation of low-level robot “intelligence” permits the cooperative manipulation of an object while enabling the human operator to stay focused on task itself and carry it out in the most natural way. The effectiveness of this approach is demonstrated in a use case in which a robot helps a human operator to successfully accomplish a writing task. System performance has been evaluated, considering several positions of the tiny vibration sensor in charge of gathering the human perception, by testing it on both the human hand and the co-manipulated object. Results suggest that the sensor provides valuable information for recognizing operator actions when it is placed either on the human hand or on the co-manipulated object. However, the sensor on the finger directly represents the operator’s perception, while the output of the sensor attached to the object changes according to the distance between the interaction point and the sensor itself. In addition, in wearing the sensor, neither the object nor the robot need to be instrumented: the operator is free to interact with a large set of objects and collaborate with any existing robot without requiring supplemental equipment.

I. INTRODUCTION

A. Background

Human-Robot Interaction (HRI) has recently received major attention in the academic community as well as in cutting-edge technology companies. The main goal is the design of robotic systems capable of being used in a natural, intuitive, and safe way by and with human beings. By definition, any interaction requires communication, and the communication between a human and a robot might take place in various forms. Considering whether the human and the robot are in close proximity to each other, we can talk about two main categories: i) remote interaction and ii) proximate interaction [1]. In the former, the human

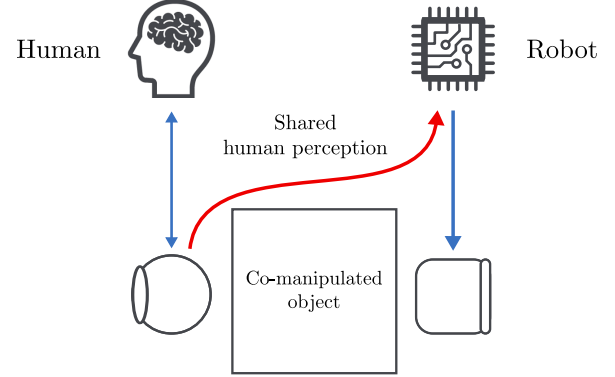


Fig. 1. Sensory-motor augmentation of the robot by the shared human perception. The robot moves according to what the human perceives.

and the robot are separated either spatially, or temporally. In the latter, both the agents are co-located and share the same workspace. Remote interaction is typically referred to as teleoperation or telemanipulation and it is broadly discussed in the literature [2], [3]. It deals with the possibility of remotely extending human manipulation capabilities by providing the user with similar conditions to those in the remote location. This can be achieved using two subsystems called master and slave. The master interface is in charge of gathering human commands and displaying back forces or any other important signal measured where the task is actually performed. The slave interface is a robot in contact with the remote environment that physically executes the task.

Only dependable robot architectures are acceptable for supporting “human-in-the-loop” conditions and human-robot teams [4]. When a robot and human are in the same room and share the same space, they can have a physical interaction either using or not using an object as a means; for instance, in the application domain of houseworks, the robot might act as an assistant or companion. In particular, in social interaction including the emotive and cognitive aspects, the ability for physical and psychological intimacy with humans need to be present [5].

The advent of robots has substantially transformed the industrial sector in terms of manufacturing capability, but challenges in integrating and programming them have restrained their widespread application across verticals. The new industry 4.0 paradigm has a prospect of more close or direct co-operation of robots with human workers [6]. Robots physically assist humans and enhance human ability in terms of force, speed and accuracy; on the other hand, the human can bring experience, global knowledge and perception and understand the proper execution of tasks [7].

The research leading to these results has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement n°688857 of the project “SoftPro” and Japan Science and Technology Agency, PRESTO n°JPMJPR14D7.

¹Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology, Gokiso-cho, 466-8555 Nagoya, Japan.

²Dept. of Information Engineering and Mathematics, University of Siena, Italy.

³JST, PRESTO, Kawaguchi, Japan.

⁴Graduate School of Media Design, Keio University, Yokohama, Kanagawa 223-8526, Japan.

⁵Dept. of Advanced Robotics, Istituto Italiano di Tecnologia, Italy and with the Dept. of Information Engineering and Mathematics, University of Siena, Italy.

In the late 80s, the concept of shared control was already present in the literature [8], [9]. It implies that the task execution is shared between the direct control of the user and some degrees of autonomy that relies on the robot's local sensory information [10]. An example using a medical scenario can be the autonomous part compensating for the patient's physiological movement, while the surgeon performs the necessary operation on a still patient. A different scenario might involve mobile robots, in which the user directly controls the desired state of a mobile agent, e.g., velocity or position, while the robot works to achieve parallel goals, such as a specific format for shape or obstacle avoidance procedures. In both of these examples, force feedback can close the control loop and significantly improve teleoperation performance: in the former, the benefit is from increasing the manipulation accuracy and decreasing the completion time and the peak and mean force exerted on the remote environment [11]; in the latter, it is from providing information back to the operator about the mismatch between commanded and actual motion.

Virtual fixtures can be considered a special application of shared control [12]. Virtual elements, such as guidance information and any general virtual element are superimposed on the master side, either from a visual or haptic point of view. These fixtures can improve task execution performance by shaping the movement of the operator along a desired path, that can be known because of prior knowledge of the environment or somehow generated online.

B. Shared perception

Although force feedback in telerobotic systems enhances telepresence, it can impose severe restrictions on system stability and performance [13]. Moreover, most bilateral solutions present in the literature require complex and ad hoc controllers together with different sorts of haptic actuators that the human is equipped with, which might constrain the workspace and limit the operator's motion.

This work proposes a configuration of human-robot collaboration in which the robot is sensory-augmented by means of a set of tactile signals coming from the human operator (see Fig. 1). The incorporation of low-level robot "intelligence" the cooperative manipulation of an object while enabling the human operator to stay focused on the task itself and carry it out in the most natural way.

In this paper we present a feasibility study. First, we test and analyze different positions of a tiny sensor, placing it on different parts of the human hand and on the co-manipulated object. Then, we present a use case -a drawing task- to prove the effectiveness of our approach. When drawing a line on a piece of paper, we typically need two hands: one uses the pen the other holds the sheet in place. What happens if one hand is not available for some reason, for instance because it is holding a different tool? The drawing action might result in only a movement of the sheet with no pen strokes, or with a pen stroke of a length or shape different from the desired one.

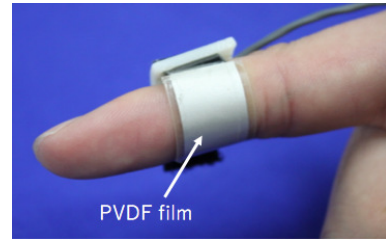


Fig. 2. Wearable vibration sensor.

In a cooperative task, the human can share with the robot his/her haptic sensation, i.e., augmenting the robot sensory experience with the operator's perception, and the robot can react to the received information by applying sufficient force to keep the piece of paper still, i.e., playing the role of the human hand is occupied. It is very important that the robot be aware of the maximum force it can apply during a joint object manipulation, because the object material might be unknown and fragile. Moreover, if the robot knows when it is necessary to apply a force, it also knows when to apply no force at all to give the operator the possibility to remote the object with no undesired impediments.

II. WEARABLE VIBRATION SENSOR

The novelty of this paper lies in the use of a single tiny sensor capable of collecting enough information to command the motion of a robot and accomplish a cooperative manipulation task. In this study, we aim to demonstrate that high-frequency vibration signals are informative enough to set up a simple cooperative human-robot interaction framework. Indeed, vibrations are generated by any mechanical interaction between two agents and also propagate far from the spot where the contact occurs. This implies a valuable advantage in cooperative human-robot interaction tasks, because it might not be strictly required to know all the contact position. Moreover, the same sensor can gather aggregated data coming from different contact sites, thus decreasing the number of sensors required. Tanaka et al. [14] used accelerometers, which detect propagated vibrations, in order to evaluate individual tactile sensations. Similarly, Shao et al. [15], investigated the propagation patterns of cutaneous vibration in the hand during interactions with touched objects by means of an array of accelerometers.

Although wearable tactile sensors equipped with accelerometers can be light, cheap, and simple, they place strict constraints on the speed of exploration, because together with skin vibrations, the sensor output includes the motion of body parts. In the proposed study, we utilized the wearable skin vibration sensor developed by Tanaka et al. [16], shown in Fig. 2. It takes advantage of a polyvinylidene fluoride (PVDF) film, known to be very flexible and light. One side of the PVDF film is attached to a rigid base which is strapped to the finger. It requires no power supply since the sensor itself generates an electrical charge resulting from an applied mechanical force.

As shown in Fig. 2, owing to its design, the sensor can also

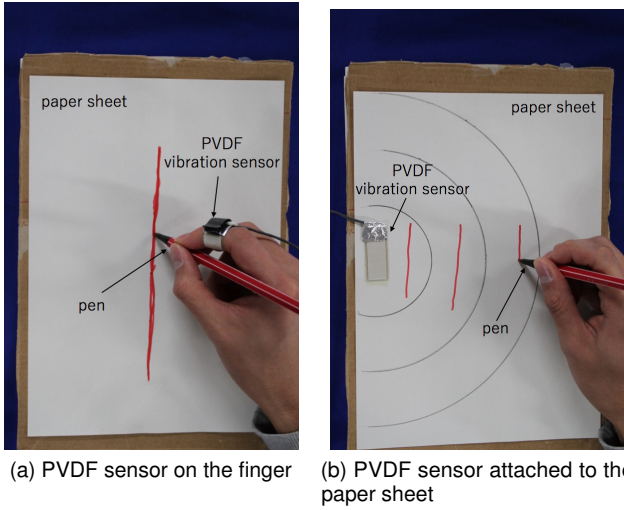


Fig. 3. Experiment #1: experimental setup. a) sensor worn on the user finger. In this specific case, the sensor is worn on the middle phalanx of the index finger. b) sensor attached to the paper sheet. In both cases, the piece of paper can be either fixed to the surface (“writing” user action) or moveable (“slipping” user action).

be easily worn on a finger phalanx to detect skin-propagated vibrations.

III. EXPERIMENTAL EVALUATION

In order to characterize the usability of a wearable vibration sensor in the considered use case, i.e., drawing a line on a piece of paper while a robot holds the sheet only in case it slides, we conducted an experiment to evaluate the performance of the tiny sensor in discerning slightly different vibrations when placed in various positions.

Once the best position of the sensor in terms of discerning performance was found, we exploited it in a paradigmatic application involving a cooperative manipulation task. It is shown how the motion of the robot can be commanded by the information collected by the sensor to help the human operator successfully accomplish the proposed writing task.

A. Experiment #1 : evaluating different sensor positions

The first experiment aimed at investigating how the position of the sensor and the distance from the pen stroke affect sensor readings and the discerning capability of users actions. Indeed, the final goal was to classify online the sensor output in order to discern three user actions: i) the paper sheet is still and the motion of the pen tip properly draws a line on it (“writing” action); ii) the paper sheet has motion because of a greater friction between the pen tip and the piece of paper than between the piece of paper and the underlying surface, and no pen mark is drawn (“slipping” action); iii) the pen tip is not in contact with the paper sheet (“no contact” action).

a) Participants: Six subjects (five males, one female, average age 23) participated in this study. None of the participants reported any deficiencies in their visual or haptic perception abilities. Participants were briefed about all the tasks and afterwards signed an informed consent form, including a declaration of having no conflicts of

interest. All were able to give the consent autonomously. Participation in the experiment did not involve the processing of genetic information or personal data (e.g., health, sexual, lifestyle, ethnicity, political opinion, religious or philosophical conviction). The experiment was approved by the Ethics Committee of Nagoya Institute of Technology.

b) Experimental setup: The setup comprised the PVDF vibration sensor, which has been already detailed in Sec. II and of a piece of paper and a pen to simulate a very weak interaction in terms of the force generated between a human operator and an object, i.e., the writing task. The vibration sensor was connected to an isolated, multifunction DAQ device, a USB-6218 by National Instrument (Texas, USA), and it was in charge of recording any interaction occurring on the object and transmitting it to the robot through serial communication.

c) Methods: Two different main positions of the sensor were evaluated for each user action mentioned above: i) the sensor was worn on the user finger; ii) the sensor was attached to the piece of paper on which the user was to write on. Fig. 3 shows the experimental setup for these two cases. Regarding the sensor worn on the finger, we tested it when placed on the distal, middle, and proximal phalanges of the index and middle fingers and on the distal and proximal phalanges of the thumb. Regarding the sensor attached to the piece of paper, three different distances were tested. The semicircles in Fig. 3b show the three ranges that participants were to write into. Their radii, representing the maximum distance from the sensor, were 4 cm (#1), 8 cm (#2), and 12 cm (#3).

In the first condition, the sensor was worn on the finger and the user was asked to draw a straight line on the piece of paper five times for any interaction action (“writing”, “slipping”, or “no contact”). A total of 120 different pen strokes were drawn on the paper per participant. It should be noted that no sensor was attached to the paper during this part of the experiment. In the second condition, the sensor was placed on the paper sheet in the center of all reference semicircles. The same user was asked again to draw a line five times within each range delimited by two adjacent semicircles. This procedure was repeated for any writing action (“writing”, “slipping”, or “no contact”), for a total of 45 different pen strokes.

While performing the “writing” action, the piece of paper was stuck to the surface in order to keep it still, but when the “slipping” action was required, the sheet was left free to move. The surface that the piece of paper was lying on had low friction, and thus the friction between the pen tip and the sheet always caused the piece of paper to slip. For the “no contact” user action, the participant was asked to keep the pen in the air without it coming in contact with the piece of paper. This gave us information about sensor noise when it was worn on the finger or attached to the sheet.

d) Data evaluation: The PVDF sensor output was collected throughout the experimental procedure. For any contact trial, data acquisition started when the participant made contact with the piece of paper and ended when the tip of the pen was pulled up. In the case of the “no contact” user action, data acquisition was manually triggered and

stopped by the experimenter. All data were high-pass filtered at 100 Hz.

For each trial, we took into account sensor output only for the central interval with a duration of 1 s to avoid possible initial and final interaction transients. Although participants were not instructed regarding the speed of their writing action, they were asked to continue the action for at least 2 s. This time interval was further divided into five parts of 0.2 s for a total of 25 samples each [sensor position - user action] condition. For all the samples we computed the fast Fourier transform (FFT) and the power spectral density (PSD), aiming to determine at which frequencies most of the energy is concentrated and to relate different user actions in the time domain with a specific behavior on the frequency domain. Fig. 4 shows the vibrations gathered by the sensor worn by a representative participant on the middle phalanx of the index finger. In the top panel is shown a complete acquisition of the sensor output. The considered time frame is that when the pen tip was in contact with the paper sheet. The central 1 s time interval depicted in blue represents the portion of the signal we took into account during our analysis, while the red parts indicate a single sample of 0.2 s, which is enlarged in the middle figure panel, and is expressed in volt. The bottom figure shows the PSD of this sample.

In analyzing all the collected data, we attempted to find some parameters that would allow us to discern different user actions according to the acquired vibrations. First, we checked how vibration intensity varies according to the user action and distance between the tip of the pen and the PVDF sensor. Thus, we computed the root mean square (RMS) of PSD of each sample as follows

$$i_{RMS} = \log \sqrt{\int_{f_1}^{f_2} PSD(f) df}, \quad (1)$$

where $f_1 = 100$ Hz and $f_2 = 1000$ Hz. We did not consider data below 100 Hz, because for such a low frequency minute body motions and heart beat significantly affect vibrations measures, while for frequencies above > 1000 Hz the sensor output does not change if different user actions are performed. In general, the intensity measure represents a relevant discerning parameter for detecting the “no contact” user action.

Furthermore, we found the following ratio to be very interesting:

$$r = \frac{A}{B}, \quad (2)$$

where A is the log(RMS) of PSD in the range [850, 1000 Hz] and B is the log(RMS) of PSD in the range $[F_{peak} \pm 75$ Hz]. F_{peak} is the frequency at which PSD is at maximum. A and B frequency ranges are shown in the bottom portion of Fig. 4 by magenta and green patches, respectively. In several pilot tests, we noticed that during the execution of the task with the sensor worn on the finger and while performing the “slipping” user action, the PSD of the sensor output was higher at about 200 Hz, identified by the F_{peak} for any participant, and lower at about 900 Hz. Meanwhile, such a difference was

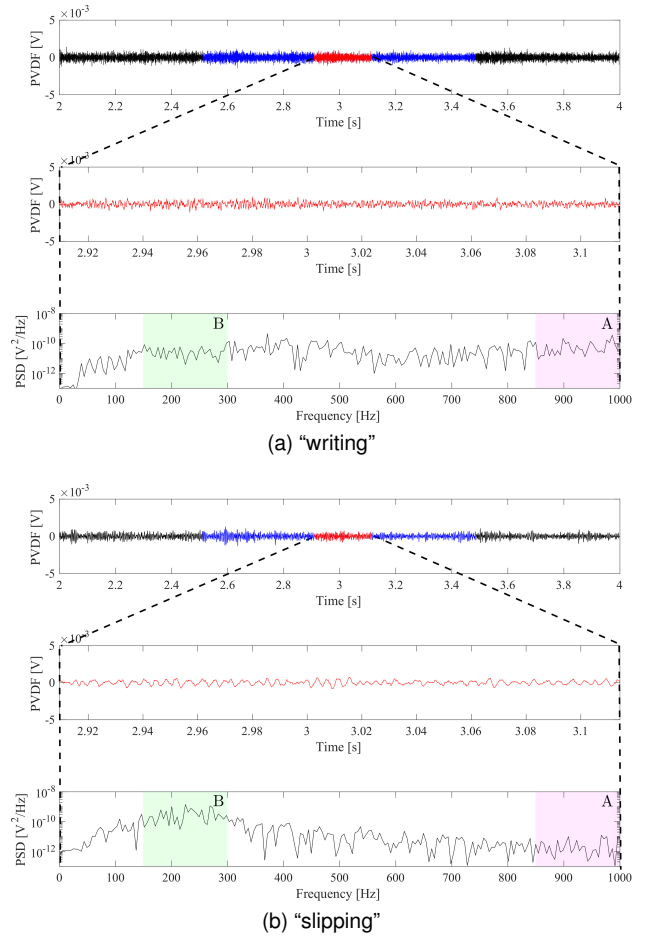


Fig. 4. Vibration data gathered by the sensor worn on the middle phalanx of the index finger of a representative participant. Top panel; the whole data acquisition during a single trial. The part in blue indicates the considered 1 s interval. Middle panel; a portion of data from a duration of 0.2 s is shown (single sample). Bottom panel; the sensor output of the middle plot has been transformed in the frequency domain and showed as power spectral density (PSD). The two colored patches represent the frequency ranges considered to compute the ratio r expressed in Equation 2.

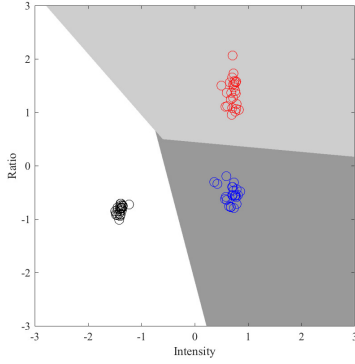
not evident while performing the “writing” user action. Thus, the ratio r plays a valuable role as a discerning parameter for distinguishing the “slipping” from “writing” user actions.

After features extraction, we use a support vector machine (SVM) [17] to define linear separators between the different user actions (one-versus-rest classification), such that the margin between the hyperplane and the points closest to it became maximal.

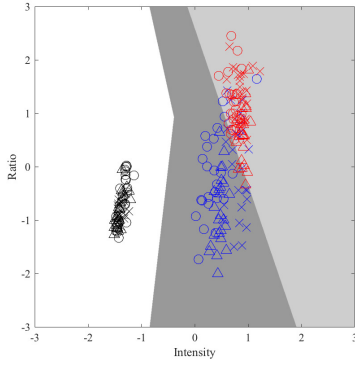
SVMs are a standard first approach before attempting more complex methods in the field of machine learning. In this experiment, we tested only the linear separator. Although the kernel method [18], may enhance the accuracy, we tried simple machine learning as the first step.

i_{RMS} and r have been used off-line to train the SVM. For SVM implementation, we used the LIBSVM integrated software [19].

e) Result: Fig. 5 shows the results of a three-class classifier on data collected for a representative participant. In the top panel (Fig. 5a) the sensor was worn on the user’s



(a) Sensor on the index finger (middle phalanx)



(b) Sensor on the paper sheet (position all)

Fig. 5. The support vector machine class boundary maximizes the margin separating three classes of data collected by a representative user. Symbols in black, red, and blue indicate “no contact”, “slipping”, and “writing” user actions, respectively. Cross, triangle, and circle symbols indicate the three considered distance ranges between the pen tip and the sensor, i.e., 0-4 cm, 4-8 cm, 8-12 cm, respectively.

index finger; in the bottom (Fig. 5b) the sensor was placed on the paper sheet.

Different user actions are distinguished by color (blue for “writing”, red for “slipping”, black for “no contact”) in Fig. 5a and the different distances between the sensor and the contact point are distinguished by symbols (cross for the closest position (#1), triangle for intermediate position (#2), circle for furthest position (#3)) in Fig. 5b. Qualitatively speaking, it is clear how the probability of the SVM for being correct while classifying user actions with the sensor placed on the finger is higher than when the sensor is attached to the paper sheet. As shown in Fig. 5a both the extracted features, r and i_{RMS} , play a crucial role in the discrimination process, in particular when discriminating the “writing” user action from the “slipping” user action; in Fig. 5b, data density appears much more scattered and the intensity parameter appears to be of primary importance.

We believe that because of the stochastic nature of the neural network training, better experimental design methods are necessary to develop its performance measures. Fig. 6 quantitatively shows the SVM performance in any tested condition. In particular, Fig. 6a depicts the results of a

stratified k-fold cross-validation, a popular model validation technique for assessing how the results of a statistical analysis generalizes to an independent data set. In our case, the entire data set is divided into five mutually exclusive folds with approximately the same class distribution as the original data set (stratified). Each fold is used once to test the performance of the classifier that is generated from the combined data of the remaining four folds, leading to five independent performance estimates. The average of the accuracy results of each user action among all the participants is shown. Overall, the accuracy in discriminating the three user actions is very high. As shown in Fig. 5, conditions in which the sensor is worn on the finger seem to slightly outperform conditions in which the sensor is attached to the paper sheet (except for the thumb proximal). An ANOVA with the aligned rank transform showed a influence of the sensor position ($F(11, 175) = 3.1105$, $p = 7.5 \times 10^{-4}$) and user action ($F(2, 175) = 49.4962$, $p = 2.22 \times 10^{-16}$). There are also interactions between sensor position and user action ($F(22, 175) = 2.0595$, $p = 5.3 \times 10^{-3}$) [20]. The accuracy of the “no contact” state is very high in all conditions. Comparing wearable and paper conditions, it seems that the wearable condition has higher accuracy in the “writing” state. The condition of the sensor placed on the proximal phalanx of the thumb has low accuracy. This might depend on the specific task proposed while collecting data; indeed, in a writing task the thumb plays a primary importance role in exerting the force on the pen, and such an interaction might affect the perception of vibration due to the contact between the pen tip and the paper sheet. Fig. 6b shows the distribution of the variance of ratio and variance of intensity on each condition. The variance of each feature was calculated from the data on each user action, and mean variance was calculated. The results show that the variance for the conditions of the sensor on the finger is smaller in both ratio and intensity than that for the condition of the sensor on the paper, and in particular, the variance of ratio for the condition of the finger is much smaller. Small variance indicates a robust localization of the parameter in the SVM space.

f) Discussion: Regarding the condition of the sensor on the finger, when writing, vibration is induced through the mechanical interaction of the paper and the pen. In contrast, when the paper slips, vibration is induced through the mechanical interaction of the paper and the table. Thus, it seems that there is a difference in the characteristic of PSD in addition to the intensity. Here, the slightly lower accuracy that resulted in the condition of the thumb proximal might be influenced by posture and motor control of the thumb during writing. When the sensor was placed on the paper, vibration was caused by the paper moving; thus, the user’s actions seem to be discriminated by intensity, not by the ratio. As shown in Fig. 5b, the intensity of the slipping state is greater than that of the writing state. But the intensity changed according to the distance between the sensor and the pen. In the case of placing the sensor in the closer position, the intensity of the writing state became similar to that of the slipping state, so it became difficult to discriminate writing and slipping state

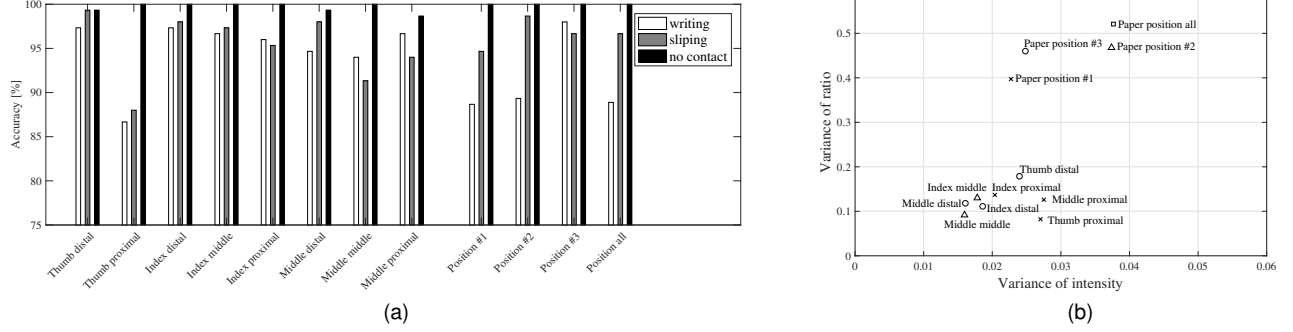


Fig. 6. SVM performance evaluation. a) k-fold cross-validation test results. SVM prediction performance are shown for any sensor position tested. For each position, the accuracy rate of each user action is shown. On the rightmost part of the figure the aggregated information on the overall accuracy for all positions of the sensor on the paper is reported. b) Data density information expressed as the variance of collected data split with respect to any tested sensor position. The aggregated information on the overall variance for all positions of the sensor on the paper is included.

and the accuracy was lower than in the furthest position. The intensity also varied depending on the writing pressure. For example, with high-pressure writing, the intensity became greater. Therefore, the ratio for discrimination is more useful than intensity. From these results, it is considered that the condition of placing the sensor on a human's fingers is most suitable for objective cooperative work. This also has the advantage that we do not have to place the sensor on every paper sheet or robot. In a writing task, the actual phenomenon is occurring in the human hand. Therefore, direct human haptic information is considered to be the most useful.

B. Shared human perception in a paradigmatic application

In this section, we show how the PVDF sensor can be exploited in a paradigmatic application, in which the robot moves according to what a human operator perceives in order to cooperatively achieve the final goal.

a) *Participants:* Two men participated in this study. None of the participants reported any deficiencies in their visual or haptic perception abilities. Participants were briefed on all the tasks and afterwards signed an informed consent form, including a declaration of having no conflicts of interest. Both were able to give this consent autonomously. Participation in the experiment did not involve the processing of genetic information or personal data (e.g., health, sexual, lifestyle, ethnicity, political opinion, religious or philosophical convictions). The experiment was approved by the Ethics Committee of the Nagoya Institute of Technology.

b) *Experimental setup:* An Omega.3 haptic interface (Force Dimension, CH) represented the robot in charge of cooperating with the human operator during the execution of the writing task. A rubber patch was attached to the bottom of the Omega.3 end-effector to increase the friction between the robot and the paper sheet when in contact. Besides the presence of the robot, the experimental setup was very similar to that for experiment #1 described in Sec. III-A: it consisted of a PVDF vibration sensor, a multifunction DAQ device, and a paper sheet. An HD camera and a ruler were used to monitor and quantify the motion of the paper sheet due

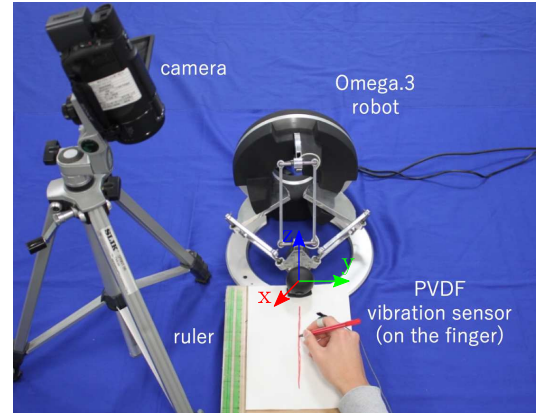


Fig. 7. Paradigmatic application: experimental setup.

to the interaction with the pen tip. The paper sheet was positioned in such a way as to guarantee contact with the robot when required; i.e., the top part of the sheet was under the robot end-effector. The initial position of the robot end-effector was set at 2 mm above the sheet. The motion of the robot end-effector was constrained along its z-axis. The main components of the experimental setup are shown in Fig. 7.

c) *Methods:* In this paradigmatic application, participants wore the PVDF sensor on the middle phalanx of the index finger and wrote a straight line on the paper sheet three times with no specific indication for the position.

Based on the online classification of the PVDF sensor output using the SVM model as presented in Sec. III-A, the robot was able to move according to what was happening during the writing task. The online classification, as well as any single robot control input, was performed at a rate of 50 Hz. Three different robot motions were related to the three possible user actions. When the “no contact” user action was recognized, the robot end-effector progressively returned to the initial position (if it was not already there), moving upwards (along the z-axis) at 0.1 mm for each classifier iteration. When the “slipping” user action was recognized, the robot end-effector moved downwards applying a force

growing at 0.05 N for each iteration until the paper slipping no longer occurred, i.e., the “no contact” or “writing” user actions were detected. When the “writing” user action was recognized, the robot end-effector applied the force computed during the previous iteration to keep the paper sheet still, but at the same time avoiding applying too high a force.

d) Data evaluation: The PVDF sensor output was collected throughout the writing task execution and was processed by the SVM to determine the state. A sliding window of 0.2 s was used for this discrimination, and this sliding window was shifted every 0.02 s. In this way, the robot could understand the current user action at every 0.02 s.

Three subsequent identical classifications were necessary to change the current behavior of the robot. We are aware that this artifact decreases the robot response time, but it makes it possible to overcome almost any miss by the SVM in classifying a user action. It is worth pointing out that a writing task does not require a high movement speed. The HD camera recorded the workspace while the task was being carried out. This video was used to make a rough evaluation of the displacement of the paper sheet during the “slipping” action. Moreover, all the forces exerted by the Omega.3 haptic interface were logged during the task execution.

e) Result: Fig. 8 illustrates how our sensory-augmented robot behaves to accomplish a paradigmatic cooperative task with a representative participant. Collected data regarding a whole writing task (drawing a line three times) are reported. For the sake of clarity and space, we neglected the initial time interval (< 10 s) during which there was no contact between the pen and the paper sheet. In the top panel, it is shown how the trained SVM classifies vibrations coming from the PVDF sensor as the three different user actions: “no contact”, “writing”, or “slipping”. As we see in the middle panel, the Omega.3 changes its behavior according to the evaluated user action. If there is no contact between the pen tip and the paper (black line), it attempts to apply a null force. This means that if its end-effector applies a force ($t \approx 14.5$ s), it moves up to the initial position (no contact between the robot and the paper sheet is present). When a “writing” user action is detected, the robotic end-effector keeps on applying the last force (blue line); this means that the force can be null or not null. In this case, the participant does not need the robot to change its behavior because it is correctly writing the requested line. In case the paper sheet slips owing to the interaction forces with the pen tip, the participant is not able to write the requested line and the robot attempts to help by applying an increasing force (red line) until the sheet is still. It is worth noting that the applied force is not always the same, but it is the minimum necessary to stop the sheet slippage. Thus, by estimating the intensity of the user action, the robot reacts accordingly. In the middle figure panel, the estimation of the sheet displacement is shown in order to prove the correct action discrimination performed by the SVM. Three pink patches highlight the time intervals in which the undesired motion of the paper sheet occurs.

f) Discussion: In this study, the proposed system relies on task-dependent training of a SVM that prevents the

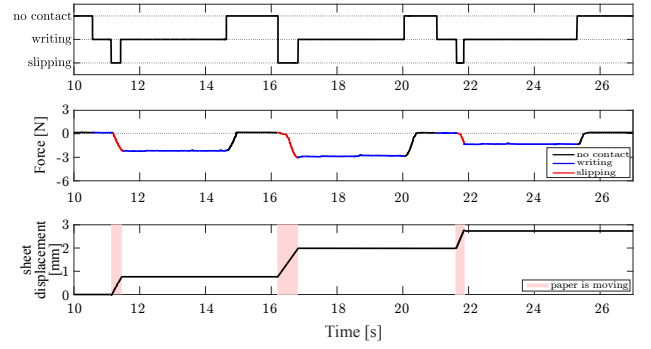


Fig. 8. How our sensory augmented robot behaves to accomplish the paradigmatic cooperative task with a representative participant. Top panel: how the SVM classifies vibrations coming from the PVDF sensor. Middle panel: the force profile commanded to the Omega.3 haptic interface to stabilize the motion of the paper sheet when necessary. Bottom panel: estimation of the sheet displacement. Pink patches highlight the time interval in which the undesired motion of the paper sheet occurs.

operator from quickly moving between various tasks. If we consider the presented paradigmatic application in detail, discrimination performance would drastically decrease if the material of either the sheet or the underlying surface was significantly different. Work is in progress to design a fast calibration procedure capable of training the SVM with a few tests on the objects involved in the task. This paper should be seen as a preliminary feasibility study on how a robot can be sensorimotor-augmented by a shared human perception. The considered use-case proves the effectiveness of our approach and can be easily reshaped for various industrial and medical applications, such as a surgical cutting task. Further research on the sensor output and on the discrimination parameters that are obtained is required to improve the current performance, in particular with an increased number of user actions to discriminate. In addition, the time necessary to discriminate different actions determines the response speed of the system. For instance, PSD requires a time series to obtain meaningful results. When we deal with a “slow” task, such as the proposed writing task, we have no problems in response speed, but if we have to discriminate highly dynamic interactions, the PSD method might not work properly. Improving the accuracy of discrimination and robustness is also necessary. The use of more sensors and the introduction of different analysis methods, such as more sophisticated kernel methods of SVM, would be of primary importance.

It is also important to allow the operator to carry out the task in more natural way. We are considering making it possible for operators to communicate with the robot. Thereby, operators can intuitively understand what the robot is doing. Haptic feedback and auditory feedback would be useful for this approach.

IV. CONCLUSION AND FUTURE WORK

In most human-robot cooperation approaches, all the agents have a direct perception of what is happening in the shared workspace through the interaction of a co-manipulated object. In this work, we propose a human-in-the-loop framework in

which robot behavior is directly based on real-time human sensorimotor information.

A tiny PVDF sensor was used to discriminate user's actions in a paradigmatic application: drawing a line on a piece of paper using a single hand. A qualitative point of view, the sensor placed on the fingers appears to slightly outperform the same sensor attached to the co-manipulated object (see Fig. 5). Sensor on the finger directly represents the operator's perception, while the output of the sensor attached to the object changes according to the distance between the interaction point and the sensor itself. In addition, by instrumenting only the operator's hand, we enable her/him to carry out human-robot collaborative tasks on any available object. If we assume wireless communication and no compatibility issues, the operator might also move around an industrial workspace equipped with several robots and perform several cooperative tasks involving different machines. Neither the robot nor the co-manipulated objects need something specific; any robot can take advantage of the operator's perception shared via a wearable sensor. Finally, the integration of multiple types of wearable sensors would allow a robot to exploit many kinds of information and improve human-robot coordination in the cooperative execution of more complex and varied tasks.

REFERENCES

- [1] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: a survey," *Foundations and trends in human-computer interaction*, vol. 1, no. 3, pp. 203–275, 2007.
- [2] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: An historical survey," *Automatica*, vol. 42, no. 12, pp. 2035–2057, 2006.
- [3] G. Salvietti, L. Meli, G. Gioioso, M. Malvezzi, and D. Prattichizzo, "Multi-contact bilateral telemanipulation with kinematic asymmetries," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 445–456, 2017.
- [4] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi, "An atlas of physical human–robot interaction," *Mechanism and Machine Theory*, vol. 43, no. 3, pp. 253–270, 2008.
- [5] J. Borenstein and R. C. Arkin, "Robots, ethics, and intimacy: the need for scientific research," in *Proc. International Association for Computing and Philosophy*, 2016.
- [6] A. Weiss, R. Buchner, M. Tscheligi, and H. Fischer, "Exploring human-robot cooperation possibilities for semiconductor manufacturing," in *Proc. IEEE Int. Conf. on Collaboration Technologies and Systems*, pp. 173–177, 2011.
- [7] O. Khatib, K. Yokoi, O. Brock, K. Chang, and A. Casal, "Robots in human environments: Basic autonomous capabilities," *International Journal of Robotics Research*, vol. 18, no. 7, pp. 684–696, 1999.
- [8] L. Conway, R. Volz, and M. Walker, "Tele-autonomous systems: Methods and architectures for intermingling autonomous and telerobotic technology," in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 4, pp. 1121–1130, 1987.
- [9] S. Hayati and S. Venkataraman, "Design and implementation of a robot control system with traded and shared control capability," in *Proc. IEEE Int. Conf. on Robotics and Automation*, pp. 1310–1315, 1989.
- [10] G. Niemeyer, C. Preusche, and G. Hirzinger, "Telerobotics," in *Springer handbook of robotics*, pp. 741–757, Springer, 2008.
- [11] L. Meli, C. Pacchierotti, and D. Prattichizzo, "Experimental evaluation of magnified haptic feedback for robot-assisted needle insertion and palpation," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 13, pp. 1809–1809, 2017.
- [12] S. A. Bowyer, B. L. Davies, and F. R. y Baena, "Active constraints/virtual fixtures: A survey," *IEEE Trans. on Robotics*, vol. 30, no. 1, pp. 138–157, 2014.
- [13] T. B. Sheridan, *Telerobotics, automation, and human supervisory control*. MIT press, 1992.
- [14] Y. Tanaka, Y. Horita, and A. Sano, "Finger-mounted skin vibration sensor for active touch," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pp. 169–174, 2012.
- [15] Y. Shao, V. Hayward, and Y. Visell, "Spatial patterns of cutaneous vibration during whole-hand haptic interactions," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 4188–4193, 2016.
- [16] Y. Tanaka, D. P. Nguyen, T. Fukuda, and A. Sano, "Wearable skin vibration sensor using a pvdf film," in *Proc. IEEE World Haptics Conference*, pp. 146–151, 2015.
- [17] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [18] A. J. Smola and B. Schölkopf, *Learning with kernels*, vol. 4. Citeseer, 1998.
- [19] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, p. 27, 2011.
- [20] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, "The aligned rank transform for nonparametric factorial analyses using only anova procedures," in *Proc. SIGCHI Human Factors in Computing Systems conference*, pp. 143–146, May 2011.